University of Warsaw
Faculty of Mathematics, Informatics and Mechanics

Paweł Pasteczka

Analytic methods in inequalities concerning means

PhD dissertation

Supervisor
dr hab. Piotr Mormul

Institute of Mathematics
University of Warsaw

April 2015

Author's declaration:
aware of legal responsibility I hereby declare that I have written this dissertation myself and all the contents of the dissertation have been obtained by legal means.

April 17, 2015 . . . .

date *Paweł Pasteczka*

Supervisor's declaration:
the dissertation is ready to be reviewed

April 17, 2015 . . . .

date *dr hab. Piotr Mormul*

Analytic methods in inequalities concerning means

Undoubtedly the most popular family of means used in mathematical analysis, statistics, probability and other branches of mathematics are Power Means. In the late 1920's and in the beginning on 1930's Kolmogorov, Nagumo and de Finetti, independently, proposed the new family being the generalization of this family – currently adopted for the name of quasi-arithmetic means. These means are defined by the equality $f^{-1}(\sum f(a_i)/n)$, where f is a continuous, strictly monotone function defined on the interval, while $(a_i)_{i=1}^n$ is a vector of arguments. For such objects, there naturally appear a whole list of questions regarding the adaptation of the classical results known for Power Means.

An example of such a problem is to adopt the classical fact, well-known for Power Means, claiming that for any fixed vector of arguments, as the parameter change among all possible arguments, one obtains (exactly once) all the intermediate values between the smallest and the largest component of the vector. In my thesis I make an attempt to resolve, using - it seems - quite advanced methods, a question when a family of quasi-arithmetic means has this property (so-called scale property).

Another important issue is the question how does a small change of the function f affecting the value of quasi-arithmetic mean generated by f. Some results in this area were obtained already in the 1960s by Cargo and Shisha (however, some additional conditions concerning regularity were done). The problem of finding necessary and sufficient conditions for convergence in the family of quasi-arithmetic means (not giving any estimate of the distance) was solved by Pàles in the late 1980's. My results provide new estimates referring to the result of Cargo and Shisha and, at the same time, generalizing Pàles'es result.

Another class of problems studied in my dissertation is a list of questions related to the Hardy means. Its history was started by Hardy's result from 1920 - the answer to a previous Hilbert's question from 1909. Hardy proved that if $\mathcal{P}_p$ is a p-th order power mean, $p \in (0,1)$ and $(a_i)_{i=1}^{\infty} \in l^1(\mathbb{R}_+)$ then $\sum_{n=1}^{\infty} \mathcal{P}_p(a_1,\dots,a_n) < (p-p^2)^{-1/p} \sum_{n=1}^{\infty} a_n$. (One year later, Landau obtained a result with the optimal constant at the right hand side.) This was, however, only a starting point for further research - currently a mean M is called *Hardy* if there exists a constant $C > 0$ such that

$$\sum_{n=1}^{\infty} M(a_1,\dots,a_n) < C \sum_{n=1}^{\infty} a_n \text{ for any sequence } a \in l^1(\mathbb{R}_+).$$

A natural question is whether some particular mean is Hardy. In the present thesis, I an going to prove this property for several families of means, as well as give a lot of negative results regarding having Hardy property. Among the already obtained results are the necessary and sufficient condition for a family which is a generalization of the arithmetic-geometric mean, considered by Gauss, and solution of hypotheses established in 2004 by Páles and Persson.

Keywords:
quasi-arithmetic mean, generalized mean, scale of means, mean, inequalities, metric, Arrow-Pratt index, Hardy means, differences among means, Gini means, Gaussian product of Power Means, generalized Power Means

AMS classification:
Primary: 26E60
Secondary: 26D15, 26D07, 47A63, 47A64

Analityczne metody w nierównościach dotyczących średnich

Niewątpliwie najbardziej popularnymi średnimi używanymi w analizie matematycznej, statystyce, rachunku prawdopodobieństwa oraz innych działach matematyki są średnie potęgowe. Na przełomie lat 20-tych i 30-tych XX wieku niezależnie Kołmogorow, Nagumo oraz de Finetti wpadli na pomysł nowych średnich będących daleko idącym uogólnieniem średnich potęgowych. Obecnie przyjęła się dla nich nazwa średnich quasi-arytmetycznych. Są to średnie postaci $f^{-1}(\sum f(a_i)/n)$, gdzie f jest funkcją określoną na odcinku, ciągłą, ściśle monotoniczną, natomiast $(a_i)_{i=1}^{n}$ jest wektorem argumentów. Dla tak zdefiniowanych wielkości pojawia się cała lista pytań dotyczących przenoszenia klasycznych wyników znanych dla średnich potęgowych.

Przykładem takiego problemu jest przenoszenie klasycznego faktu znanego dla średnich potęgowych, mówiącego, że dla dowolnego ustalonego wektora argumentów, przy przebieganiu parametrem wszystkich wartości rzeczywistych otrzymujemy (dokładnie raz) wszystkie wartości pośrednie pomiędzy najmniejszą i największą składową wektora. W moim doktoracie podejmuję próbę rozstrzygnięcia, przy użyciu – wydaje się – dość zaawansowanych metod, kiedy dana rodzina średnich quasi-arytmetycznych posiada wymienioną własność (tzw. własność skali).

Kolejnym kluczowym zagadnieniem jest pytanie, w jaki sposób zmiany funkcji f wpływają na zmianę wartości średniej quasi-arytmetycznej pochodzącej od f. Pewne wyniki w tym zakresie były uzyskiwane już w latach 1960-tych (przy pewnych dodatkowych warunkach dot. regularności) przez Cargo oraz Shishę. Problem znalezienia warunków koniecznych i dostatecznych dla zbieżności w rodzinie średnich quasi-arytmetycznych (niedający jakiegokolwiek oszacowania odległości) został rozwiązany przez Pálesa pod koniec lat 1980-tych. Moje dotychczas uzyskane rezultaty dają nowe oszacowania nawiązujące do prac Cargo oraz Shishy i równocześnie uogólniające wyniki Pálesa.

Inną klasą problemów badaną w mojej pracy doktorskiej jest lista pytań związanych ze średnimi Hardy'ego. Początkiem był tu rezultat Hardy'ego z roku 1920 - odpowiedź na wcześniejsze pytanie Hilberta z roku 1909. Hardy pokazał, że jeśli $\mathcal{P}_p$ jest średnią potęgową rzędu p to, gdy $p \in (0,1)$ oraz $(a_i)_{i=1}^{\infty} \in l^1(\mathbb{R}_+)$ wtedy ma miejsce nierówność $\sum_{n=1}^{\infty} \mathcal{P}_p(a_1,\dots,a_n) < (p-p^2)^{-1/p} \sum_{n=1}^{\infty} a_n$ (Rok później Landau uzyskał rezultat z lepszą, optymalną stałą po prawej stronie.) Stanowiło to jednakże jedynie punkt wyjścia do dalszych badań – aktualnie średnią M nazywamy Hardy'ego jeśli istnieje stała $C > 0$ taka, że

$$\sum_{n=1}^{\infty} M(a_1,\dots,a_n) < C \sum_{n=1}^{\infty} a_n \text{ dla dowolnego ciągu } a \in l^1(\mathbb{R}_+).$$

Naturalnym pytaniem jest, czy dana średnia jest Hardy'ego. W pracy udowadniam w/w własności dla kilku rodzin średnich, jak również podaję wiele negatywnych dotyczących własności Hardy'ego. Wśród już uzyskanych wyników są: warunek konieczny i dostateczny dla rodziny będącej uogólnieniem średniej arytmetyczno-geometrycznej rozważanej jeszcze przez Gaussa oraz rozwiązanie hipotezy postawionej w 2004 roku przez Perssona oraz Pálesa.

Słowa kluczowe:
średnia quasi-arytmetyczna, uogólniona średnia, skala średnich, średnia, nierówności, metryka, indeks Arrowa-Pratta, różnice między średnimi, średnie Hardy'ego, średnie Gini'ego, iloczyn gaussa średnich potęgowych, uogólnione średnie potęgowe

Klasyfikacja AMS:
Główna: 26E60
Pomocnicza: 26D15, 26D07, 47A63, 47A64

Contents

Chapter 1

Overview

Presented dissertation consists of seven chapters that was written during my PhD studies at the Faculty of Mathematics, Informatics and Mechanics, University of Warsaw. They are concerning theory of means. First part applies to the quasi-arithmetic means, while the next one to so-called Hardy property.

The idea of quasi-arithmetic means emerged in the beginning of 1930s in three nearly simultaneous papers [12; 25; 31]. They are defined for any continuous, strictly monotone function $f\colon U \to \mathbb{R}$ (U is an interval) by the equality $f^{-1}(\sum w_i f(a_i))$, where $a \in U^n$ are *entries*, $w_i > 0$ satisfying $\sum w_i = 1$ – *weights*. They are a natural generalization of Power Means. Moreover, many known results concerning Power Means could be adapted to the family of quasi-arithmetic means.

In part I, we will prove some estimates of the differences among quasi-arithmetic means. These results are corresponding with earlier papers of G. T. Cargo and O. Shisha [10] and [11]. An important tool will be the earlier idea of Mikusiński and, independently, Łojasiewicz [29, footnote 2].

Chapter 3 was published in Real Analysis Exchange, chapter 4 is currently under referee process .

Next part are concerning so-called Hardy property. In 1920, G. H. Hardy, [20], proved that whenever $p < 1$ and $\mathcal{P}_p$ is a power mean then there exists a constant c_p satisfying

$$\sum_{n=1}^{\infty} \mathcal{P}_p(a_1, \dots, a_n) < c_p \sum_{n=1}^{\infty} a_n \text{ for any } a \in l^1(\mathbb{R}_+).$$

He proved that for $p > 0$ the constant $\tilde{c}_p = (p - p^2)^{-1/p}$ fits. Moreover he had claimed that the constant $c_p = (1 - p)^{-1/p}$ is optimal. His conjecture was confirmed one year later by Landau [27]. Two years later Carleman [8] independently proved that $\sum_{n=1}^{\infty} \mathcal{P}_0(a_1, \dots, a_n) < e \sum_{n=1}^{\infty} a_n$ and the constant e is optimal. In 1928, Knopp [23], proved that the constant $c_p = (1 - p)^{-1/p}$ is optimal for $p < 0$ too. This result fulfilled the problem of optimal constant for power means. [It could be easily verified that for $p \geqslant 1$ power means do not satisfied similar property.]

This property was named in 2004 by Zs. Páles and L.-E. Persson [36]. A mean $\mathfrak{A} \colon \bigcup_{n=1}^{\infty} U^n \to \mathbb{R}_+$, U – an interval, $\inf U = 0$ is *Hardy* if there exists a constant c such that

$$\sum_{n=1}^{\infty} \mathfrak{A}(a_1, \dots, a_n) < c \sum_{n=1}^{\infty} a_n \quad \text{for any } a \in l^1(U).$$

We have already known that $\mathcal{P}_p$ is Hardy if and only if $p < 1$.

Next step in the development of Hardy means was the paper by Mulholland [30]. He proved an if and only if condition for quasi-arithmetic means to be Hardy. It still remains one of the most general result in this theory. Let us note that, however most of presented results are classical, the problem of Hardy property remains open for many families of means. Many results in theory of Hardy means have appeared ever since 1920s, it could be seen in by-now-classical reviews [13; 42; 33] and a recent book [26].

In chapter 6, I am going to characterize Hardy property among three-parameters generalization of power means introduced in 1971 by Carlson, Meany and Nelson [9]. Later, in chapter 7, I am going to completely characterize Hardy property for Gini means [16] (it is a solution of conjecture established by

Pàles and Persson in 2004 [36]) and Gaussian product of power means.

Papers containing pertinent results, [P2; P3] are awaiting publication.

In the following two sections I am going to describe in detail my results concerning quasi-arithmetic and Hardy means, respectively.

1.1 Quasi-arithmetic means

Quasi-arithmetic means were defined in three nearly simultaneous papers [12; 25; 31] in the beginning of 1930s. For a continuous, strictly monotone function $f\colon U \to \mathbb{R}$ (U–an interval) we define a quasi-arithmetic mean by

$$\mathfrak{M}_f(a,w) := f^{-1}\left(\sum_{i=1}^{n} w_i f(a_i)\right),$$

where $a \in U^n$ are *entries*; $w_i > 0$ satisfying $\sum w_i = 1$ – *weights.*

This family of means is a natural generalization of power means. This fact was announced by Knopp in late 1920s [23]. Indeed, this family are so closely related, that the classical proof of inequalities between power means could be easily adopted to quasi-arithmetic means. Namely, let U be an interval, $f,\ g\colon U \to \mathbb{R}$ – continuous, strictly monotone functions. Let us assume without loss of generality that f is an increasing function ($\mathfrak{M}_f = \mathfrak{M}_{-f}$). Then the following conditions are equivalent:
(i) $\mathfrak{M}_f(a,w) \geqslant \mathfrak{M}_g(a,w)$ for any admissible a, w,
(ii) $f \circ g^{-1}$ is convex.
Moreover, having some additional smoothness assumption, by Jensen inequality, one obtains an additional iff condition:
(iii) $A_f(x) \geqslant A_g(x)$ for $x \in I$, where $A_f := f''/f'$.
This condition was significantly generalized by Mikusiński

Theorem ([29])**.** *Let U be an interval, $f,\ g\colon U \to \mathbb{R}$ – twice derivable functions, $f' \cdot g' \neq 0$. The following conditions are* *equivalent:* (i) $\mathfrak{M}_f(a,w) \geqslant \mathfrak{M}_g(a,w)$ *for all admissible vectors* a *and corresponding weights* w*, with both sides equal only if* a *is a constant vector,* (ii) $A_f > A_g$ *on a dense subset of* U*.*

Estimate of the differences among quasi-arithmetic means. In 1960s Cargo and Shisha [10; 11] considered some estimates of the difference among quasi-arithmetic means. They introduced a metric

$$\rho(\mathfrak{M}_f, \mathfrak{M}_g) := \sup\{|\mathfrak{M}_f(a,w) - \mathfrak{M}_g(a,w)| : a,\ w \text{ admissible } \}.$$

Many upper estimates were proved i.e. a classical estimation $\rho(\mathfrak{M}_f, \mathfrak{M}_g) < 2\omega_{f^{-1}}(\|f-g\|_\infty)$ (ω denotes a modulus of continuity).

In chapter 3, using some results from [10; 11], I obtained an upper estimation of the difference for the means generated by twice differentiable functions with nowhere vanishing first derivative. It was expressed in terms of Mikusiński's operator.

For the purpose of the present overview, let me quote, in what follows, from the ulterior chapters 3 and 4.

Theorem (3.3 in chapter 3)**.** *Let* U *be an interval,* $f,\ g \colon U \to \mathbb{R}$ *– twice derivable functions,* $f' \cdot g' \neq 0$*. Then*

$$\rho(\mathfrak{M}_f, \mathfrak{M}_g) \leqslant |U| \exp(2\,\|A_f\|_1) \sinh 2\,\|A_g - A_f\|_1 .$$

Later, as an application to the solution of the problem presented by Zs. Páles [34], this result was strengthened. (It will be proved in chapter 4)

Theorem (4.2 in chapter 3)**.** *Let* U *be an interval,* $f,\ g \colon U \to \mathbb{R}$ *– twice derivable functions,* $f' \cdot g' \neq 0$*. Then*

$$\rho(\mathfrak{M}_f, \mathfrak{M}_g) \leqslant |U| \exp\|A_f\|_* \Big(\exp\|A_f - A_g\|_* - 1\Big),$$

where $\|u\|_* := \sup_{a,b\in U} \left|\int_a^b u(x)dx\right|$.

Moreover, one of theorem in chapter 4 allows us to estimate a differences among quasi-arithmetic means, which are *not* generated by derivable functions. This result was inspired by problem presented in 2013 by Pàles [34]. It is worded in terms of his operator $P_f(x,y,z) := \frac{f(x)-f(y)}{f(x)-f(z)}$ (compare [35]):

Fact (4.1 in chapter 4)**.** *Let U be an interval, f, $g\colon U \to \mathbb{R}$. Let $\Delta_\alpha := \{(x,y,z) \in U^3\colon\ |x-z| \geqslant \alpha\}$. If $\|P_f - P_g\|_{\infty,\Delta_\alpha} < 1$ for some $\alpha > 0$ then $\rho(\mathfrak{M}_f, \mathfrak{M}_g) < \alpha$.*
($\|\cdot\|_{\infty,\Delta_\alpha}$ denotes the standard norm in the space $L^\infty(\Delta_\alpha)$.)

Scale. Some other direction of my research consists in adapting some classical property of power means to quasi-arithmetic means. Namely, for any fixed, non-constant vector of entries and corresponding weights as we consider all possible power means, one obtains all values between minimal and maximal entry of considering vector of entries. In my thesis I'm going to present a very general, however no an iff, result for a subfamily of quasi-arithmetic means to admit this property (*scale property*). In chapter 3, I proved

Theorem (3.1, 3.2)**.** *Let I, U be an open intervals, $(k_\alpha)_{\alpha\in I}$ – family of $\mathcal{C}^2$ functions, $k_\alpha\colon U \to \mathbb{R}$, $k_\alpha \neq 0$, $\alpha \in I$.*

- *If $I \ni \alpha \mapsto A_{k_\alpha}(x)$ if increasing, 1–1 on a dense subset of U and 'onto' for all $x \in U$ then $(\mathfrak{M}_{k_\alpha})_{\alpha\in I}$ is a scale;*
- *If $(\mathfrak{M}_{k_\alpha})_{\alpha\in I}$ is a scale then there exist a dense subset $X \subset U$ such that $I \ni \alpha \mapsto A_{k_\alpha}(x)$ is increasing, 1–1 and 'onto' for any $x \in X$.*

1.2 Hardy means

History of Hardy means, formally introduced ten years ago in [36], emerged in 1920, when Hardy published a Riesz's proof of the following inequality:

$$\sum_{n=1}^{N}\left(\frac{1}{n}\sum_{k=1}^{n}a_k\right)^q \leqslant \left(\frac{q^2}{q-1}\right)^q \sum_{k=1}^{n}a_k^q, \text{ where } q>1 \text{ and } a_k \geqslant 0.$$

By this inequality one could easily obtain

$$\sum_{n=1}^{\infty}\mathcal{P}_p(a_1,\dots,a_n) < (p-p^2)^{-1/p}\sum_{n=1}^{\infty}a_n, \quad p\in(0,1),\ a\in l^1(\mathbb{R}_+).$$

Today we say briefly: $\mathcal{P}_p$ is Hardy for $p<1$. However one year later, [27], it was already known that the constant presented above is not optimal, but in current thesis we will not deal with optimality of constant appearing on the right-hand-side of Hardy inequality. There is also a similar result by Mulholland from 1932 in the same spirit

Theorem ([30]). *Let $U \subset \mathbb{R}_+$ be an interval,* $\inf U = 0$. *Let $f\colon U \to \mathbb{R}$ be a continuous, strictly monotone function. Then $\mathfrak{M}_f$ is Hardy if and only if there exist: numbers $A>1$, $k>1$, and a convex function $\varphi\colon f(I)\to\mathbb{R}$ satisfying*

$$\varphi(y) \leqslant (f^{-1}(y))^{1/k} \leqslant A\cdot\varphi(y) \quad \text{for } y\in f(I).$$

Nevertheless, the problem of being or not being Hardy remains open for many families of means. In this thesis I'm going to present a necessary and sufficient condition to be Hardy for *(i)* Gini means and *(ii)* Gaussian product of power means to be Hardy. Moreover, some advanced consideration for generalized power means will be done.

Generalized power means. In 1971, [9], Carlson, Meany and Nelson introduced the following generalization of power means: For $k\in\mathbb{N}$, $s\,,q\in\mathbb{R}$ one defines

$$\tilde{\mathcal{P}}_{k,s,q}(v_1\dots v_n) := \begin{cases}\mathcal{P}_s\Big(\mathcal{P}_q(v_{i_1},\dots,v_{i_k})\colon 1\leqslant i_1<\dots<i_k\leqslant n\Big) & \text{if } k\leqslant n\,,\\ \mathcal{P}_q(v_1,\dots,v_n) & \text{if } k>n\,.\end{cases}$$

During my studies, Corollary 6.3.1, I proved that, for any fixed $k \in \mathbb{N}$,

- $\tilde{\mathcal{P}}_{k,s,q}$ is Hardy for no $s \geqslant 1$, $q > 0$,
- $\tilde{\mathcal{P}}_{k,s,q}$ is Hardy for no $s \geqslant k$ and $q \in \mathbb{R} \cup \{\pm\infty\}$,
- $\tilde{\mathcal{P}}_{k,1,q}$ is Hardy for $q \leqslant 0$,
- $\tilde{\mathcal{P}}_{k,s,q}$ is Hardy for $s < 1$ and $q \in \mathbb{R} \cup \{\pm\infty\}$.

Gini means. Gini means were proposed in 1938 [16] as a generalization of power means. This family is defined by

$$\mathfrak{G}_{p,q}(a_1, \dots, a_n) := \begin{cases} \left(\frac{\sum_{i=1}^n a_i^p}{\sum_{i=1}^n a_i^q} \right)^{1/(p-q)} & \text{if } p \neq q\,, \\ \exp\left(\frac{\sum_{i=1}^n a_i^p \ln a_i}{\sum_{i=1}^n a_i^p} \right) & \text{if } p = q\,. \end{cases}$$

Indeed, it could be easily verified that upon taking $q = 0$ one obtains the p-th power mean. Zs. Páles and L.-E. Persson [36] proved that

- If $\mathfrak{G}_{p,q}$ is Hardy then $\min(p, q) \leqslant 0$ and $\max(p, q) \leqslant 1$.
- If $\min(p, q) \leqslant 0$ and $\max(p, q) < 1$ then $\mathfrak{G}_{p,q}$ is Hardy.

In chapter 7 I will prove that the second condition is also a necessary one (cf. Theorem 7.3). It confirms the hypothesis announced in [36].

Gaussian product of power means. This definition, which generalize Arithmetic–geometric mean, was defined [in a slightly more general version] by Gustin [17]. Let $p \in \mathbb{N}$, $\lambda \in \mathbb{R}^{p+1}$ and v be an all-positive-components vector. One defines

$$\begin{aligned} v^{(0)} &= v, \\ v^{(i+1)} &= \Big(\mathcal{P}_{\lambda_0}(v^{(i)}), \mathcal{P}_{\lambda_1}(v^{(i)}), \dots, \mathcal{P}_{\lambda_p}(v^{(i)})\Big)\,, \quad i = 0,\, 1,\, 2, \dots. \end{aligned}$$

Then, for any $0 \leqslant k \leqslant p$, there exists a limit $\lim_{i\to\infty} v_k^{(i)}$. Moreover it does not depend on k. We will call it a Gaussian product of power means and denote by $\mathcal{P}_{\lambda_0} \otimes \mathcal{P}_{\lambda_1} \otimes \cdots \otimes \mathcal{P}_{\lambda_p}(v)$. In chapter 7, I will prove (cf. Theorem 7.2) that

$$\mathcal{P}_{\lambda_0} \otimes \mathcal{P}_{\lambda_1} \otimes \cdots \otimes \mathcal{P}_{\lambda_p} \text{ is Hardy} \iff \max \lambda < 1.$$

Attention: Some effective estimates for the speed of convergence have been recently given in [P5].

Part I

Quasi-arithmetic means

$$f^{-1}\left(\frac{f(a_1)+f(a_2)+\ldots+f(a_n)}{n}\right)$$

Chapter 2

Introduction

One of the most popular families of means encountered in the literature consists of quasi-arithmetic means (Q-A for short). That mean is defined for any continuous strictly monotone function $f\colon U \to \mathbb{R}$, U – an open interval. When $a = (a_1, \ldots, a_n)$ is a sequence of points in U and $w = (w_1, \ldots, w_n)$ is a sequence of *weights* ($w_i > 0$, $w_1 + \cdots + w_n = 1$), then the mean $\mathfrak{M} = \mathfrak{M}_f(a, w)$ is defined by the equality

$$\mathfrak{M}_f(a, w) := f^{-1}\left(\sum_{i=1}^{n} w_i f(a_i)\right) .$$

According to [21, pp. 158–159], this family of means was dealt with for the first time in the papers [12; 25; 31], just a couple of years before the coming out of that benchmark contribution [21]. Among the names of independent, if simultaneous, contributors there is that of Kolmogorov. He had explained in [25] the *naturality* of the above construction. In fact, an extremely short list of his most natural postulates [to be satisfied by a mean] forces the existence of a continuous function governing that mean, exactly as in the definition above. The issue is also discussed in the by-now-classical encyclopaedic publications [6] and [7]. It is underlined there that one thus naturally generalizes the *Power Means*. Indeed, the latter family, containing the most popular means: arithmetic, geometric, quadratic, harmonic, is encompassed by this approach upon taking the functions

$$k_\alpha(x) = \begin{cases} x^\alpha & \text{if } \alpha \neq 0, \\ \ln x & \text{if } \alpha = 0, \end{cases} \tag{2.1}$$

for $x \in U = (0, +\infty),\ \alpha \in I = \mathbb{R}$.

In the 1960s Cargo and Shisha [10; 11] introduced a metric among quasi-arithmetic means. Namely, for f and g both continuous, strictly monotone and with the same domain, they have defined a distance

$$\rho(\mathfrak{M}_f, \mathfrak{M}_g) := \sup\{|\mathfrak{M}_f(a, w) - \mathfrak{M}_g(a, w)| : a \text{ and } w \text{ admissible}\}.$$

They also furnished some majorizations for $\rho(\mathfrak{M}_f, \mathfrak{M}_g)$. One of their results is the proposition below; hereafter $\|\cdot\|_p$ denotes the standard L^p standard norm in the space $L^p(X)$ over a suitable space X $(1 \leqslant p \leqslant \infty)$.

If not otherwise stated, the interval under consideration is arbitrary.

Theorem 2.1 ([11],Theorem 4.2)**.** *Let U be an interval, $g \in \mathcal{C}(U)$ be strictly monotone, $f \in \mathcal{C}^1(U)$, $\inf|f'| > 0$. Then* $\rho(\mathfrak{M}_f, \mathfrak{M}_g) \leqslant \frac{2\|f-g\|_\infty}{\inf|f'|}$.

Departing from another observation in [11], we will prove in chapter 3 an alternative estimate for the distance between two quasi-arithmetic means satisfying certain smoothness conditions. An important tool for that will be the operator introduced by Mikusiński in [29]. The relevant result will be presented in Theorem 3.3.

Remark 2.1. Note that the left hand side is symmetric with respect to f and g, while the right one is not. One could clearly symmetrize using the min function. Nevertheless, this operation will be omitted to keep the notation compact. The same remark applies to many a result within the present thesis.

In fact, let U be an interval and $\mathcal{C}^{2\neq}(U)$ be the class of functions from $\mathcal{C}^2(U)$ with the first derivative vanishing nowhere in U (if a boundary point belongs to U, as will happen in section 3.3, then we will assume the existence of the corresponding one-sided, second derivative, and the one-side first derivative nonzero at that point). Within this class one defines the operator $A\colon \mathcal{C}^{2\neq}(U) \to \mathcal{C}(U)$ sending f to A_f, by the formula[1]

$$A_f := \frac{f''}{f'}\,.$$

This operator has, due to [29], numerous applications in the comparison of means, as exemplified in Proposition 3.3.1. In fact, it will enable us to compare means in huge families, not only in pairs. Precisely this kind of comparison was being advanced by Polish mathematicians in the late 1940s.

One particularly important fact concerning the operator A was discovered by Mikusiński, who published his result, [29, (5)], in the first post-war issue of "Studia Mathematica"[2]. It is quite surprising that such a useful result has not been included in the referential book [6].

Let us not that some of the results included in Part I have been inspired by a benchmark paper [35] by Páles. Namely, he proved, using the three parameters' operator

$$P_f(x,y,z) := \frac{f(x)-f(y)}{f(x)-f(z)}$$

defined on $\{(x,y,z)\in U^3\colon x\neq z\} =: \Delta$, the following

---

[1]Mikusiński and, independently, Łojasiewicz, proved that comparability of quasi-arithmetic means might be easily expressed in terms of operator A. Besides, in the mathematical economy, the *negative* of this operator happenes to be called the *Arrow-Pratt measure of risk aversion*.

[2]the flagship journal of the pre-war Lvov Mathematical School, established by H. Steinhaus and S. Banach.

Theorem 2.2 ([35], Corollary 1)**.** *Let U be an interval, f and f_n, $n \in \mathbb{N}$, be continuous, strictly monotone functions defined on U.*

Then $\Big(\mathfrak{M}_{f_n} \to \mathfrak{M}_f$ *pointwise* $\Big) \iff \Big(P_{f_n} \to P_f$ *pointwise on* $\Delta \Big)$.

Remark 2.2. Let U be an interval and $f,\, g \in \mathcal{C}^{2\neq}(U)$. Then the following conditions are equivalent:

(i) $A_f(x) = A_g(x)$ for all $x \in U$,

(ii) $f = \alpha g + \beta$ for some $\alpha,\, \beta \in \mathbb{R},\; \alpha \neq 0$,

(iii) $\mathfrak{M}_f(a,\, w) = \mathfrak{M}_g(a,\, w)$ for all vectors $a \in U^n$ and arbitrary corresponding weights w

(iv) $P_f = P_g$ on Δ

(see, for instance, [21, p. 66], [29]).

The results enclosed in chapter 3 was published at the beginning of 2013. Few mounths later, during the 15th International Conference on Functional Equations and Inequalities[3] Páles himself asked about possible generalizations, of the ($\Leftarrow$) part of his theorem. In fact, he asked for a majorization of the distance $\rho(\mathfrak{M}_f, \mathfrak{M}_g)$ in terms of the operator P. Related with this is another problem presented in chapter 3 in Example 4.1. On the other hand, it is natural to look for possible strengthening of Theorem 3.3 and Corollary 3.3.2.

In chapter 4 we are going to propose such an estimate which not only implies the ($\Leftarrow$) part of Páles' result but also leads to a handy strengthening of Theorem 3.3; compare Corollary 4.3.1 and Theorem 4.2, respectively.

---

[3]The conference held in May 2013, while the results enclosed in chapter 3 was published in the beginning of 213

Chapter 3

When is a family of quasi-arithmetic means a scale?

For a family $\{k_\alpha \,|\, \alpha \in I\}$ of real $\mathcal{C}^2$ functions defined on U (I, U – open intervals) and satisfying some mild regularity conditions, we will work towards justifying when the mapping $I \ni \alpha \mapsto \mathfrak{M}_{k_\alpha}(a, w)$ is a continuous bijection between I and $(\min a, \max a)$, for every fixed non-constant sequence $a = \big(a_i\big)_{i=1}^n$ with values in U and weights $w = \big(w_i\big)_{i=1}^n$. In such a situation one says that the family of functions $\{k_\alpha\}$ generates a *scale* on U.

This chapter is based on the paper [P1].

3.1 Scales

If a non-constant vector $a \in U^n$ and weights w are fixed then the mapping $f \mapsto \mathfrak{M}_f(a,\, w)$ takes continuous monotone functions $f\colon U \to \mathbb{R}$ to the interval $(\min a, \max a)$. One is interested in finding such families of functions $\{k_\alpha\colon U \to \mathbb{R}\}_{\alpha\in I}$, where I is an interval, that for every non-constant vector a with values in

U and arbitrary fixed corresponding weights w, the mapping $I \ni \alpha \mapsto \mathfrak{M}_{k_\alpha}(a, w)$ be a *bijection* onto $(\min a, \max a)$. Every such a family of means $\mathfrak{M}_{k_\alpha}$ is called *scale on* U.

The problem of finding conditions, for a family of means, equivalent to its being a scale has been discussed for various families. For instance, a set of conditions pertinent for Gini means was presented in [1]. Many results concerning means may be expressed in a compact way in terms of scales. Probably the most famous is the fact that the family of power means is a scale on $(0, +\infty)$. It was proved for the first time (for arbitrary weights) in [2]. More about the underlying history, as well as another proof, was given in [6, p. 203]. In the last section of the present note we will present a new, extremely short proof of this classical fact.

3.2 Comparison of means

Dealing with means, we would like to know whether (a) one mean is not smaller than the other, whenever both are defined on the same interval and computed on same, but arbitrary, set of arguments. And, when (a) holds true, whether (b) the two means, evaluated on same arguments, are equal only when $a_1 = a_2 = \cdots = a_n$. With (a) and (b) holding true, we would say that the first mean is *greater* than the second. (Note that this relation slightly strengtens the common comparablity between means.)

As long as quasi-arithmetic means are concerned, the comparability of $\mathfrak{M}_f$ and $\mathfrak{M}_g$ as such turns out to be intimately related to the convexity of the function $f \circ g^{-1}$; see items (ii) and (iii) in Proposition 3.3.1 below.

Unfortunately, however, when it comes to scales, the family of objects to handle becomes uncountable. Hence one is forced to use another tool, allowing to tell something about uncountable families of means. Its concept goes back to a seminal paper [29].

A key operator from [29], denoted in this thesis by A, is used in item (i) in our technically crucial Proposition 3.3.1.

Prior to that we will present both necessary and sufficient conditions, for a family of functions $\{k_\alpha\}_{\alpha\in I}$ defined on a common interval U, to generate a scale on U. The key conditions in our Theorems 3.1 and 3.2 are given, naturally enough, in terms of the operator A. Reiterating, it is handy to compare means with its help. So we begin with

Theorem 3.1. *Let U be an interval, I – an open interval, $(k_\alpha)_{\alpha\in I}$ – a family of functions on U, $k_\alpha \in \mathcal{C}^{2\neq}(U)$ for all α.*

If $I \ni \alpha \mapsto A_{k_\alpha}(x) \in \mathbb{R}$ is increasing and 1–1 on a dense subset of U, and is onto for all $x \in U$, then $(\mathfrak{M}_{k_\alpha})_{\alpha\in I}$ is an increasing scale on U.

A proof of this theorem is given in section 3.4. As a matter of fact, we will need for applications a wider version of the above theorem. Namely, we extend the setup as follows.

In the definition of a scale one may replace $\min a$ and $\max a$ by arbitrary bounds $L(a, w)$ and $H(a, w)$ respectively, with some functions L and H.[1] Then such a modified family of means is called a *scale between L and H*. Such generalization is very natural and is frequently used, e. g. in [6, pp. 323, 364].

Bounds in a scale, in most cases, are either quasi-arithmetic means or min, or max. In order to make the notation more homogeneous, we introduce two extra symbols $\perp$ and $\top$, and write henceforth, purely formally, $\mathfrak{M}_\perp = \min$ and $\mathfrak{M}_\top = \max$. We also adopt the convention that $A_\perp = -\infty$ and $A_\top = +\infty$.

Attention. In some papers scales may as well be decreasing. In fact, we do not lose generality if we assume that all scales are increasing, because whenever a family $\{k_\alpha\}_{\alpha\in I}$ generates a decreasing scale and $\varphi\colon J \to I$ is continuous, decreasing, 1–1

[1] We slightly abuse the notation here, as most of the researchers active in the field of means do.

and onto, then the family $\{k_{\varphi(\alpha)}\}_{\alpha\in J}$ generates an increasing scale (see, e. g., Proposition 3.5.4 in Section 3.5).

Moreover, using this property, we will assume in all proofs that $I=\mathbb{R}$.

Corollary 3.2.1 (Bounded Scale). *Let U be an interval, I – an open interval. Let $l,\, h \in \mathcal{C}^{2\neq}(U)\cup\{\bot,\, \top\}$ and $(k_\alpha)_{\alpha\in I}$ be a family of functions, $k_\alpha \in \mathcal{C}^{2\neq}(U)$ for all $\alpha\in I$.*

If $I \ni \alpha \mapsto A_{k_\alpha}(x) \in \mathbb{R}$ is increasing (decreasing) and 1–1 on a dense subset of U, and is onto $(A_l(x),\, A_h(x))$ for all $x\in U$, then $(\mathfrak{M}_{k_\alpha})_{\alpha\in I}$ is an increasing (decreasing) scale between $\mathfrak{M}_l$ and $\mathfrak{M}_h$.

The proof is just a specification of the proof of Theorem 3.1.

Remark. If, in the above corollary, $l,\, h\in\mathcal{C}^{2\neq}(U)$, then it would be enough to assume that the mapping $\alpha\mapsto A_{k_\alpha}(x)$ be onto for almost all $x\in U$. (Because, then, by Corollary 3.3.2, one gets the convergence in $L^1(U)$.)

The strength of Theorem 3.1 is visible in the following example (or, rather, exercise).

Example 3.1. Let $U=(\frac{1}{e},\,+\infty)$ and $k_\alpha(x)=x^{\alpha x}$ for $\alpha\in\mathbb{R}\setminus\{0\}$.
Find such a function k_0 that the completed family $(k_\alpha)_{\alpha\in\mathbb{R}}$ generates a scale on U.

By the definition of the operator A, for $\alpha\neq 0$ there holds

$$A_{k_\alpha}(x)=\frac{1}{x(\ln x+1)}+\alpha(\ln x+1)\,.$$

In view of Theorem 3.1 we will be done, provided $\alpha\mapsto A_{k_\alpha}(x)$ is increasing, 1–1 and onto $\mathbb{R}$ for all $x\in U$. However, changing the view point from x to α, we clearly have

$$\mathbb{R}\setminus\{0\}\ni\alpha\mapsto A_{k_\alpha}(x)\in\mathbb{R}\setminus\left\{\frac{1}{x(\ln x+1)}\right\},\quad x\in U\,.$$

Hence it is natural to take $k_0 = A^{-1}\left(\frac{1}{x(\ln x+1)}\right)$. Then the pattern $A^{-1}(A_f) = \int e^{\int A_f}$ gives automatically $k_0(x) = x\ln x$.
Therefore, an increasing scale on $(\frac{1}{e}, +\infty)$ is generated by the family

$$k_\alpha = \begin{cases} x \mapsto x^{\alpha x} & \text{if } \alpha \neq 0\,, \\ x \mapsto x\ln x & \text{if } \alpha = 0\,. \end{cases}$$

Moreover, it is now immediate to note that, in turn, the same family of functions generates a *decreasing* scale on $(0, \frac{1}{e})$.

Now, how about a possible reversing of Theorem 3.1 ? This point is rather fine; the existence of a scale implies a somehow weaker set of properties than the one assumed in Theorem 3.1. To the best of author's knowledge, the problem of finding a set of conditions *exactly* equivalent to generating a scale is much more sophisticated than it is presented here.

Theorem 3.2. *Let U be an interval, I – an open interval, $(k_\alpha)_{\alpha\in I}$ – a family of functions on U, $k_\alpha \in \mathcal{C}^{2\neq}(U)$ for all α.*

If $(\mathfrak{M}_{k_\alpha})_{\alpha\in I}$ is an increasing scale then there exists a dense subset $X \subset U$ such that the mapping $I \ni \alpha \mapsto A_{k_\alpha}(x) \in \mathbb{R}$ is increasing, 1–1 and onto for all $x \in X$.

A proof of this theorem is given in Section 3.4, immediately after the proof of Theorem 3.1.

3.3 An Arrow-Pratt like operator in Quasi-Arithmetic means

In what follows we will extensively use the operator A. Here we recall, after [29], some of its key properties. We also rephrase in the terms of A an important result from [11].

All this will be instrumental in showing that many nontrivial families of functions do generate scales. We will also deduce about the limit properties of our quasi-arithmetic means, stating a new result (Proposition 3.5.3) inspired, to some extent, by the paper [24].

Regarding scales as such, many examples of them were furnished in [6, p. 269]. Scales were also used by the old Italian school of statisticians; see, e. g., [3; 4; 5; 15; 43; 45]. One of significant results from that last group of works will be presented, with a new and compact proof, in Proposition 3.5.4. That new approach will, we hope, show how quickly one can nowadays prove old results.

Let $f \in \mathcal{C}^1(U)$ be a strictly monotone function such that $f'(x) \neq 0$ for all $x \in U$. Then there either holds $f'(x) < 0$ for all $x \in U$, or else $f'(x) > 0$ for all $x \in U$. So we define the sign $\mathrm{sgn}(f')$ of the first derivative of f to be $\mathrm{sgn}(f')(x)$, where x is any point in U. The key tool in our approach is

Proposition 3.3.1 (Basic comparison). *Let U be an interval, $f,\ g \in \mathcal{C}^{2\neq}(U)$. Then the following conditions are equivalent:*

(i) $A_f > A_g$ *on a dense set in* U *,*

(ii) $(\mathrm{sgn} f') \cdot (f \circ g^{-1})$ *is strictly convex,*

(iii) $\mathfrak{M}_f(a,\ w) \geqslant \mathfrak{M}_g(a,\ w)$ *for all vectors $a \in U^n$ and weights w, with both sides equal only when a is a constant vector.*

For the equivalence of (i) and (iii), see [29, p. 95] (this characterization of comparability of means had been independently obtained by S. Łojasiewicz – see footnote 2 in [29]). For the equivalence of (ii) and (iii), see, for instance, [10, p. 1053].

In the course of comparing means, one needs to majorate the difference between two means. If the interval U is unbounded then, of course, the difference between any given two means can be unbounded (for example such is the difference between the arithmetic and geometric mean). In order to eliminate this drawback, we will henceforth suppose that the means are always defined on a compact interval. It will be with no loss of generality, because it is easy to check that a family of means defined on U is a scale on U if and only if those means form a scale on D, when treated as functions $D \to \mathbb{R}$, for every closed subinterval $D \subset U$. Indeed, if a is a vector with values in U, then a is also a vector with values in D for some closed subinterval D of U.

So, from now on, we have U – a compact interval, $g \in \mathcal{C}^{2\neq}(U)$ increasing, and, clearly, $A_g \in L^1(U)$. The following theorem is of utmost technical importance.

Theorem 3.3. *Let U be a closed, bounded interval and $f, g \in C^{2\neq}(U)$. Then*

$$\rho(\mathfrak{M}_f, \mathfrak{M}_g) \leqslant |U| \exp(2 \|A_f\|_1) \sinh\Big(2 \|A_g - A_f\|_1\Big)$$

for all a and w ($\|\cdot\|_1$ is taken in the space $L^1(U)$).

Remark. This theorem and corollary below will be generalized in chapter 4 (Theorem 4.2). However, the proof given here below has been enclosed in the original paper [P1].

Proof. Fix a and w. Replacing the initial function f by $\alpha f + \beta$ (compare Remark 2.2) with $\alpha = \frac{1}{f'(\min a)}$, $\beta = -\frac{f(\min a)}{f'(\min a)}$, we can assume without loss of generality that

$$f(x) = \int_{\min a}^{x} \exp\left(\int_{\min a}^{s} A_f(t)dt\right) ds, \quad x \in U. \tag{3.1}$$

Moreover, let us make the same simplification for g. Then $f(\min a) = g(\min a) = 0$ and both functions are positive and increasing on $(\min a, \max a)$.

Then, much like in [11, pp. 215–216], we have

$$\begin{aligned}&\mathfrak{M}_f(a, w) - \mathfrak{M}_g(a, w)\\ &= (f^{-1})'(\gamma) \sum_{1 \leqslant i < j \leqslant n} w_i w_j \Big(g(a_i) - g(a_j)\Big)\Big(\theta(z_i) - \theta(z_j)\Big),\end{aligned}$$

where $\theta = (f \circ g^{-1})'$, for certain $\gamma \in [\min a, \max a]$ and $z_i \in g(U)$, $i = 1, \ldots, n$. The vector w denotes, as usual, weights so, naturally, $\sum_{1 \leqslant i < j \leqslant n} w_i w_j < \frac{1}{2}$. Hence

$$\begin{aligned}
&|\mathfrak{M}_f(a,w) - \mathfrak{M}_g(a,w)| \\
&= \left| (f^{-1})'(\gamma) \sum_{1 \leqslant i < j \leqslant n} w_i w_j \Big(g(a_i) - g(a_j) \Big) \Big(\theta(z_i) - \theta(z_j) \Big) \right| \\
&\leqslant \frac{\|(f^{-1})'\|_\infty}{2}\, g(\max a) \sup_{z,v \in g(U)} |\theta(z) - \theta(v)|
\end{aligned}$$

Putting $\varepsilon := \|A_f - A_g\|_1$, we assuredly have

$$\frac{f'}{g'} = e^{\int A_f - A_g} \in (e^{-\varepsilon},\, e^{\varepsilon})\,.$$

Thus $\theta(z) = (f \circ g^{-1})'(z) = \frac{f' \circ g^{-1}(z)}{g' \circ g^{-1}(z)} \in (e^{-\varepsilon},\, e^{\varepsilon})$. What is more,

$$g(\max a) = \int_{\min a}^{\max a} g'(s)\, ds \leqslant \int_{\min a}^{\max a} e^{\varepsilon} f'(s)\, ds = e^{\varepsilon} f(\max a)\,.$$

Whence, estimating further,

$$\begin{aligned}
&|\mathfrak{M}_f(a,w) - \mathfrak{M}_g(a,w)| \\
&\leqslant \frac{\|(f^{-1})'\|_\infty}{2}\, g(\max a) \sup_{z,v \in g(U)} |\theta(z) - \theta(v)| \\
&\leqslant \frac{\|(f^{-1})'\|_\infty\, e^{\varepsilon}}{2} f(\max a) \left(e^{\varepsilon} - e^{-\varepsilon} \right) \\
&= \frac{f(\max a)}{\inf f'} \cdot \frac{e^{2\varepsilon} - 1}{2} \\
&\leqslant \frac{f(\max a)}{\inf f'} \sinh 2\varepsilon\,.
\end{aligned}$$

But, by (3.1), we also know that

$$\begin{aligned}
f(\max a) &= \int_{\min a}^{\max a} \exp\Big(\int_{\min a}^{s} A_f\Big) \\
&\leqslant |U| \exp(\|A_f\|_1) \qquad (3.2)
\end{aligned}$$

and

$$\inf f' = \inf_{s\in U} \exp(\int_{\min a}^{s} A_f) \geqslant \exp(-\|A_f\|_1). \tag{3.3}$$

So, prolonging the previous chain of estimations and using (3.2) and (3.3),

$$|\mathfrak{M}_f(a,w) - \mathfrak{M}_g(a,w)| \leqslant |U| \exp(2\|A_f\|_1) \sinh 2\|A_g - A_f\|_1 .$$

□

Remark. Theorem 3.3 would still hold true, for any vector a, if $\|\cdot\|_1$ denote the standard norm in the space $L^1(\min a, \max a)$.

One immediately gets the following

Corollary 3.3.2. *Let U be a closed, bounded interval and $f \in \mathcal{C}^{2\neq}(U)$. Moreover, let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions from $\mathcal{C}^{2\neq}(U)$ satisfying $A_{f_n} \to A_f$ in $L^1(U)$. Then $\mathfrak{M}_{f_n} \rightrightarrows \mathfrak{M}_f$ uniformly with respect to a and w.*

Heading towards the main results of the present chapter, we state now

Proposition 3.3.3. *Let U be a closed bounded interval, $(k_\alpha)_{\alpha\in\mathbb{R}}$ – a family of functions from $\mathcal{C}^{2\neq}(U)$.*
(A) *If $(\mathfrak{M}_{k_\alpha})_{\alpha\in\mathbb{R}}$ is an increasing scale then $(A_{k_\alpha})_{\alpha\in\mathbb{R}}$ satisfies all the conditions* (a) *through* (d) *listed below.*

(a) *if $\alpha_i \to \alpha$, then $A_{k_{\alpha_i}} \to A_{k_\alpha}$ on a dense subset of U (independent of α and (α_i)),*

(b) *if $\alpha < \beta$, then $A_{k_\alpha} < A_{k_\beta}$ on a dense subset of U (independent of α and β),*

(c) *if $\alpha \to -\infty$, then $A_{k_\alpha}(x) \to -\infty$ on a dense subset of U (independent of the sequence α),*

(d) *if $\beta \to +\infty$, then $A_{k_\beta}(x) \to +\infty$ on a dense subset of U (independent of the sequence β).*

(B) *Strengthening conditions* (a), (c) *and* (d) *to*

(e) *if* $\alpha_i \to \alpha$, *then* $A_{k_{\alpha_i}} \to A_{k_\alpha}$,

(f) $(\alpha \to -\infty \Rightarrow A_{k_\alpha}(x) \to -\infty)$ *and* $(\beta \to \infty \Rightarrow A_{k_\beta}(x) \to +\infty)$ *for all* $x \in U$

suffices to reverse the implication: (b), (e), *and* (f) *imply* $(\mathfrak{M}_{k_\alpha})_{\alpha\in\mathbb{R}}$ *being an increasing scale.*

Proof. To simplify the notation, having a and w fixed, we write shortly

$$F(\alpha) = \mathfrak{M}_{k_\alpha}(a,\, w),$$

$F\colon \mathbb{R} \to (\min a,\, \max a)$. And then one simply checks step by step:

(a) With no loss of generality one may consider $\alpha_i \to \alpha+$.

Suppose the converse – that there exists an open subset $V \subset U$, that $A_{k_{\alpha_i}} \not\to A_{k_\alpha}$ on V. Then there exists $m \in C^{2\neq}(V)$ such that $A_{k_{\alpha_i}} < A_m < A_{k_\alpha}$. Hence for all i and non-constant vector a with corresponding weights w, $\mathfrak{M}_{k_{\alpha_i}} < \mathfrak{M}_m < \mathfrak{M}_{k_\alpha}$. Contradiction.

At this moment the produced dense set possibly depends on α.

We will show how to produce an independent-of-α dense set. Let

$$X_{(a)} := \Big\{x\colon \text{for an arbitrary } \alpha, \text{ if } \alpha_i \to \alpha, \text{ then } A_{k_{\alpha_i}}(x) \to A_{k_\alpha}(x)\Big\}.$$

That is to say

$$X_{(a)} = \Big\{x\colon \forall_{\alpha\in\mathbb{R}}\ \forall_{\varepsilon>0}\ \exists_{\delta>0}\ \forall_{\beta\in B(\alpha,\delta)}\ \Big|A_{k_\alpha}(x) - A_{k_\beta}(x)\Big| < \varepsilon\Big\},$$

or equivalently (using the monotonicity of the mapping $\alpha \mapsto A_{k_\alpha}(x)$ for all $x \in U$) one obtain

$$\begin{aligned} X_{(a)} &= \Big\{x \colon \forall \alpha \in \mathbb{R},\ \forall \varepsilon > 0,\ \exists \delta > 0,\ \Big|A_{k_{\alpha-\delta}}(x) - A_{k_{\alpha+\delta}}(x)\Big| < \varepsilon\Big\} \\ &= \Big\{x \colon \forall \alpha \in \mathbb{Q},\ \forall \varepsilon \in \mathbb{Q}_+,\ \exists \delta > 0,\ \Big|A_{k_{\alpha-\delta}}(x) - A_{k_{\alpha+\delta}}(x)\Big| < \varepsilon\Big\} \\ &= \bigcap_{\substack{\alpha \in \mathbb{Q} \\ \varepsilon \in \mathbb{Q}_+}} \Big\{x \colon \exists \delta > 0,\ \Big|A_{k_{\alpha-\delta}}(x) - A_{k_{\alpha+\delta}}(x)\Big| < \varepsilon\Big\}. \end{aligned}$$

But, if $\alpha_i \to \alpha$, then $A_{k_{\alpha_i}} \to A_{k_\alpha}$ on a dense subset of U. Thus

$$\Big\{x \colon \exists \delta > 0,\ \Big|A_{k_{\alpha-\delta}}(x) - A_{k_{\alpha+\delta}}(x)\Big| < \varepsilon\Big\}$$

is dense and open for all $\varepsilon > 0$ and $\alpha \in \mathbb{R}$. In the outcome, $X_{(a)}$ is a dense G_δ-set.

(b) if $\alpha < \beta$, we have $F(\alpha) \leqslant F(\beta)$ and the equality holds iff a is constant. So by Proposition 3.3.1 we have $A_{k_\alpha} < A_{k_\beta}$ on a dense set.
Let

$$\begin{aligned} X_{(b)} &:= \Big\{x \in U : \forall \alpha,\ \forall \beta \neq \alpha,\ A_{k_\alpha}(x) \neq A_{k_\beta}(x)\Big\}, \\ E_{\alpha,\beta} &:= \Big\{x \in U \colon A_{k_\alpha}(x) \neq A_{k_\beta}(x)\Big\}. \end{aligned}$$

We have that if $[\alpha',\, \beta'] \subset [\alpha,\, \beta]$ then $E_{\alpha,\beta} \subset E_{\alpha',\beta'}$, and $E_{\alpha,\beta}$ is an open dense set. Thus

$$X_{(b)} = \bigcap_{\substack{\alpha,\beta \in \mathbb{R} \\ \alpha \neq \beta}} E_{\alpha,\beta} = \bigcap_{\substack{\alpha,\beta \in \mathbb{Q} \\ \alpha \neq \beta}} E_{\alpha,\beta}$$

is a dense G_δ-set.

(c) The proof is completely similar to that of (d) given below.

(d) Let
$$X_{(d)} = \{x : \lim_{\beta\to\infty} A_{k_\beta}(x) \to +\infty\}.$$

We shall prove that if $X_{(d)}$ were not a dense set, then there would exist a closed, non-trivial interval D such that

$$M_D := \sup_{x\in D} \lim_{\beta\to\infty} A_{k_\beta}(x) < \infty\,.$$

Arguing by contradiction, let us assume that for any non-trivial closed interval D one has $M_D = \infty$.

We will prove that for any non-trivial, closed interval D_0 one could take non-trivial, closed intervals $D_0 \supset D_1 \supset \dots$ satisfying

$$\lim_{\beta\to\infty} A_{k_\beta}(x) > j, \text{ for any } j \in \mathbb{N}_+ \text{ and } x \in D_j.$$

Indeed, for any $j \in \mathbb{N}$, in view of $M_{D_j} = +\infty$, there exists $x_j \in D_j$ and β_j such that $A_{k_{\beta_j}}(x_j) > j+1$. In particular, one can take a closed, non-trivial interval $D_{j+1} \ni x_j$, $D_{j+1} \subset D_j$ satisfying

$$A_{k_{\beta_j}}(x) > j+1 \text{ for any } x \in D_{j+1}.$$

Whence, $X_{(d)} \supset \bigcap_{j=0}^{\infty} D_j \neq \emptyset$, so that $X_{(d)} \cap D_0 \neq \emptyset$.

Eventually, upon taking a set D such that $M_D < +\infty$, one gets
$$\mathfrak{M}_{k_\beta}(v,q) \leqslant \mathfrak{M}_{\exp(M_D\cdot x)}(v,q) < \max v$$
for all β and v, q such that $v \in D^n$. Hence the family $\{k_\beta\}$ would not generate a scale on U. So $X_{(d)}$ is a dense set.

To prove part (B) one needs to show that, under (e) and (f), $(\mathfrak{M}_{k_\alpha})_{\alpha\in\mathbb{R}}$ is a scale on U.

By Proposition 3.3.1 we know that F is 1–1. Additionally, when arguing to this side, we know that if $\alpha \nearrow \alpha_0$ then $A_{k_\alpha} \nearrow$

$A_{k_{\alpha_0}}$. So $A_{k_\alpha} \rightrightarrows A_{k_{\alpha_0}}$ on $[\min a, \max a]$. Therefore, by Corollary 3.3.2, we have $\mathfrak{M}_{k_\alpha} \rightrightarrows \mathfrak{M}_{k_{\alpha_0}}$ with respect to a and w. Thus F is continuous and 1–1.

To complete the proof, it sufficies to show that

$$\lim_{\alpha\to-\infty} F(\alpha) = \min a\,, \qquad \lim_{\beta\to+\infty} F(\beta) = \max a\,.$$

We know that $A_{k_\beta} \to +\infty$ on the closed interval U. So $A_{k_\beta} \rightrightarrows +\infty$ on U. Therefore, for any $M \in \mathbb{R}$ there exists $\beta_M \in I$ such that

$$F(\beta) \geqslant \mathfrak{M}_{e^{Mx}}(a, w), \text{ for all } \beta > \beta_M.$$

Now, letting $M \to +\infty$, and knowing that $\{e^{tx}\colon t \neq 0\} \cup \{x\}$ generates a scale on $\mathbb{R}$ (a folk-type theorem, see also Remark 3.1 below) we get

$$F(\beta) \xrightarrow[\beta\to+\infty]{} \max a\,.$$

One may similarly prove that

$$F(\alpha) \xrightarrow[\alpha\to-\infty]{} \min a\,.$$

So F is a continuous bijection between $\mathbb{R}$ and $(\min a, \max a)$. Hence $(\mathfrak{M}_{k_\alpha})_{\alpha\in\mathbb{R}}$ is a scale on U. □

Remark 3.1. To prove that the family $\{e^{tx}\colon t \neq 0\} \cup \{x\}$ generates a scale on $\mathbb{R}$ it is enough, having data a, w, to consider the all-positive-components-vector $v = (e^{a_1}, \dots, e^{a_n})$. And then to use the fact that the family of power means evaluated on v with weights w is a scale on $\mathbb{R}_+$.

Corollary 3.3.4 (strengthening of Proposition 3.3.3)**.** *Let U be an interval, $(k_\alpha)_{\alpha\in\mathbb{R}}$– a family of functions, $k_\alpha \in \mathcal{C}^{2\neq}(U)$ for all α.*

(A) *If $(\mathfrak{M}_{k_\alpha})_{\alpha\in\mathbb{R}}$ is an increasing scale then there exists a dense set $X \subset U$ such that*

(a) *if $\alpha_i \to \alpha$, then $A_{k_{\alpha_i}} \to A_{k_\alpha}$ on X,*

(b) *if* $\alpha < \beta$, *then* $A_{k_\alpha} < A_{k_\beta}$ *on* X,

(c) *if* $\alpha \to -\infty$, *then* $A_{k_\alpha}(x) \to -\infty$ *on* X,

(d) *if* $\beta \to +\infty$, *then* $A_{k_\beta}(x) \to +\infty$ *on* X.

(B) *Under the stronger condition*

(e) *if* $\alpha_i \to \alpha$, *then* $A_{k_{\alpha_i}} \to A_{k_\alpha}$,

(f) *(*$\alpha \to -\infty \Rightarrow A_{k_\alpha}(x) \to -\infty$*)* *and* *(*$\beta \to +\infty \Rightarrow A_{k_\beta}(x) \to +\infty$*)* *for all* $x \in U$

the entire implication of the corollary can be reversed: (b), (e), *and* (f) *imply* $(\mathfrak{M}_{k_\alpha})_{\alpha\in\mathbb{R}}$ *being an increasing scale.*

This corollary says that, in Proposition 3.3.3, one can have a single common subset (X) of U on which the conditions (a) through (d) hold.

Proof. The (B) parts in Proposition 3.3.3 and Corollary 3.3.4 are the same except the notion of X given to the dense set in (B) in Proposition 3.3.3. We are going to show the implication (A).

We might assume that U is a closed interval (compare the comment paragraph below Proposition 3.3.1).

Let us denote a dense sets appearing in (a) through (d) in the Proposition 3.3.3 by $X_{(a)}, \dots, X_{(d)}$. Our aim is to prove that each of these sets is a dense G_δ-set. By [the proof of] Proposition 3.3.3, $X_{(a)}$ and $X_{(b)}$ are declared to be dense G_δ-sets. By definition

$$X_{(d)} = \{x \in U\colon \lim_{\beta\to+\infty} A_{k_\beta}(x) \to +\infty\}.$$

Due to Proposition 3.3.3 we know that $X_{(d)}$ is dense. Let

$$Y_s := \{x \in U\colon \lim_{\beta\to+\infty} A_{k_\beta}(x) > s\}.$$

Observe that Y_s is dense (because $Y_s \supset X_{(d)}$). Moreover, for all $x_0 \in Y_s$ there holds $A_{k_{\beta_0}}(x_0) > s+\delta$ for some $\beta_0 \in \mathbb{R}$ and $\delta > 0$.

Hence one may take an open neighborhood $P \ni x_0$ satisfying $A_{k_{\beta_0}}(x) > s + \frac{1}{2}\delta$ for all $x \in P$, implying $P \subset Y_s$. So Y_s is open. But the mapping $\beta \mapsto A_{k_\beta}(x)$ is nondecreasing for all $x \in U$. Hence $X_{(d)} = \bigcap_{s=1}^{\infty} Y_s$ is a dense G_δ-set. So is the set $X_{(c)}$.

Now one may take $X := X_{(a)} \cap X_{(b)} \cap X_{(c)} \cap X_{(d)}$. X is clearly dense (being a countable intersection of open dense sets). □

3.4 Proofs of Theorems 3.1 and 3.2

Proof of Theorem 3.1. Let U be an interval, $I = \mathbb{R}$, and X be that dense subset of U upon witch the mapping given in the wording of theorem is increasing and 1–1. We work with the family of functions $(k_\alpha)_{\alpha\in\mathbb{R}}$, $k_\alpha \in \mathcal{C}^{2\neq}(U)$ for all $\alpha \in \mathbb{R}$.

Let us take an arbitrary $x_0 \in X$. We know that $\mathbb{R} \ni \alpha \mapsto A_{k_\alpha}(x_0)$ is increasing, 1–1, and onto $\mathbb{R}$. Next, let us single out the function $\Phi : \mathbb{R} \to \mathbb{R}$ such that $A_{k_{\Phi(\alpha)}}(x_0) = \alpha$. This function is increasing as well.

Then for $\alpha < \beta$ we have $A_{k_{\Phi(\alpha)}} < A_{k_{\Phi(\beta)}}$ on X. But, due to the fact that $\mathbb{R} \ni \alpha \mapsto A_{k_\alpha}(x) \in \mathbb{R}$ is onto, we have

$$\lim_{\alpha\to-\infty} A_{k_{\Phi(\alpha)}}(x) = -\infty \qquad \text{and} \qquad \lim_{\beta\to+\infty} A_{k_{\Phi(\beta)}}(x) = +\infty$$

everywhere on U. So one is in a position to use the part (B) of Corollary 3.3.4. Thus the family of means $(\mathfrak{M}_{k_\alpha})_{\alpha\in\mathbb{R}}$ is an increasing scale on U. □

Proof of Theorem 3.2. Let us take X from Corollary 3.3.4. Let then fix any $x_0 \in X$. Let $\{s_p\}_{p\in\mathbb{R}}$ be the reparameterized family $\{k_\alpha\}_{\alpha\in I}$, with restriction

$$s_p = k_\alpha \text{ , where } p = A_{k_\alpha}(x_0).$$

Then we know that the mapping

$$\mathbb{R} \ni p \mapsto A_{s_p}(x) \in \mathbb{R}$$

is 1–1 and onto for all $x \in X$, and, if $p > q$,

$$A_{s_p}(x) > A_{s_q}(x).$$

Moreover, due to the fact that $A_{s_p}(x_0)$ is onto, we have

$$\lim_{p \to -\infty} A_{s_p}(x_0) = -\infty \qquad \text{and} \qquad \lim_{p \to +\infty} A_{s_p}(x_0) = +\infty$$

for all x_0. So $p \mapsto A_{s_p}(x)$ is increasing, 1–1, and onto $\mathbb{R}$ for every $x \in X$. □

Remark. The most recent research provide that the property $\mathfrak{M}_{f_\alpha} \to \max$ cannot be exhaustively characterize by the set

$$X_{(d)} = \{x \in U \colon \lim_{\beta \to +\infty} A_{k_\beta}(x) \to +\infty\}.$$

(cf. [P5])

3.5 Applications

Proposition 3.5.1 (power means do form a scale)**.** $(\mathcal{P}_\alpha)_{\alpha \in \mathbb{R}}$ *(see* (2.1) *for a definition) is a scale on* $\mathbb{R}_+$*.*

Proof. We compute the functions A_{k_α},

$$A_{k_\alpha}(x) = \frac{\alpha - 1}{x},$$

and see that the mapping $\alpha \mapsto A_{k_\alpha}(x)$ is increasing, 1–1 and onto for every $x \in \mathbb{R}_+$. So the assumptions in Theorem 3.1 hold, implying that the family (k_α) generates an increasing scale on $\mathbb{R}_+$. □

Before giving our second application, we reproduce a result which is now 14 years old.

Proposition 3.5.2 ([24])**.** *Let* $k \colon [0,1] \to \mathbb{R}$ *be a continuous monotone function. Writing* $k_\alpha(x) := k(x^\alpha)$ *for every* $\alpha > 0$*, there hold:*

(i) *if there exists the one side, nonzero derivative* $k'(0+)$ *then*

$$\lim_{\alpha\to+\infty} \mathfrak{M}_{k_\alpha} = \max\,,$$

(ii) *if there exists the one side, nonzero derivative* $k'(1-)$ *then*

$$\lim_{\alpha\to 0+} \mathfrak{M}_{k_\alpha} = \mathcal{P}_0\,.$$

We prove a somehow similar (yet not so close) result.

Proposition 3.5.3. *Let* $k \in \mathcal{C}^{2\neq}[0,1] \to (0,+\infty)$ *and* $k_\alpha(x) := k(x^\alpha)$, $\alpha \in (0,+\infty)$. *Then*

$$\lim_{\alpha\to 0+} \mathfrak{M}_{k_\alpha} = \mathcal{P}_0 \qquad \text{and} \qquad \lim_{\alpha\to+\infty} \mathfrak{M}_{k_\alpha} = \max\,. \tag{3.4}$$

If, in addition, k *is convex,*[2] *then* $(k_\alpha)_{\alpha\in(0,+\infty)}$ *generates a scale between the geometric mean and* max.

Proof. We have to prove that the mapping $(0,+\infty) \ni \alpha \mapsto A_{k_\alpha}(x) \in \mathbb{R}$ is 1–1 and onto for all $x \in (0,1)$. Let us fix an arbitrary $x \in (0,1)$. Then we have

$$A_{k_\alpha}(x) = \alpha x^{\alpha-1} A_k(x^\alpha) + \frac{\alpha-1}{x}\,.$$

When $\alpha \to 0+$, then

$$A_{k_0}(x) := \lim_{\alpha\to 0+} A_{k_\alpha}(x) = \frac{-1}{x}\,.$$

In turn, when $\alpha \to +\infty$, there holds

$$A_{k_\alpha}(x) = \underbrace{\alpha x^{\alpha-1}\frac{k''(0)}{k'(0)}}_{>-\infty} + \frac{\alpha-1}{x} \to +\infty\,.$$

The proof of formulas (3.4) is now completed.

When, additionally, g is convex, then $A_k \geqslant 0$ and, by Corollary 3.2.1, the family $\{k_\alpha\}_{\alpha\in\mathbb{R}_+}$ generates a scale on $(0,1)$ between the geometric mean and max. □

[2] in this situation one could just assume that $k \in \mathcal{C}^2[0,1]$ and k is strictly monotone, instead of assuming $k \in \mathcal{C}^{2\neq}[0,1]$

To close the present chapter, we would like to present one classical result of the Italian school of statisticians from 1910-20s. That result has been reported in [6, p. 269]. We now give it a new short proof based on Corollary 3.2.1.

Proposition 3.5.4 (Radical Means)**.** *Let* $U = \mathbb{R}_+$ *and* $(k_\alpha)_{\alpha\in\mathbb{R}_+}$*,* $k_\alpha(x) = \alpha^{1/x}$ *for* $\alpha \neq 1$*, completed by* $k_1(x) = 1/x$*, be the family of radical functions. Then this family generates a decreasing scale on* $\mathbb{R}_+$*.*

Proof. The proof happens to be quite close to that of Proposition 3.5.1. Indeed, we quickly compute

$$A_{k_\alpha}(x) = -\frac{2x + \ln\alpha}{x^2},$$

finding that the mapping $\alpha \mapsto A_{k_\alpha}(x)$ is decreasing, 1–1 and onto for every $x \in \mathbb{R}_+$. So the assumptions in Corollary 3.2.1 hold, and hence the family $(k_\alpha)_{\alpha\in\mathbb{R}_+}$ generates a decreasing scale on $\mathbb{R}_+$. □

Chapter 4

Estimates for the distance between quasi-arithmetic means

+

In the 1960s Cargo and Shisha proved some majorizations for the distance among quasi-arithmetic means. Nearly thirty years later, in 1991, Pàles presented an iff condition for a sequence of quasi-arithmetic means to converge to another QA mean. It was closely related with the three parameters' operator $(f(x) - f(y))/(f(x) - f(z))$.

So it is natural to look for similar estimate(s) in the case of the underlying functions *not* being smooth. For instance, by the way of using Pàles' operator. This is done in the present chapter. Moreover, the estimates from previous chapter.

This chapter is based on the paper [P2].

4.1 Initial example

Note that the implication converse to that in Corollary 3.3.2 does not hold – closeness of quasi-arithmetic means do not imply closeness between their arrow-pratt indexes in L^1 norm. It might already be observed in the following

Example 4.1. Let $U = [0, 2\pi]$, $f_n(x) = x + n^{-2}\sin(nx)$, $n \geqslant 2$ and $f(x) = x$ for $x \in U$. Then, by Theorem 2.1, $\rho(\mathfrak{M}_f, \mathfrak{M}_{f_n}) \leqslant 2n^{-2}$. On the other hand, it can be straightforwardly estimated that $\|A_{f_n} - A_f\|_1 = 2\,n\ln(n+1) \geqslant 4\ln 3$ for every $n \geqslant 2$.

This drawback is implied by the fact that „the first norm does not see cancellation of positive and negative part in the integral". On the other hand, in the previous chapter a couple of additional monotonicity assumptions was made from the very beginning making no point *there* to care about the traps signalled an instant ago.

In the present chapter it is otherwise. As we will see, in order to handle examples like the one above, it is more convenient to use another norm, $\|\cdot\|_*$ (defined in section 4.3.1).

In the non-smooth case, however, we will extensively deal with the operator P (defined in Introduction to the present part).

4.2 Main result

The main idea is to use the elementary fact that on compact sets the pointwise convergence of monotone functions coincides with the uniform one. However, Δ is not compact (even if U is). Therefore, finding suitable compact subsets of Δ has seemed to be of utmost importance in the search for an estimate for the distance among means.

We observe that, when x approaches z, the operator

$$P_f \colon \Delta \ni (x, y, z) \mapsto \frac{f(x) - f(y)}{f(x) - f(z)}$$

becomes unbounded. So it is natural to consider those points of Δ for which the coordinates x and z are separated one from the other. For any $\alpha > 0$ define

$$\Delta_\alpha := \{(x, y, z) \in U^3\colon\ |x - z| \geqslant \alpha\} \subset \Delta\,.$$

We are going to prove the following

Theorem 4.1. *Let U be an interval, f and g be two continuous, strictly monotone functions defined on U, and $\alpha > 0$. Then $\|P_f - P_g\|_{\infty,\Delta_\alpha} < 1$ implies $\rho(\mathfrak{M}_f, \mathfrak{M}_g) < \alpha$.*

Before starting a proof, it will be handy to recall some basic properties of the operator P. Namely, for any f,

$$P_f(x, y, z) + P_f(z, y, x) = 1 \quad \text{for all } (x, y, z) \in \Delta, \tag{4.1}$$

$$\sum_i w_i P_f\Big(\mathfrak{M}_f(a, w), a_i, z\Big) = 0 \quad \text{for all } a,\ w, \text{ and admissible } z. \tag{4.2}$$

Proof of Theorem 4.1. Fix $a \in U^n$ with corresponding weights w and write shortly

$$F := \mathfrak{M}_f(a, w) \quad \text{and} \quad G := \mathfrak{M}_g(a, w).$$

It is sufficient to find $i \in \{1, \dots, n\}$ such that $(F, a_i, G) \notin \Delta_\alpha$. Then, by the very definition of Δ_α, $|F - G| < \alpha$.

Suppose conversely that $(F, a_i, G) \in \Delta_\alpha$ for all $i \in \{1, \dots, n\}$. In particular,

$$|P_f(F, a_i, G) - P_g(F, a_i, G)| < 1 \quad \text{for all } i \in \{1, \dots, n\}.$$

Hence, upon using (4.1) and (4.2), one obtains

$$\begin{aligned} -1 &< \sum_{k=1}^n w_k \Big(P_f(F, a_k, G) - P_g(F, a_k, G)\Big) \\ &= \sum_{k=1}^n w_k P_f(F, a_k, G) + \sum_{k=1}^n w_k \Big(-1 + P_g(G, a_k, F)\Big) \\ &= -1 + \sum_{k=1}^n w_k P_g(G, a_k, F) = -1\,. \end{aligned}$$

This contradiction ends the proof. □

4.3 Applications

Corollary 4.3.1. *Let U be an interval, f and f_n, $n \in \mathbb{N}$, be strictly monotone functions defined on U, $P_{f_n} \to P_f$ pointwise in Δ. Then $\mathfrak{M}_{f_n}(a, w) \to \mathfrak{M}_f(a, w)$ for every fixed a and w.*

Moreover, if U is compact then $\mathfrak{M}_{f_n} \to \mathfrak{M}_f$ uniformly with respect to a and w in their respective ranges.

Proof. Fix: an arbitrary $a \in U^n$ with corresponding weights w, a compact interval $K \subseteq U$ such that $a \in K^n$, and a positive constant α. We are going to prove that $P_{f_n} \to P_f$ uniformly in $\Delta_\alpha \cap K^3$ and then use Theorem 4.1.

To that end fix first $p, q \in K$, $p \neq q$. Then note that P and $\mathfrak{M}$ do not change under affine transformations of f and f_n, $n \in \mathbb{N}$. So (like it is usually done in dealing with quasi-arithmetic means) assume that $f(p) = f_n(p) = 0$ and $f(q) = f_n(q) = 1$.

One then has $f_n(\cdot) = P_{f_n}(p, \cdot, q)$ and $f(\cdot) = P_f(p, \cdot, q)$. So, by the assumption, $f_n \to f$ pointwise in K. But f and f_n, $n \in \mathbb{N}$ are continuous and strictly monotone. Hence one knows (cf., e.g., [32]) that this convergence is uniform in K.

Then $f_n(x) - f_n(y) \to f(x) - f(y)$ uniformly in $\Delta_\alpha \cap K^3$, as functions of (x, y, z). Similarly $f_n(x) - f_n(z) \to f(x) - f(z)$ uniformly in the same set, as functions of (x, y, z).

Now, to prove that $P_{f_n} \to P_f$ uniformly in $\Delta_\alpha \cap K^3$, it is only needed to guarantee that $f(x) - f(z)$, as a function of (x, y, z), is bounded away from 0 in $\Delta_\alpha \cap K^3$. But it is a continuous, non-vanishing function defined on a compact set.

Therefore, there exists an integer n_α such that

$$\|P_{f_n} - P_f\|_{\infty, \Delta_\alpha \cap K^3} < 1 \quad \text{for all } n > n_\alpha \,.$$

Hence, by Theorem 4.1, one obtains

$$\rho(\mathfrak{M}_{f_n}|_K, \mathfrak{M}_f|_K) < \alpha \quad \text{for all } n > n_\alpha \,,$$

where $\mathfrak{M}_g|_K$ stands for the mean defined for a function g and vectors taking values in the relevant Cartesian products of K. This is more than needed in corollary's first statement.

As for the 'moreover' statement, one just takes $K = U$ in the above. □

Corollary 4.3.2. *Let U be a compact interval, f and f_n, $n \in \mathbb{N}$, be strictly monotone functions defined on U. Then the following conditions are equivalent:*

(i) $P_{f_n} \to P_f$ *pointwise in* Δ,

(ii) $\mathfrak{M}_{f_n} \to \mathfrak{M}_f$ *pointwise,*

(iii) $\mathfrak{M}_{f_n} \to \mathfrak{M}_f$ *uniformly with respect to a and w.*

Obviously (iii) ⇒ (ii). Moreover, by Theorem 2.2, (i) $\iff$ (ii), while, by Corollary 4.3.1, (i) ⇒ (iii).

Corollary 4.3.3. *If, in Theorem 4.1, the assumed inequality is not sharp, $\|P_f - P_g\|_{\infty,\Delta_\alpha} \leqslant 1$, then $\rho(\mathfrak{M}_f, \mathfrak{M}_g) \leqslant \alpha$.*

4.3.1 Strengthening of Theorem 3.3

Now we are going to propose some solution to the problem hinted at in Example 4.1. Recalling, that problem arose from the fact that the closeness of functions does not imply closeness of their derivatives. Therefore, Theorem 3.3 is completely useless in the situation having occurred in that example. Hence, we take possibly weaker topology to strengthen Theorem 3.3 and, consequently, Corollary 3.3.2. Instead of using the first norm, one needs to define some other norm which would circumvent the mentioned drawback of the L^1 norm.

Let U be an interval, $f\colon U \to \mathbb{R}$ be an arbitrary continuous function, and the 'oscillation' norm be defined by

$$\|f\|_* := \sup_{a,b \in U} \left| \int_a^b f(x)dx \right|.$$

We are going to justify that Theorem 3.3 might be strengthened to

Theorem 4.2. *Let U be a closed, bounded interval and $f, g \in \mathcal{C}^{2\neq}(U)$. Then*

$$\rho(\mathfrak{M}_f, \mathfrak{M}_g) \leqslant |U| \exp \|A_f\|_* \Big(\exp \|A_f - A_g\|_* - 1 \Big).$$

It might be proved that the right hand side of the above inequality can be majorized by the one appeared in Theorem 3.3. This theorem also has a corollary, which is a strengthening of Corollary 3.3.2, using $\|\cdot\|_*$ instead of L^1, but it will be worded nowhere in this thesis. This time the drawback discussed in the beginning of the section does not appear; cf. Example 4.2 later on. Moreover, $\|\cdot\|_* \leqslant \|\cdot\|_1$, hence the above theorem holds if one replaces $\|\cdot\|_*$ by $\|\cdot\|_1$; Remark 2.1 is applicable here.

Proof of Theorem 4.2. Fix any $(x, y, z) \in \Delta$. We would like to majorize the value of $|P_f(x, y, z) - P_g(x, y, z)|$. By Remark 2.2, let us suppose without loss of generality that

$$f(s) = \int_x^s \exp\Big(\int_x^t A_f(u)du\Big)\, dt,$$
$$g(s) = \int_x^s \exp\Big(\int_x^t A_g(u)du\Big)\, dt.$$

Then

$$\begin{aligned} f(y) - f(x) &= \int_x^y \exp\Big(\int_x^t A_f(u)du\Big)\, dt \\ &= \int_x^y \exp\Big(\int_x^t A_f(u) - A_g(u)du\Big) \exp\Big(\int_x^t A_g(u)du\Big)\, dt \\ &= \int_x^y \exp\Big(\int_x^t A_f(u) - A_g(u)du\Big)\, g'(t)dt. \end{aligned}$$

By the mean value theorem, there exists $\xi \in I$ such that

$$\begin{aligned} f(y) - f(x) &= \exp\left(\int_x^\xi A_f(u) - A_g(u)du\right) \int_x^y g'(t)dt \\ &= \exp\left(\int_x^\xi A_f(u) - A_g(u)du\right) (g(y) - g(x)). \end{aligned}$$

Similarly, there exists $\eta \in I$ such that

$$f(z) - f(x) = \exp\left(\int_x^\eta A_f(u) - A_g(u)du\right)(g(z) - g(x)).$$

Therefore,

$$P_f(x, y, z) = \exp\left(\int_\eta^\xi A_f(u) - A_g(u)du\right) P_g(x, y, z).$$

So

$$\exp\left(\|A_f - A_g\|_*\right) \cdot |P_g(x, y, z)| \leqslant |P_f(x, y, z)|$$
$$\exp\left(-\|A_f - A_g\|_*\right) \cdot |P_g(x, y, z)| \geqslant |P_f(x, y, z)|$$

But $\operatorname{sign} P_f(x, y, z) = \operatorname{sign} P_g(x, y, z)$ for any admissible x, y and z. Hence one obtains

$$|P_f(x, y, z) - P_g(x, y, z)| \leqslant |P_f(x, y, z)| \left(\exp \|A_f - A_g\|_* - 1\right). \tag{4.3}$$

Now we are going to majorize the value of $|P_f(x, y, z)|$. But

$$|P_f(x, y, z)| = \left|\frac{f(x) - f(y)}{f(x) - f(z)}\right| = \left|\frac{x - y}{x - z}\right| \frac{f'(p)}{f'(q)} \text{ for some } p,\ q \in U.$$

Moreover $|x - y| \leqslant |U|$ and

$$\left|\ln\left(\frac{f'(p)}{f'(q)}\right)\right| = \left|\int_q^p A_f(x)dx\right| \leqslant \|A_f\|_*.$$

So

$$|P_f(x, y, z)| \leqslant \frac{|U|}{|x - z|} \exp \|A_f\|_*.$$

Hence, in view of (4.3),

$$|P_f(x, y, z) - P_g(x, y, z)| \leqslant \frac{|U|}{|x - z|} \exp \|A_f\|_* \left(\exp \|A_f - A_g\|_* - 1\right).$$

Therefore, applying Corollary 4.3.3 with proper α immediately gives

$$\rho(\mathfrak{M}_f, \mathfrak{M}_g) \leqslant |U| \exp \|A_f\|_* \left(\exp \|A_f - A_g\|_* - 1\right).$$

□

Example 4.2. Let us take U, f and f_n like in Example 4.1. Then $A_f \equiv 0$ so $\|A_f\|_* = 0$,

$$\|A_{f_n} - A_f\|_* = \sup_{a,b \in [0,2\pi]} \int_a^b \frac{-n\Delta \sin nx}{n + \cos nx} = \ln\left(\frac{n+1}{n-1}\right).$$

So, by Theorem 4.2, $\rho(\mathfrak{M}_f, \mathfrak{M}_{f_n}) \leqslant \frac{4\pi}{n-1}$. This estimate is still much worse than one could expect (it is $\mathcal{O}(n^{-1})$ instead of $\mathcal{O}(n^{-2})$ ascertained in Example 4.1) but it is better than the one implied by Theorem 3.3 (a trivial one).

Part II

Hardy means

$$\sum_{n=1}^{\infty} \mathfrak{A}(a_1, \ldots, a_n) < C \cdot \sum_{n=1}^{\infty} a_n$$

Chapter 5

Introduction

Power Means have been investigated ever since their conception in the 19th century. One of the classical results concerning them is the following inequality, in its final form reproduced below due to Hardy [20] in the course of his commenting some still earlier results of Hilbert

$$\sum_{n=1}^{\infty} \mathcal{P}_p(a_1, \ldots, a_n) < (p{-}p^2)^{-1/p} \|a\|_1 \text{ for } p \in (0,1) \text{ and } a \in l_1(\mathbb{R}_+)\,.$$

Hardy realized that the constant appearing on the right hand side is not optimal and claimed that this inequality is satisfied with a constant equal $(1-p)^{-1/p}$. He also realized that the constant $(1-p)^{-1/p}$, whenever satisfied, cannot be diminished for any p.

One year later Landau ,[27], proved that, whenever $p \in (0,1)$,

$$\sum_{n=1}^{\infty} \mathcal{P}_p(a_1, \ldots, a_n) < (1-p)^{-1/p} \|a\|_1 \text{ for any } a \in l_1(\mathbb{R}_+)\,. \quad (5.1)$$

For example, upon putting $p = \frac{1}{2}$, one has

$$\sum_{n=1}^{\infty} \mathcal{P}_{1/2}(a_1, \ldots, a_n) < 4 \|a\|_1 \text{ for any } a \in l_1(\mathbb{R}_+)\,. \quad (5.2)$$

In 1923 Carleman, [8], establish the similar inequality for geometric mean

$$\sum_{n=1}^{\infty} \mathcal{P}_0(a_1,\dots,a_n) < e\,\|a\|_1 \text{ for any } a \in l_1(\mathbb{R}_+)\,.$$

In the end of 1920's Knopp, [23], fulfilled the problem of optimal constant for Power Means – he proved that the inequality (5.1) is satisfied also for $p < 0$ and, moreover, this constant is optimal.

Let me note that these result has their integral, closely related, version

$$\int_0^\infty \left(\frac{1}{x}\int_0^x f(t)dt\right)^q dx \leqslant \left(\frac{q}{q-1}\right)^q \int_0^\infty f(x)^q dx$$

for $q > 1$ and $f \in L^q(\mathbb{R}_+,\mathbb{R}_+)$, which is also known as Hardy inequality. Connection between this inequality and a classical Hardy problem becomes move obvious after substitutions $q \leftarrow 1/p$ and $f(x) \leftarrow [g(x)]^p$, where $p \in (0,1)$.

Nowadays these results are few of many in the emerging "theory of Hardy means". In this way, except of finding a means which sattisfiend are Hardy, there exists a pararell, widely develop, area of research; namely there are significants resuts (e.g. [22; 44; 28]) where the following generalization is consider

$$\sum_{n=1}^{\infty} \nu_n \mathcal{P}_p(a_1,\dots,a_n) < \sum_{n=1}^{\infty} \mu_n a_n \text{ for any } a \in l_1(\mathbb{R}_+)\,,$$

where

$$\nu_n \geqslant 1 \text{ and } \begin{cases} \mu_n \leqslant (1-p)^{-1/p} & p \neq 0, \\ \mu_n \leqslant e & p = 0. \end{cases}$$

Such a result were optained few times for example assuming $p = 0$ one could put

Kaluza and Szegö 1927 [22]	$\nu_n = n \cdot (e^{1/n} - 1);\ \mu_n = e$
Redheffer 1983 [44, p. 138]	$\nu_n = 1 + \frac{1}{2n};\ \mu_n = e$
Redheffer 1983 [44, p. 138]	$\nu_n = 1;\ \mu_n = e \cdot (1 - \frac{1}{2n} + \mathcal{O}(\frac{1}{n^2}))$
Love 1991 [28, §4]	$\nu_n = 1;\ \mu_n = e \cdot \sum_{m=n}^{\infty} \frac{n - 1/2}{m^2}$

In the year 2004 Páles and Persson, [36], put forward the following definition. [However it had been felt in the air since the 1920s.] Let $I \subset \mathbb{R}_+$ be an interval, $\inf I = 0$. A mean $\mathfrak{A}\colon \bigcup_{n=1}^{\infty} I^n \to \mathbb{R}_+$ (no additional assumption is given) is *Hardy* whenever there exists a (positive) constant C such that for any $a \in l^1(I)$

$$\sum_{n=1}^{\infty} \mathfrak{A}(a_1, \dots, a_n) < C \sum_{n=1}^{\infty} a_n \,.$$

This definition was introduced under some additional assumptions on $\mathfrak{A}$, however they could be easily omitted. (Most often some additional properties are being assumed for a function to call it a mean; e. g., its value should lie between the minimal and maximal entry of every vector of arguments, like it was done, e.g., in [36]. In this part, however – we reiterate – such extra assumptions are neither needed nor made.)

These authors proposed in [36] certain conditions sufficient for a mean to be Hardy. Those conditions are relatively mild and are satisfied by the means in a considerable number of families.

Hence it is natural to ask what other means are Hardy. In fact, this question was extensively dealt with decades before the formal definition appeared. The detailed history of the events related to, and facts implied by above inequalities is sketched in catching surveys [42; 13; 33], and in a recent book [26].

Unfortunately, for many families of means the problem if they are Hardy remains open. Such a problem for the two-parameter family of Gini means was, for instance, considered in [36], where many special subcases were solved. This open problem was explicitly worded three years later in [26, p. 89]. – It will be solved in chapter 7.

Another interesting family which could be considered in this context is a natural generalization of Power Means proposed in 1971 by Carlson, Meany and Nelson. For any fixed parameters $k \in \mathbb{N}$, s, $q \in \mathbb{R}$ and positive vector $(v_1, \ldots, v_n)$, $n \geqslant k$, they take the q-th power means of all possible k-tuples $(v_{i_1}, \ldots, v_{i_k})$, $1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n$, and then calculate the s-th power mean of the resulting vector of length $\binom{n}{k}$ (for the purpose of this thesis, we will denote it by $\tilde{\mathcal{P}}_{k,s,q}$).

In chapter 6, we will be working towards a complete answer to the question when these means satisfy inequalities resembling the classical Hardy inequality. Within a large part of the parameter space (k, s, q) the answer is definitive.

Chapter 6

Generalized power means

We discuss properties of a natural generalization of Power Means, $\tilde{\mathcal{P}}_{k,s,q}$ (see below for the definition) proposed in 1971 by Carlson, Meany and Nelson. We work towards a complete answer to the question when these means ar Hardy. Within a large part of the parameter space (k, s, q) the answer is definitive.

This chapter is based on the paper [P3].

6.1 Basic definition

We are going to analyse certain multi-parameter family of means. Namely, in 1971 Carlson, Meany and Nelson [9], among other things, proposed the following family, which in one time encompasses Power Means, Hamy means and Hayashi means:

$$\tilde{\mathcal{P}}_{k,s,q}(v_1 \dots v_n) := \begin{cases} \mathcal{P}_s\Big(\mathcal{P}_q(v_{i_1}, \dots, v_{i_k}) \colon 1 \leqslant i_1 < \dots < i_k \leqslant n\Big) & \text{if } k \leqslant n\,, \\ \mathcal{P}_q(v_1, \dots, v_n) & \text{if } k > n\,. \end{cases}$$

Those authors were interested in certain inequalities binding the means $\tilde{\mathcal{P}}_{k,s,q}$ when the order of parameters s and q was being

reversed.
These means are analysed here from the point of view of being or not being Hardy. Namely, we are going to prove that these means are Hardy for

$$(s,q) \in \Big((-\infty,1)\times\mathbb{R}\Big) \cup \Big((\mathbb{R}_- \cup \{0\})\times\{1\}\Big) \text{ and any } k \geqslant 2$$

(see Figure 6.1 below for a better visualisation).

Let us note that, by [46], there hold the following inequalities

$$\tilde{\mathcal{P}}_{k,s,q} \leqslant \tilde{\mathcal{P}}_{k,t,p} \text{ for } s \leqslant t \text{ and } q \leqslant p\,, \tag{6.1}$$

$$\tilde{\mathcal{P}}_{k,s,q} \leqslant \tilde{\mathcal{P}}_{k-1,s,q} \text{ for } s > q\,. \tag{6.2}$$

6.2 Main result

In our main Theorem 6.1 we are going to prove a seemingly isolated fact that $\tilde{\mathcal{P}}_{2,1,0}$ is a Hardy mean. Obviously, all means majorized by some Hardy mean (or, more generally, majorized up to some constant coefficient) are Hardy, too. Therefore, one time $\tilde{\mathcal{P}}_{2,1,0}$ being Hardy, the inequalities (6.2) and (6.1) imply that $\tilde{\mathcal{P}}_{k,s,q}$ are Hardy, too, for a vast family of parameters. This is precisely worded in Corollary 6.3.1 below.

That corollary is a fairy wide extension of our 'trendsetting' Theorem 6.1. Its Hardy-negative part subsumes regions in the parameter plane (s,q) which have until recently seemed to be a kind of challenge – see for instance the second item in Corollary 6.3.1, and especially the subregion $s \geqslant k$, $q \leqslant 0$ encompassed by that item.

Theorem 6.1. $\tilde{\mathcal{P}}_{2,1,0}$ *is a Hardy mean and*

$$\sum_{n=1}^{\infty} \tilde{\mathcal{P}}_{2,1,0}(a_1,\dots,a_n) < 4\,\|a\|_1 \text{ for every } a \in l^1(\mathbb{R}_+)\,.$$

Proof. We will show that $\tilde{\mathcal{P}}_{2,1,0}$ is majorized by $\mathcal{P}_{1/2}$. Indeed,

$$\begin{aligned}
\tilde{\mathcal{P}}_{2,1,0}(a_1,\ldots,a_n) &= \binom{n}{2}^{-1} \sum_{1\leqslant i<j\leqslant n} \sqrt{a_i a_j} \\
&= \tfrac{1}{n(n-1)} \sum_{1\leqslant i<j\leqslant n} 2\sqrt{a_i a_j} \\
&= \tfrac{n}{n-1} \left(\left(\tfrac{1}{n} \sum_{i=1}^{n} \sqrt{a_i} \right)^2 - \frac{1}{n^2} \sum_{i=1}^{n} a_i \right) \\
&= \tfrac{n}{n-1} \left(\mathcal{P}_{1/2}(a_1,\ldots,a_n) - \tfrac{1}{n}\mathcal{P}_1(a_1,\ldots,a_n) \right) \\
&\leqslant \tfrac{n}{n-1} \left(\mathcal{P}_{1/2}(a_1,\ldots,a_n) - \tfrac{1}{n}\mathcal{P}_{1/2}(a_1,\ldots,a_n) \right) \\
&= \mathcal{P}_{1/2}(a_1,\ldots,a_n)\,. \qquad (6.3)
\end{aligned}$$

Hence, by (5.2), one obtains

$$\sum_{n=1}^{\infty} \tilde{\mathcal{P}}_{2,1,0}(a_1,\ldots,a_n) \leqslant \sum_{n=1}^{\infty} \mathcal{P}_{1/2}(a_1,\ldots,a_n) < 4\,\|a\|_1\,.$$

□

Remark. The constant 4 in the above theorem cannot be diminished. Indeed, upon taking $a_n = \frac{1}{n}$, a simple calculation yields

$$\lim_{n\to\infty} a_n^{-1} \tilde{\mathcal{P}}_{2,1,0}(a_1,\ldots,a_n) = 4\,.$$

Then the machinery originally devised for the Power Means in [21, pp. 241–242] becomes applicable. It gives, by taking N arbitrary large and considering the auxiliary sequence

$$(a_1,\ldots,a_N,(N+1)^{-2},(N+2)^{-2},(N+3)^{-2},\ldots),$$

that the constant cannot be smaller than 4.

Theorem 6.2. *Let $k \in \mathbb{N}_+$. Then $\tilde{\mathcal{P}}_{k,k,-\infty}$ is not a Hardy mean.*

Proof. Let us take any decreasing sequence $a \in l^1(\mathbb{R}_+)$. For any $n \geqslant k$ one obtains

$$\begin{aligned}
&\tilde{\mathcal{P}}_{k,k,-\infty}(a_1,\dots,a_n)\\
&=\left(\binom{n}{k}^{-1}\sum_{1\leqslant i_1<\ldots<i_k\leqslant n}\min(a_{i_1},\dots,a_{i_k})^k\right)^{1/k}\\
&>\left(\binom{n}{k}^{-1}\min(a_1,\dots,a_k)^k\right)^{1/k}\\
&=\binom{n}{k}^{-1/k}a_k\\
&>n^{-1}a_k.
\end{aligned}$$

Hence $\sum_{n=1}^{\infty}\tilde{\mathcal{P}}_{k,k,-\infty}(a_1,\dots,a_n)=+\infty$.

□

6.3 Discussion of parameters

We know that for any fixed $q > 0$ the inequality

$$\begin{aligned}
\mathcal{P}_q(v_1,\dots,v_k) &= \left(\tfrac{1}{k}\sum v_i^q\right)^{1/q} > \left(\tfrac{1}{k}\max(v_1^q,\dots,v_n^q)\right)^{1/q}\\
&= k^{-1/q}\max(v_1,\dots,v_k) = C(k,q)\max(v_1,\dots,v_k)
\end{aligned}$$

holds for any $v \in \mathbb{R}_+^k$, with $C(k,q) := k^{-1/q}$. In particular, for any $r \in \mathbb{R}\cup\{\pm\infty\}$

$$\mathcal{P}_q(v_1,\dots,v_k) > C(k,q)\mathcal{P}_r(v_1,\dots,v_k) \tag{6.4}$$

(cf. also [6, p. 237]). The above inequalities are instrumental in proving the following

Corollary 6.3.1. *For any* $k \geqslant 2$

- $\tilde{\mathcal{P}}_{k,s,q}$ *is a Hardy mean for no* $s \geqslant 1$ *and* $q > 0$,
- $\tilde{\mathcal{P}}_{k,s,q}$ *is a Hardy mean for no* $s \geqslant k$ *and* $q \in \mathbb{R} \cup \{\pm\infty\}$,
- $\tilde{\mathcal{P}}_{k,1,q}$ *is a Hardy mean for any* $q \leqslant 0$,
- $\tilde{\mathcal{P}}_{k,s,q}$ *is a Hardy mean for any* $s < 1$ *and* $q \in \mathbb{R} \cup \{\pm\infty\}$

(see Figure 6.1).

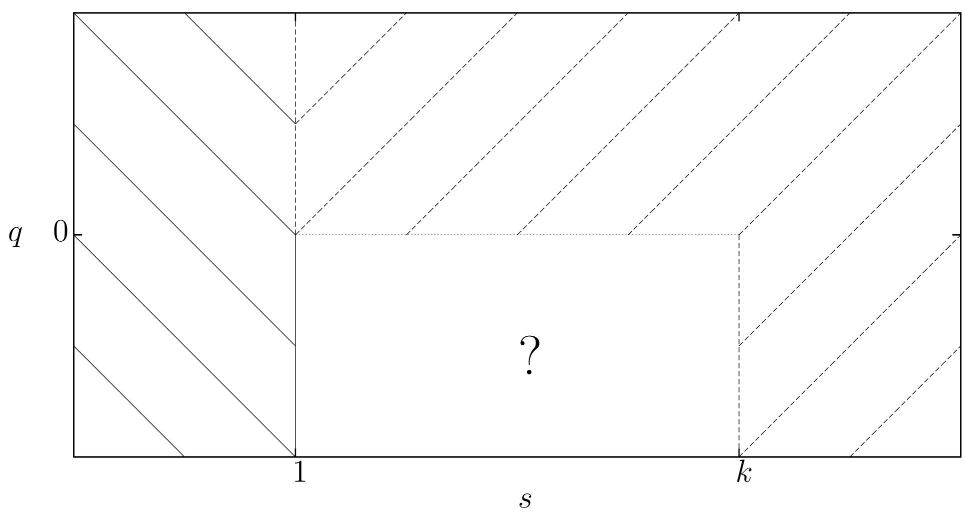

Figure 6.1: Space of parameters, for which the mean $\tilde{\mathcal{P}}_{k,s,q}$ is Hardy (solid lines), and for which it is not Hardy (dashed lines); k is fixed.

Proof. Let us recall that the length of vectors in $\mathcal{P}_q$ in the definition of $\tilde{\mathcal{P}}_{k,s,q}$ is fixed (and equal to k, whenever $k \leqslant n$). Moreover, if a mean could be majorized by some Hardy mean up to a constant coeficient, then it is Hardy, too. Therefore the use of (6.4) is very natural to in the investigation of behaviour of $\tilde{\mathcal{P}}_{k,s,q}$ while the parameter q is changed.

First item follows from the fact that $\tilde{\mathcal{P}}_{k,1,1}$ is an arithmetic mean. So it is not Hardy. But $\tilde{\mathcal{P}}_{k,s,q} \geqslant \tilde{\mathcal{P}}_{k,1,q} \geqslant C\tilde{\mathcal{P}}_{k,1,1}$ for some constant C, hence $\tilde{\mathcal{P}}_{k,s,q}$ is not Hardy, too.

Second item is an immediate corollary from Theorem 6.2. Third one is implied by Theorem 6.1 and (6.1).

Fourth item is being proved in two steps. First, with no loss of generality we may assume that $s \in (0,1)$. We know that $\tilde{\mathcal{P}}_{k,s,s} = \mathcal{P}_s$, so it is a Hardy mean. Second, applying (6.4), one gets $\tilde{\mathcal{P}}_{k,s,s} > C(k,s)\,\tilde{\mathcal{P}}_{k,s,q}$. So $\tilde{\mathcal{P}}_{k,s,q}$ is Hardy as well. □

Corollary 6.3.2. *For any $k \geqslant 2$ the Hamy mean $\mathfrak{ha}^{[k]} := \tilde{\mathcal{P}}_{k,1,0}$ (cf. [18], [6, pp. 364–365] for more details) is a Hardy mean.*

Corollary 6.3.3. *For any $k \geqslant 2$ the Hayashi mean $\mathfrak{hy}^{[k]} := \tilde{\mathcal{P}}_{k,0,1}$ (cf. [19] and [6, pp. 365–366] for more details) is a Hardy mean.*

6.3.1 Remaining cases

The problem whether $\tilde{\mathcal{P}}_{k,s,q}$ is a Hardy mean for $k \geqslant 2$, $s \in (1,k)$ and $q \leqslant 0$ remains open (see the central part in Figure 6.1).

Chapter 7

Certain negative results

We give a new necessary condition for a mean to be a Hardy mean. This condition is then applied to completely characterize the Hardy property among: (1) the Gini means, (2) Gaussian products of power means, and (3) symmetric polynomial means.

This chapter is based on the paper [P4].

7.1 Main result

Many means were shown to be Hardy in the course of years. As for the present chapter, there are obtained some results going in the opposite direction. In fact, we are going to give a *necessary* condition for a mean to be Hardy. Namely,

Theorem 7.1. *Let* $I \subset \mathbb{R}_+$ *be an interval,* $\inf I = 0$*. Let* $\mathfrak{A}$ *be a mean defined on* I *and* $(a_n)_{n=1}^{\infty}$ *be a sequence of numbers in* I *satisfying* $\sum a_n = +\infty$*.*
If $\lim a_n^{-1}\mathfrak{A}(a_1, \dots a_n) = +\infty$ *then* $\mathfrak{A}$ *is not Hardy.*

Proof. Suppose conversely that $\mathfrak{A}$ is a Hardy mean with a constant $C > 0$.

We will show that the constant C is not good for $\mathfrak{A}$. Let us pick n_0 and $n_1 > n_0$ such that

$$a_n^{-1}\mathfrak{A}(a_1, \dots a_n) > 2\,C \text{ for every } n > n_0, \tag{7.1}$$

$$\sum_{n=n_0+1}^{n_1-1} a_n > \sum_{n=1}^{n_0} a_n. \tag{7.2}$$

Define a new sequence (b_n), $b_n = \begin{cases} a_n & \text{, for } n \leqslant n_1, \\ a_{n_1} 2^{-n} & \text{, for } n > n_1 \end{cases}$.
The constant C is not good for the sequence $(b_n) \in l^1(I)$. Indeed, one has

$$\begin{aligned} 2\sum_{n=n_0+1}^{n_1} a_n &= \sum_{n=n_0+1}^{n_1-1} a_n + a_{n_1} + \sum_{n=n_0+1}^{n_1} a_n \\ &> \sum_{n=1}^{n_0} a_n + \sum_{n=n_1+1}^{\infty} a_{n_1} 2^{-n} + \sum_{n=n_0+1}^{n_1} a_n \qquad \text{by (7.2)} \\ &= \sum_{n=1}^{n_0} b_n + \sum_{n=n_1+1}^{\infty} b_n + \sum_{n=n_0+1}^{n_1} b_n \\ &= \sum_{n=1}^{\infty} b_n. \end{aligned} \tag{7.3}$$

Whence

$$\begin{aligned} \sum_{n=1}^{\infty} \mathfrak{A}(b_1, \dots, b_n) &> \sum_{n=n_0+1}^{n_1} \mathfrak{A}(b_1, \dots, b_n) \\ &= \sum_{n=n_0+1}^{n_1} \mathfrak{A}(a_1, \dots, a_n) \\ &> 2\,C \sum_{n=n_0+1}^{n_1} a_n \qquad \text{by (7.1)} \\ &> C \sum_{n=1}^{\infty} b_n \qquad \text{by (7.3).} \end{aligned}$$

□

7.2 Applications

We are going to show how Theorem 7.1 can be applied to three fairly known families of means. Necessary and sufficient conditions for a mean to be Hardy will be presented for each family. The relevant proofs will be given in the following section.

7.2.1 Gaussian product of Power Means

Power means were generalized in different ways by many authors (cf., e. g., [6, chap. III-VI] for details). In particular, in 1947, Gustin [17] proposed an extension of a famous Gauss' concept of the arithmetic-geometric mean, recalled in detail in [14, pp. 361–403].

Recalling from Overwiew, for $\lambda = (\lambda_0, \ldots, \lambda_p) \in \mathbb{R}^{p+1}$ and v – all-positive-components vector, Gustin defines (in fact, in a more setup) the sequence of length-$(p+1)$ vectors (except of $v^{(0)}$):

$$\begin{aligned} v^{(0)} &= v, \\ v^{(i+1)} &= \Big(\mathcal{P}_{\lambda_0}(v^{(i)}), \mathcal{P}_{\lambda_1}(v^{(i)}), \ldots, \mathcal{P}_{\lambda_p}(v^{(i)})\Big), \quad i = 0,\, 1,\, 2, \ldots. \end{aligned}$$

He proved that, for each $0 \leqslant k \leqslant p$, the limit $\lim_{i\to\infty} v_k^{(i)}$ exists and does not depend on k. This common limit is denoted by $\mathcal{P}_{\lambda_0} \otimes \cdots \otimes \mathcal{P}_{\lambda_p}(v)$. Because of various Gauss' results on $\mathcal{P}_1 \otimes \mathcal{P}_0$, such means are called Gaussian Means (or, more descriptively, the Gaussian products of Power Means).

Note that, for any $\lambda \in \mathbb{R}^{p+1}$, the Gaussian product applied to a vector of length $p+1$,

$$K := \mathcal{P}_{\lambda_0} \otimes \cdots \otimes \mathcal{P}_{\lambda_p} \colon \mathbb{R}_+^{p+1} \to \mathbb{R}_+$$

is the *only* function satisfying

$$K(v) = K\Big(\mathcal{P}_{\lambda_0}(v),\, \mathcal{P}_{\lambda_1}(v), \ldots,\, \mathcal{P}_{\lambda_p}(v)\Big), \quad v \in \mathbb{R}_+^{p+1}, \tag{7.4}$$

$$\min(v) \leqslant K(v) \leqslant \max(v), \quad v \in \mathbb{R}_+^{p+1}. \tag{7.5}$$

We are going to give a necessary and sufficient condition for a Gaussian mean to be Hardy. More precisely, we are going to prove

Theorem 7.2. *Let* $p \in \mathbb{N}$ *and* $\lambda \in \mathbb{R}^{p+1}$. *Then* $\mathcal{P}_{\lambda_0} \otimes \mathcal{P}_{\lambda_1} \otimes \cdots \otimes \mathcal{P}_{\lambda_p}$ *is Hardy if and only if* $\max(\lambda_0, \ldots, \lambda_p) < 1$.

7.2.2 Gini Means

Recall, gini means is a family defined on all-postitve-components vectors by

$$\mathfrak{G}_{p,q}(a_1, \ldots, a_n) := \begin{cases} \left(\frac{\sum_{i=1}^n a_i^p}{\sum_{i=1}^n a_i^q} \right)^{1/(p-q)} & \text{if } p \neq q\,, \\ \exp\left(\frac{\sum_{i=1}^n a_i^p \ln a_i}{\sum_{i=1}^n a_i^p} \right) & \text{if } p = q\,. \end{cases}$$

Reiterating, for $q = 0$ one gets here the p-th power mean. For this family many problems are still open. One of them is to ascertain the set of parameters (p, q) such that $\mathfrak{G}_{p,q}$ is Hardy. A recent approach (dating from 2004), due to Páles and Persson, states the following

Proposition 7.2.1 ([36], Theorem 2)**.** *Let* $p, q \in \mathbb{R}$. *If* $\mathfrak{G}_{p,q}$ *is a Hardy mean, then*

$$\min(p, q) \leqslant 0 \textit{ and } \max(p, q) \leqslant 1.$$

Conversely, if

$$\min(p, q) \leqslant 0 \textit{ and } \max(p, q) < 1,$$

then $\mathfrak{G}_{p,q}$ *is a Hardy mean.*

The authors put also forward a conjecture [36, Open Problem 3] that the sufficient condition in the proposition above is also a necessary one. In what follows we will justify this conjecture. Namely, we will prove the following

Theorem 7.3. *Let* $p, q \in \mathbb{R}$. *Then* $\mathfrak{G}_{p,q}$ *is a Hardy mean if and only if* $\min(p, q) \leqslant 0$ *and* $\max(p, q) < 1$.

7.2.3 Symmetric Polynomial Means

This family arose in the study of algebraic equations. Many results have been given already by Newton and Campbell (cf. [6, chap. V] for details). Namely for r, $n \in \mathbb{N}$, $n \geqslant r$ and an all-positive-components vector a, let us define

$$\mathfrak{S}_r(a_1, \dots, a_n) = \left(\binom{n}{r}^{-1} \sum_{1 \leqslant k_1 < k_2 < \dots < k_r \leqslant n} a_{k_1} a_{k_2} \cdots a_{k_r} \right)^{1/r}.$$

Assuming that the mean does not exceed the maximal plugged in argument, the fact whether it is Hardy does not depend on any finite number of the initial summands in the definition of Hardy mean. However, in order to have a definition of a mean fulfilled, it should be well defined for $n < r$, too. This might be done in any way; let us simply assume that

$$\mathfrak{S}_r(a_1, \dots, a_n) = \sqrt[n]{a_1 \cdots a_n} \text{ for } n < r.$$

We will prove the following

Theorem 7.4. $\mathfrak{S}_r$ *is Hardy for no* $r \in \mathbb{N}$.

In view of the equality $\mathfrak{S}_r = \tilde{\mathcal{P}}_{r,r,0}$ this theorem was obtained in more general context in the previous chapter. However, it remains a handy example of Theorem 7.1.

7.3 Proofs

In proofs used will be the elementary estimations

$$\sum_{i=1}^{n} i^k \leqslant n^{k+1} \quad \text{for every } n \in \mathbb{N}, \tag{7.6}$$

$$\sum_{i=1}^{n} \tfrac{1}{i} \geqslant \ln n \quad \text{for every } n \in \mathbb{N}. \tag{7.7}$$

7.3.1 Proof of Theorem 7.2

Prior to the proof, let us note that if $\lambda_i \leqslant \lambda_i'$ for every $i = 0, 1, 2, \dots, p$, then

$$\mathcal{P}_{\lambda_0} \otimes \mathcal{P}_{\lambda_1} \otimes \cdots \otimes \mathcal{P}_{\lambda_p} \leqslant \mathcal{P}_{\lambda_0'} \otimes \mathcal{P}_{\lambda_1'} \otimes \cdots \otimes \mathcal{P}_{\lambda_p'}. \tag{7.8}$$

Therefore, the ($\Leftarrow$) part is simply implied by the fact that the mean $\mathcal{P}_{\lambda_0} \otimes \mathcal{P}_{\lambda_1} \otimes \cdots \otimes \mathcal{P}_{\lambda_p}$ is majorized by a Hardy mean $\mathcal{P}_{\max(\lambda_0,\dots,\lambda_p)}$, so it is Hardy too (recall that $\max(\lambda_0, \dots, \lambda_p) < 1$).

Now we are going to prove the ($\Rightarrow$) implication. One may assume that $i \mapsto \lambda_i$ is non-increasing. Moreover, by (7.8), having $\lambda_0 \geqslant 1$ we estimate from below: λ_0 by 1 and $\lambda_1, \lambda_2, \dots, \lambda_p$ by $-\lambda$ for certain $\lambda > 0$ and we are going to prove that

$$\mathfrak{A} := \mathcal{P}_1 \otimes \underbrace{\mathcal{P}_{-\lambda} \otimes \cdots \otimes \mathcal{P}_{-\lambda}}_{p}$$

is not Hardy.

To this end, let us consider a two variable function $F(a, b) := \mathfrak{A}(a, \underbrace{b, \dots, b}_{p})$ and fix $\theta > 1$. Then, using monotonicity of $\mathfrak{A}$ with respect to each variable and inequality (7.4), for $a > \theta b$,

$$\begin{aligned}
F(a,b) &= \mathfrak{A}\left(\frac{a+pb}{p+1}, \underbrace{\left(\frac{p+1}{a^{-\lambda}+pb^{-\lambda}}\right)^{1/\lambda}, \dots, \left(\frac{p+1}{a^{-\lambda}+pb^{-\lambda}}\right)^{1/\lambda}}_{p}\right)\\
&\geqslant \mathfrak{A}\left(\tfrac{1}{p+1}a, \underbrace{\left(\frac{p+1}{(\theta b)^{-\lambda}+pb^{-\lambda}}\right)^{1/\lambda}, \dots, \left(\frac{p+1}{(\theta b)^{-\lambda}+pb^{-\lambda}}\right)^{1/\lambda}}_{p}\right)\\
&= \mathfrak{A}\left(\tfrac{1}{p+1}a, \underbrace{\left(\frac{p+1}{\theta^{-\lambda}+p}\right)^{1/\lambda} b, \dots, \left(\frac{p+1}{\theta^{-\lambda}+p}\right)^{1/\lambda} b}_{p}\right)\\
&= F\left(\tfrac{1}{p+1}a, \left(\frac{p+1}{\theta^{-\lambda}+p}\right)^{1/\lambda} b\right).
\end{aligned} \tag{7.9}$$

Introduce two more mappings

$$\tau\colon (a,b) \mapsto \left(\tfrac{1}{p+1}a, \left(\frac{p+1}{\theta^{-\lambda}+p}\right)^{1/\lambda} b\right),$$

$$G\colon (a,b) \mapsto \left(a^{\log_{p+1}\left(\frac{p+1}{\theta^{-\lambda}+p}\right)} b^{\lambda}\right)^{1/\left(\lambda+\log_{p+1}\left(\frac{p+1}{\theta^{-\lambda}+p}\right)\right)}.$$

Inequality (7.9) assumes now a compact form

$$F\circ\tau(a,b) \leqslant F(a,b) \qquad \text{for any} \qquad a,\, b,\, a>\theta b. \tag{7.10}$$

We will prove in a moment a technical, if important inequality

$$F(a,b) > \tfrac{1}{\theta(p+1)} G(a,b) \qquad \text{for any} \qquad a,\, b\, a>b. \tag{7.11}$$

Postponing the proof of this inequality, using it altogether with: –the fact that the mean $\mathfrak{A}$ is greater then its minimal argument, –homogeneity of F, –inequality (7.7), one obtains

$$\begin{aligned}
(\tfrac{1}{n})^{-1}\mathfrak{A}(1,\tfrac12,\tfrac13,\dots,\tfrac1n) &= nF(\mathcal{P}_1(1,\tfrac12,\tfrac13,\dots,\tfrac1n),\mathcal{P}_{-\lambda}(1,\tfrac12,\tfrac13,\dots,\tfrac1n))\\
&\geqslant nF\left(\frac{\ln n}{n},\frac{1}{n}\right)\\
&= F(\ln n, 1)\\
&\geqslant \tfrac{1}{\theta(p+1)} G(\ln n, 1)\\
&\geqslant \tfrac{1}{\theta(p+1)} (\ln n)^{\frac{\log_{p+1}\left(\frac{p+1}{\theta^{-\lambda}+p}\right)}{\lambda+\log_{p+1}\left(\frac{p+1}{\theta^{-\lambda}+p}\right)}}.
\end{aligned}$$

But, for any $\theta > 1$ and $\lambda > 0$, the exponent in the right-most term is positive. Whence this term tends to infinity when $n \to +\infty$. So, by Theorem 7.1, $\mathfrak{A}$ is not Hardy.

Remark. Often the right-most term tends to infinity very slowly. For example ($\lambda = 5$ and $p = 3$) one obtains (taking $\theta = \frac32$)

$$n\cdot \mathcal{P}_1 \otimes \mathcal{P}_{-5} \otimes \mathcal{P}_{-5} \otimes \mathcal{P}_{-5}(1,\tfrac12,\tfrac13,\dots,\tfrac1n) > \tfrac16 (\ln n)^{0.0341}.$$

The right hand side exceeds 1 only for $n > 10^{2.86\cdot 10^{22}}$.

Proof of the inequality (7.11)

There clearly hold

- $G(a,b) \in (\min(a,b), \max(a,b))$,
- $F(a,b) \in (\min(a,b), \max(a,b))$,
- $G \circ \tau(a,b) = G(a,b)$.

The case when $\frac{a}{b} < \theta(p+1)$ is simply implied by the first and second property.

Otherwise, let $a_0 = a$, $b_0 = b$, $(a_{i+1}, b_{i+1}) = \tau(a_i, b_i)$. By the definition of τ, $a_n \searrow 0$ and $b_n \nearrow +\infty$. Denote by N the smallest natural number such that $a_N \leqslant \theta b_N$. Obviously, $a_{N-1} > \theta b_{N-1}$. Thus

$$\begin{aligned} a_N = \tfrac{1}{p+1} a_{N-1} > \tfrac{\theta}{p+1} b_{N-1} &= \tfrac{\theta}{p+1} \left(\frac{\theta^{-\lambda} + p}{p+1} \right)^{1/\lambda} b_N \\ &> \tfrac{\theta}{p+1} \left(\frac{\theta^{-\lambda} + \theta^{-\lambda} p}{p+1} \right)^{1/\lambda} b_N = \tfrac{1}{p+1} b_N. \end{aligned}$$

Therefore $\frac{1}{p+1} b_N < a_N \leqslant \theta b_N$, and so

$$\min(a_N, b_N) > \tfrac{1}{\theta(p+1)} \max(a_N, b_N).$$

Hence, using inequality (7.10) and the facts under • above, one gets

$$\begin{aligned} F(a,b) = F(a_0, b_0) &\geqslant F \circ \tau^N (a_0, b_0) \\ &= F(a_N, b_N) \\ &\geqslant \min(a_N, b_N) \\ &> \tfrac{1}{\theta(p+1)} \max(a_N, b_N) \\ &> \tfrac{1}{\theta(p+1)} G(a_N, b_N) \\ &= \tfrac{1}{\theta(p+1)} G \circ \tau^N (a_0, b_0) \\ &= \tfrac{1}{\theta(p+1)} G(a_0, b_0) = \tfrac{1}{\theta(p+1)} G(a,b). \end{aligned}$$

7.3.2 Proof of Theorem 7.3

The ($\Leftarrow$) implication is implied by Proposition 7.2.1.

Working towards the ($\Rightarrow$) implication, let us assume that $\max(p,q) \geqslant 1$. We will prove that $\mathfrak{G}_{p,q}$ is not Hardy. By [6, pp.249–250], we know that

$$\begin{aligned} &\mathfrak{G}_{p,q} \leqslant \mathfrak{G}_{p',q'} \qquad \text{for} \qquad p \leqslant p' \text{ and } q \leqslant q', \\ &\mathfrak{G}_{p,q} = \mathfrak{G}_{q,p} \qquad \text{for any} \qquad p,\, q \in \mathbb{R}. \end{aligned}$$

Whence, most likely it was done at the beginning of section 7.3.1, we can assume without loss of generality that $p = 1$ and $q < 0$.

Upon taking the sequence $a_i = \frac{1}{i}$, by (7.6) and (7.7), one obtains

$$\begin{aligned} a_n^{-1}\mathfrak{G}_{1,q}(a_1,\dots,a_n) &= n\left(\frac{\sum_{i=1}^n \frac{1}{i}}{\sum_{i=1}^n i^{-q}}\right)^{1/(1-q)} \\ &\geqslant n\left(\frac{\ln n}{n^{1-q}}\right)^{1/(1-q)} = (\ln n)^{1/(1-q)}. \end{aligned}$$

But $(\ln n)^{1/(1-q)} \to +\infty$ so, by Theorem 7.1, the Gini mean $\mathfrak{G}_{1,q}$ is Hardy for no $q < 0$. Hence $\max(p,q) \geqslant 1$ implies $\mathfrak{G}_{p,q}$ not being Hardy.

7.3.3 Proof of Theorem 7.4

Let us fix $r \in \mathbb{N}$ and take the sequence $a_n = \frac{1}{n}$. Then, for $n > r$, one obtains

$$
\begin{aligned}
& a_n^{-1}\mathfrak{S}_r(a_1, \ldots, a_n) \\
&= n\left(\binom{n}{r}^{-1} \sum_{1\leqslant k_1<k_2<\ldots<k_r\leqslant n} a_{k_1}a_{k_2}\cdots a_{k_r}\right)^{1/r} \\
&\geqslant n\binom{n}{r}^{-1/r}\Big(a_1a_2\cdots a_{r-1}\cdot(a_r+a_{r+1}+\ldots+a_n)\Big)^{1/r} \\
&\geqslant n\left(\frac{(n-r)^r}{r!}\right)^{-1/r}\left(\frac{1}{(r-1)!}\right)^{1/r}\cdot(\ln n-(1+\tfrac{1}{2}+\ldots+\tfrac{1}{r-1}))^{1/r} \\
&= \frac{n}{n-r}r^{1/r}\cdot(\ln n-(1+\tfrac{1}{2}+\ldots+\tfrac{1}{r-1}))^{1/r}
\end{aligned}
$$

By Theorem 7.1, upon taking $n \to +\infty$, the above inequality implies $\mathfrak{S}_r$ not being Hardy.

Bibliography

Bibliography

[1] P. R. Bessac and J. Pečarič, *On Jensen's inequality for convex functions II*, J. Math. Anal. Appl. **118** (1986), 61–65.

[2] D. Besso, *Teoremi elementari sui massimi i minnimi*, Annuari Ist. Tech. Roma (1879), 7–24.

[3] C. E. Bonferroni, *La media esponenziale in matematica finanziaria*, Annuario R. Ist. Sc. Econ. Comm. Bari (1923–24), 1–14.

[4] C. E. Bonferroni, *Della vita matematica come media esponenziale*, Annuario R. Ist. Sc. Econ. Comm. Bari (1924–25), 3–16.

[5] C. E. Bonferroni, *Sulle medie dedotte de funzioni concave*, Giorn. Mat. Finanz. **1** (1927), 13–24.

[6] P. S. Bullen, *Handbook of Means and Their Inequalities*, Mathematics and Its Applications, vol. **560**, Kluwer Acad. Publ., Dordrecht 2003.

[7] P. S. Bullen, D. S. Mitrinović and P. M. Vasić, *Means and Their Inequalities*, D. Reidel Publishing Company, Dordrecht 1987.

[8] T. Carleman, *Sur les fonctions quasi-analytiques*, Conférences faites au cinquième congres des mathématiciens Scandinaves, Helsinki (1923), 181-196.

[9] B. C. Carlson, R. K. Meany, S. A. Nelson, *Mixed arithmetic and geometric means*, Pacif. J. Math. **38** (1971), 343–349.

[10] G. T. Cargo and O. Shisha, *On comparable means*, Pacific J. Math. **14** (1964), 1053–1058.

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

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

[13] J. Ducan, C. M. McGregor, *Carleman's Inequality*, Amer. Math. Monthly **110(5)** (2003), 424–431.

[14] C. F. Gauss, *Werke* **3**, Göttingen-Leipzig, 1866.

[15] C. Gini, *The contributions of Italy to modern statistical methods*, J. Roy. Statist. Soc. **89** (1926), 703–721.

[16] C. Gini, *Di una formula compressiva delle medie*, Metron **13** (1938), 3–22.

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

[18] M. Hamy, *Sur le théorème de la moyenne*, Bull. Sci. Math. **14** (1890), 103–104.

[19] T. Hayashi, *Un théorème relatif aux valeurs moyennes*, Nouv. Ann. Math. **5** (1905), 355–357.

[20] G. H. Hardy, *Note on a theorem of Hilbert*, Math. Zeitschrift **6** (1920), 314–317.

[21] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge University Press, Cambridge 1934.

[22] T. Kaluza, G. Szegö, *Über Reihen mit lauter positiven Gliedern*, J. London Math. Soc. **2** (1927), 266–272.

[23] K. Knopp, *Über Reihen mit positiven Gliedern*, J. London Math. Soc. **3** (1928), 205–211.

[24] A. Kolesárová, *Limit properties of quasi-arithmetic means*, Fuzzy Sets and Systems **124** (2001), 65–71.

[25] A. Kolmogoroff, *Sur la notion de la moyenne*, Rend. Accad. dei Lincei **6** (1930), 388–391.

[26] A. Kufner, L. Maligranda, L.-E. Persson, *The Hardy Inequality: About its History and Some Related Results*, Vydavatelský Servis, Pilsen 2007.

[27] E. Landau, *A note on a theorem concerning series of positive terms*, J. London Math. Soc. **1** (1921), 38–39.

[28] E. R. Love, *Inequalities related to Carleman's inequality*, in: W. Norrie Everitt (ed.), *Inequalities: Fifty years on from Hardy, Littlewood and Pòlya*, 135–141, Lecture notes in pure and applied mathematics vol. 129, Marcel Dekker, Inc., New York, 1991.

[29] J. G. Mikusiński, *Sur les moyennes de la forme* $\psi^{-1}\left[\sum q\psi(x)\right]$, Studia Math. **10** (1948), 90–96.

[30] H. P. Mulholland, *On the generalization of Hardy's inequality*, J. Lond. Math. Soc. **7** (1932), 208–214.

[31] M. Nagumo, *Uber eine Klasse der Mittelwerte*, Jap. Journ. of Math. **7** (1930), 71–79.

[32] M. Nosarzewska *On uniform convergence in some classes of functions*, Fundamenta Mathematicae **39** (1952), 38–52.

[33] J. A. Oguntuase, L-E. Persson, *Hardy type inequalities via convexity – the journey so far*, Aust. J. Math. Anal. Appl. **7** (2011), 1–19.

[34] Zs. Páles *Report of Meeting 15th International Conference on Functional Equations and Inequalities*, Ann. Univ. Paedag. Crac. Stud. Math. **XII** (2013), 121–122.

[35] Zs. Páles, *On the convergence of Means*, J. Anal. Appl. **156** (1991), 52–60.

[36] Zs. Páles, L.-E. Persson, *Hardy-type inequalities for means*, Bull. Austral. Math. Soc. **70** (2004), 521–528.

[P1] P. Pasteczka, *When is a Family of Generalized Means a Scale?*, Real Analysis Exchange **38** (2013), 193–210.

[P2] P. Pasteczka, *A new estimate of the difference among quasi-arithmetic means*, arXiv:1310.7212.

[P3] P. Pasteczka, *On some type of Hardy inequality involving generalized power means*, Publ. Math. Debrecen, to appear.

[P4] P. Pasteczka, *On negative results concerning Hardy means*, Acta Math. Hungar., to appear.

[P5] P. Pasteczka, *Limit properties in a family of quasi-arithmetic means*, arXiv:1504.02371.

[42] J. Pečarič, K. B. Stolarsky, *Carleman's inequality: history and new generalizations*, Aequationes Math. **61** (2001), 49–62.

[43] E. Pizzetti, *Osservazione sulle medie esponenziali e baso-esponenziali*, Metron **13**, No. 4 (1939), 3–15

[44] R. M. Redheffer, *Easy proofs of hard inequalities*, in: E. F. Beckenbach, W. Walter (eds.) *General Inequalities 3*, 123—140, Birkhäuser Verlag, Basel, 1983.

[45] U. Ricci, *Confronti fra medie*, Giorn. Econ. Rev. Statist. **3** (36), No. 11 (1915), 38–66.

[46] R. Kh. Sadikova, *Comparison of discrete mixed means containing symmetric functions*, Math. Notes, **80** (2006), 254–260 (English translation).