Wokół średnich

Justyna Jarczyk

Instytut Matematyki
Uniwersytet Zielonogórski

Wirtualne Seminarium PolWoMaths

1. Wprowadzenie - pojęcie średniej i klasyczne przykłady
2. Średnie w badaniach dawnych matematyków - od Babilończyków i Archimedesa do Gaussa i Borchardta
3. Uogólnione algorytmy Gaussa i Archimedesa-Borchardta
4. Niezmienniczość średnich
5. Ważone średnie quasi-arytmetyczne

Idea średniej jest równie stara jak pojęcie liczby. Mając dane wielkości $x_1, \ldots, x_p$ intuicyjnie poszukujemy ich średniej $M(x_1, \ldots, x_p)$ gdzieś pomiędzy ich skrajnymi wartościami:

$$\min\{x_1, \ldots, x_p\} \leqslant M(x_1, \ldots, x_p) \leqslant \max\{x_1, \ldots, x_p\}. \tag{1}$$

Warunek ten to, prawdopodobnie, pierwsze formalna definicja średniej. Została podana przez Cauchy'ego w traktacie [6] w roku 1821. Wśród klasycznych przykładów, doskonale znanych już w starożytności, pojawiają się

średnia arytmetyczna $A : \mathbb{R}^p \to \mathbb{R}$:

$$A(x_1, \ldots, x_p) = \frac{x_1 + \ldots + x_p}{p},$$

średnia geometryczna $G : (0, +\infty)^p \to \mathbb{R}$:

$$G(x_1, \ldots, x_p) = \sqrt[p]{x_1 \cdot \ldots \cdot x_p},$$

średnia harmoniczna $H : (0, +\infty)^p \to \mathbb{R}$:

$$H(x_1, \ldots, x_p) = \frac{p}{\frac{1}{x_1} + \ldots + \frac{1}{x_p}}.$$

Jak wiemy, podstawowy związek pomiędzy tymi średnimi opisują poniższe nierówności

$$H(x_1, \ldots, x_p) \leqslant G(x_1, \ldots, x_p) \leqslant A(x_1, \ldots, x_p)$$

dla wszystkich $x_1, \ldots, x_p \in (0, +\infty)^p$.

Dokładniej, mając ustalony przedział I, każdą funkcję $M : I^p \to I$, spełniającą nierówności (1) dla wszystkich $x_1, \ldots, x_p \in I$, nazywamy *średnią* (p *zmiennych*) *na przedziale* I. Średnią $M : I^p \to I$ nazywamy *ścisłą*, gdy nierówności (1) są silne o ile $\min\{x_1, \ldots, x_p\} < \max\{x_1, \ldots, x_p\}$ *symetryczną*, gdy

$$M\left(x_{\sigma(1)}, \ldots, x_{\sigma(p)}\right) = M\left(x_1, \ldots, x_p\right)$$

dla wszystkich $x_1, \ldots, x_p \in I$ i dla każdej permutacji $\sigma : \{1, \ldots, p\} \to \{1, \ldots, p\}$, *jednorodną*, gdy $I = \mathbb{R}$ lub $I = (0, +\infty)$ oraz

$$M\left(tx_1, \ldots, tx_p\right) = tM\left(x_1, \ldots, x_p\right)$$

dla wszystkich $x_1, \ldots, x_p \in I$ oraz $t \in I$. W ogólności *średnią na przedziale* I nazywamy dowolną funkcję $M : \bigcup_{p=1}^{\infty} I^p \to I$ spełniającą nierówności (1) dla wszystkich $x_1, \ldots, x_p \in I$ i $p \in \mathbb{N}$. Wiele ważnych rodzin średnich, a także związki między nimi, są omawiane w książce Bullena [5]. Bogactwo informacji na temat średnich można też znaleźć w klasycznych książkach Hardy'ego, Littlewooda i Pólyi [14] i braci Borwain [4], a także w pracy przeglądowej Daróczyego i Pálesa [10].

Iteracyjne procedury, wśród nich te pozwalające przybliżać różne liczby niewymierne za pomocą średnich, są znane od starożytności. Jeden z klasycznych algorytmów iteracyjnych ma postać

$$x_{k+1} = \frac{1}{2}\left(x_k + \frac{a}{x_k}\right), \quad k \in \mathbb{N}_0. \tag{2}$$

Dziś wiążemy go z nazwiskiem Newtona, ale przecież jest to formalizacja babilońskiej metody wyznaczania pierwiastka kwadratowego liczby dodatniej a. Startując z dowolnej liczby dodatniej x_0 ciąg $(x_k)_{k\in\mathbb{N}_0}$ jest (ściśle) malejący i ograniczony, a więc zbieżny - przybliża liczbę $\sqrt{a}$. Przyjmując $y_k := a/x_k$ widzimy, że warunek (2) jest równoważny warunkowi

$$x_{k+1} = \frac{1}{2}(x_k + y_k) \quad \text{i} \quad \frac{1}{y_{k+1}} = \frac{1}{2}\left(\frac{1}{x_k} + \frac{1}{y_k}\right), \quad k \in \mathbb{N}_0,$$

czyli

$$x_{k+1} = A(x_k, y_k) \quad \text{i} \quad y_{k+1} = H(x_k, y_k), \quad k \in \mathbb{N}_0, \tag{3}$$

gdzie A i H oznaczają odpowiednio średnią arytmetyczną i harmoniczną. Z definicji ciągu $(y_k)_{k\in\mathbb{N}_0}$ wynika, że ściśle rośnie do $\sqrt{a}$.

Zatem

$$y_k < y_{k+1} < x_{k+1} < x_k, \quad k \in \mathbb{N}_0,$$

oraz

$$\lim_{k\to\infty} x_k = \lim_{k\to\infty} y_k = \sqrt{a}.$$

Oczywiście algorytm (3) jest szczególnym przypadkiem rekurencji

(4) $$x_{k+1} = M(x_k, y_k) \quad \text{i} \quad y_{k+1} = N(x_k, y_k), \quad k \in \mathbb{N}_0,$$

gdzie $M, N : I^2 \to I$ są średnimi na przedziale I.

Innym ważnym przykładem rekurencyjnego algorytmu (4) jest ten, gdzie $M = A$ i $N = G$, czyli M i N są, odpowiednio, średnimi arytmetyczną i geometryczną:

(5) $$x_{k+1} = \frac{1}{2}(x_k + y_k) \quad \text{i} \quad y_{k+1} = (x_k y_k)^{1/2}, \quad k \in \mathbb{N}_0.$$

Oba powyższe ciągi maja wspólną granicę zwana średnią arytmetyczno-geometryczną (*medium arithmeticum-geometricum*) liczb x_0 i y_0, oznaczaną $A \otimes G(x_0, y_0)$. Przepiszmy równości (5) w postaci

$$(x_{k+1}, y_{k+1}) = (A, G)(x_k, y_k), \quad k \in \mathbb{N}_0.$$

Wtedy

$$(x_k, y_k) = (A, G)^k (x_0, y_0), \quad k \in \mathbb{N}_0,$$

czyli

$$(A, G)^k \longrightarrow (A \otimes G, A \otimes G). \tag{6}$$

Algorytm (5) po raz pierwszy pojawił się w roku 1784 w pracy [23] Lagrange'a w związku z szacowaniem całek eliptycznych. Jednak, jak się wydaje, to Gauss odkrył, że ten algorytm dostarcza iteracyjne rozwiązanie problemu prostowalności łuku lemniskaty Bernoulliego. W szczególności pomogło to podać krótki dowód twierdzenia Fagnano, pokazującego jak podwoić łuk lemniskaty, za pomocą cyrkla i linijki. W roku 1799 Gauss odnotował, że

$$A \otimes G\left(1, \sqrt{2}\right) = \frac{\pi}{2}\left(\int_0^{\pi/2} \frac{1}{\left(1 + \sin^2\vartheta\right)^{1/2}} d\vartheta\right)^{-1}.$$

Wartość średniej arytmetyczno-geometrycznej w dowolnym punkcie $(x_0, y_0) \in (0, +\infty)^2$ została wyznaczona przez Gaussa w roku 1818:

$$A \otimes G\,(x_0, y_0) = \frac{\pi}{2}\left(\int_0^{\pi/2} \frac{1}{\left(x_0^2 \cos^2\vartheta + y_0^2 \sin^2\vartheta\right)^{1/2}}\,d\vartheta\right)^{-1}.$$

Szczegółowy opis powyższej teorii Gaussa podał Cox w pracy [8] z roku 1984. O innych średnich złożonych, otrzymanych jako wspólne granice ciągów zdefiniowanych w warunku (4), można się dowiedzieć z książki braci Borwein [4].

W roku 1800 Gauss zaproponował badanie rekurencji

$$(7) \qquad x_{k+1} = \frac{1}{2}(x_k + y_k) \quad \text{i} \quad y_{k+1} = (x_{k+1}y_k)^{1/2}, \quad k \in \mathbb{N}_0,$$

która jest pozornie podobna do algorytmu (5). Jak się wydaje zdawał sobie sprawę, z tego że, ciągi te mają wspólną granicę. Okazuje się jednak, że nie wyraża się ona poprzez funkcje eliptyczne tak jak w przypadku średniej arytmetyczno-geometrycznej, ale przy użyciu funkcji trygonometrycznych i hiperbolicznych. W roku 1880 algorytm (7) i jego podstawowe własności zostały na nowo odkryte przez Borchardta w pracy [3]. Od tego czasu nosi on nazwę algorytmu Borchardta.

Zamieniając w algorytmie Borchardta (7) średnią arytmetyczną średnią harmoniczną dochodzimy do rekurencji

$$(8) \qquad x_{k+1} = \frac{2}{\frac{1}{x_k} + \frac{1}{y_k}} \quad \text{i} \quad y_{k+1} = (x_{k+1} y_k)^{1/2}, \quad k \in \mathbb{N}_0.$$

Także tutaj ciągi $(x_k)_{k\in\mathbb{N}}$ i $(y_k)_{k\in\mathbb{N}}$ zmierzają do wspólnej granicy. W szczególności, startując z punktu $(2\sqrt{3}, 3)$ otrzymujemy algorytm przypisywany Archimedesowi, aproksymujący liczbę π. Zauważmy, że ciągi $(x_k)_{k\in\mathbb{N}}$ i $(y_k)_{k\in\mathbb{N}}$ spełniają algorytm Borchardta wtedy i tylko wtedy, gdy ciągi $(1/x_k)_{k\in\mathbb{N}}$ i $(1/y_k)_{k\in\mathbb{N}}$ spełniają rekurencję Archimedesa (8). Zatem wystarczy rozpatrywać tylko jeden z tych algorytmów, na przykład (7). Oba procesy: Borchardta (7) i Archimedesa (8) są szczególnymi przypadkami algorytmu

$$(9) \qquad x_{k+1} = M(x_k, y_k) \quad \text{i} \quad y_{k+1} = N(x_{k+1}, y_k), \quad k \in \mathbb{N}_0,$$

gdzie M i N są średnimi dwóch zmiennych na przedziale I i $x_0, y_0 \in I$. Więcej przykładów rekurencji Archimedesa-Borchardta (9) zawiera praca [13] Fostera i Phillipsa z roku 1984.

Jeśli średnie $M, N : I^2 \to I$ są średnimi dwóch zmiennych na przedziale I, to para (M, N) odwzorowuje kwadrat I^2 w siebie, a więc możemy ją iterować. Zauważmy, że uogólniony algorytm Gaussa (4) można teraz zapisać w postaci

$$(x_k, y_k) = (M, N)^k (x_0, y_0), \quad k \in \mathbb{N}_0,$$

gdzie $(M, N)^k$ oznacza k-tą iteratę odwzorowania (M, N). Tym samym badanie granicznego zachowania ciągów $(x_k)_{k\in\mathbb{N}_0}$ i $(y_k)_{k\in\mathbb{N}_0}$ to po prostu badanie zbieżności ciągu iterat $\left((M, N)^k\right)_{k\in\mathbb{N}_0}$.

Sformułujemy teraz twierdzenie opisujące to graniczne zachowanie w ogólnym przypadku, gdy p jest ustaloną liczbą naturalną, a $M_1, \ldots, M_p : I^p \to I$ średnimi p-zmiennych.

Twierdzenie 1

Niech I będzie przedziałem, a $M_1, \ldots, M_p : I^p \to I$ średnimi ciągłymi. Załóżmy, że dla każdego $\mathbf{x} = (x_1, \ldots, x_p) \in I^p$ jeśli

$$\min\left\{M_1(\mathbf{x}), \ldots, M_p(\mathbf{x})\right\} = \min\left\{x_1, \ldots, x_p\right\} \tag{10}$$

oraz

$$\max\left\{M_1(\mathbf{x}), \ldots, M_p(\mathbf{x})\right\} = \max\left\{x_1, \ldots, x_p\right\}, \tag{11}$$

to $x_1 = \ldots = x_p$. Wtedy istnieje taka średnia ciągła $L : I^p \to I$, że

$$\lim_{k \to \infty} (M_1, \ldots, M_p)^k = (L, \ldots, L)$$

jednostajnie na każdym zwartym podzbiorze kostki I^p; ponadto L jest jedyną ciągła średnią spełniającą warunek

$$L \circ (M_1, \ldots, M_p) = L.$$

Powyższe twierdzenie pochodzi z pracy [27] Matkowskiego, opublikowanej w roku 2009. We wcześniejszych jego pracach pojawiło się silniejsze, ale za to nieco czytelniejsze założenie, że co najwyżej jedna ze średnich $M_1, \ldots, M_p$ nie jest ścisła. Nie jest trudno zauważyć, że wtedy spełnione jest założenie Twierdzenia 2: jeśli $\mathbf{x} = (x_1, \ldots, x_p) \in I^p$ i zachodzą równości (10) i (11) to $x_1 = \ldots = x_p$. Oczywiście w ciągu 200 lat, jakie dzielą badania Gaussa i Matkowskiego, problem uogólnienia oryginalnego wyniku Gaussa powracał wielokrotnie. Dotyczy to głównie minionego półwiecza, choć nie tylko. Wśród prac, które koniecznie należy tu wymienić, są artykuły Sutô [30] i [31] z roku 1914, Lehmera [24] z roku 1971, Schoenberga [29], który ukazał się w roku 1982, wspomniana już wcześniej praca [13] Fostera i Philipsa z roku 1984, w końcu książka [4] braci Borwein opublikowana 3 lata później.

W ciągu ostatnich kilku lat Twierdzenie 1 doczekało się różnych uogólnień. Jedno z nich, rozszerzające algorytm Gaussa na średnie z parametrem, zostało opublikowane w mojej pracy [16] przed sześcioma laty. Mając dany przedział I, liczbę naturalną p i niepusty zbiór Ω, przez średnią z parametrem (p-zmiennych na przedziale I) będziemy rozumieć każdą taką funkcję $M : I^p \times \Omega \to I$, że dla każdego $\omega \in \Omega$ funkcja $M(\cdot, \omega)$ jest średnią. W głównym twierdzeniu pracy [16] używam iterat $(M_1, \ldots, M_p)^n$ odwzorowania $(M_1, \ldots, M_p) : I^p \times \Omega \to I^p$. Stosuję tu następującą definicję pochodzącą z pracy [2] Barona i Kuczmy z roku 1977.

Przyjmijmy $\Omega^\infty := \Omega^{\mathbb{N}}$. Mając dany zbiór X i funkcję $f : X \times \Omega \to X$, kolejne iteraty $f^n : X \times \Omega^\infty \to X$, $n \in \mathbb{N}$, określamy przy pomocy równości

$$f^1(x, \omega_1, \omega_2, \ldots) := f(x, \omega_1)$$

i rekurencji

$$f^{n+1}(x, \omega_1, \omega_2, \ldots) := f(f^n(x, \omega_1, \omega_2, \ldots), \omega_{n+1})$$

postulowanych dla wszystkich $x \in X$, $(\omega_1, \omega_2, \ldots) \in \Omega^\infty$ i $n \in \mathbb{N}$. Zauważmy, że w rzeczywistości n-ta iterata $f^n(\cdot, \omega_1, \omega_2, \ldots)$ zależy jedynie od pierwszych n parametrów $\omega_1, \ldots, \omega_n$.

W głównym wyniku pracy [16], czyli Theorem 3.5, pod założeniem zwartości zbioru parametrów Ω uzyskałam tezę analogiczną do tej z Twierdzenia 1. Prosty przykład podany w tej samej pracy pokazuje że założenie zwartości zbioru Ω jest tu istotne. W przypadku, gdy na zbiorze Ω zadana jest tu istotnie. W przypadku, gdy na zbiorze Ω zadana jest miara probabilistyczna, można rozpatrywać pewną istotną modyfikację omawianych tu wyników. Średnie z parametrem nazywamy wtedy średnimi losowymi. Praca [18], napisana wspólnie z Witoldem Jarczykiem, kładzie nacisk na probabilistyczną stronę algorytmu Gaussa dla średnich losowych.

Przejdźmy teraz do uogólnionego algorytmu Archimedesa-Borchardta dla średnich p zmiennych. Mając dane średnie $M_1, \ldots, M_p : I^p \to I$ i punkty $x_{0,1}, \ldots x_{0,p} \in I$ rozpatrzmy proces rekurencyjny

$$(12)\quad \begin{array}{l} x_{k+1,1} = M_1\left(x_{k,1}, \ldots, x_{k,p}\right), \;\; k \in \mathbb{N}_0, \\ x_{k+1,i} = M_i\left(x_{k+1,1}, \ldots, x_{k+1,i-1}, x_{k,i}, \ldots, x_{k,p}\right), \;\; i = 2, \ldots, p, \;\; k \in \mathbb{N}_0, \end{array}$$

który uogólnia zarówno algorytmy (7), jak i (9). W pracy [19], napisanej wspólnie z Witoldem Jarczykiem, udowodniłam następujący rezultat

Twierdzenie 2

Jeśli $x_{0,1}, \ldots, x_{0,p}$ są punktami przedziału I, a $M_1, \ldots, M_p : I^p \to I$ ścisłymi średnimi ciągłymi, to ciągi $\left(x_{k,1}\right)_{k \in N}, \ldots, \left(x_{k,p}\right)_{k \in N}$, określone równościami (12), zmierzają do wspólnej granicy. Zbieżność ta jest jednostajna ze względu na punkt początkowy $\mathbf{x}_0 = \left(x_{0,1}, \ldots, x_{0,p}\right)$ w dowolnym podzbiorze zwartym kostki I^p.

Dowód Twierdzenia 2 biegnie poprzez zredukowanie rekurencji Archimedesa-Borchardta (12) do odpowiedniego uogólnionego algorytmu Gaussa. Kluczową rolę odgrywa tu następujący lemat.

Mając dany przedział I i dowolne funkcje $M_1, \dots, M_p : I^p \longrightarrow I$, rekurencyjnie określmy funkcje $N_1, \dots, N_p : I^p \to I$ przy pomocy równości

$$(13)\quad \begin{aligned} &N_1(\mathbf{x}) = M_1(\mathbf{x}),\\ &N_i(\mathbf{x}) = M_i\left(N_1(\mathbf{x}), \dots, N_{i-1}(\mathbf{x}), x_i, \dots, x_p\right), \qquad i = 2, \dots, p, \end{aligned}$$

spełnionych dla wszystkich $\mathbf{x} = (x_1, \dots, x_p) \in I^p$.

Lemat 3

Niech $N_1, \dots, N_p : I^p \longrightarrow I$ będą określone równościami (13). Wtedy:

(i) jeśli $M_1, \dots, M_p$ są średnimi, to także $N_1, \dots, N_p$ są średnimi;

(ii) jeśli średnie $M_1, \dots, M_p$ są ścisłe, to także średnie $N_1, \dots, N_p$ są ścisłe;

(iii) jeśli funkcje $M_1, \dots, M_p$ są ciągłe, to także $N_1, \dots, N_p$ są ciągłe;

(iv) dla każdego punktu $\mathbf{x}_0 = (x_{0,1}, \dots, x_{0,p}) \in I^p$ ciągi $(x_{k,i})_{k \in \mathbb{N}_0}$, $i = 1, \dots, p$, spełniają rekurencję (13) wtedy i tylko wtedy, gdy

$$(14)\quad x_{k+1,i} = N_i\left(x_{k,1}, \dots, x_{k,p}\right), \quad i = 1, \dots, p,\ k \in \mathbb{N}_0.$$

Prosta obserwacja opisana w Lemacie 3(iv) pokazuje, że badanie procesu Archimedesa-Borchardta (12) można zredukować do znanego nam już algorytmu Gaussa (14). Przypomnijmy, że równości (14) są równoważne warunkowi

$$\mathbf{x}_k = (N_1, \ldots, N_p)^k (\mathbf{x}_0), \qquad k \in \mathbb{N}_0.$$

Tak więc ciąg $(\mathbf{x}_k)_{k\in\mathbb{N}_0}$, spełniający rekurencję (14), jest ciągiem iteracyjnym. Zauważmy jednak, że zupełnie nie jest to widoczne jeśli patrzymy na ten ciąg jako rozwiązanie równań (12).

Dowód Twierdzenia 2. Weźmy dowolny punkt $\mathbf{x}_0 = (x_{0,1}, \ldots, x_{0,p}) \in I^p$ i niech $(x_k)_{k\in\mathbb{N}_0}$, gdzie $\mathbf{x}_k = (x_{k,1}, \ldots, x_{k,p})$, będzie ciągiem określonym równościami (12), w których $M_1, \ldots, M_p : I^p \to I$ są ścisłymi średnimi ciągłymi. Określmy funkcje $N_1, \ldots, N_p : I^p \to I$ równościami (13). Wobec Lematu 3 są one ścisłymi średnimi ciągłymi, a ciąg $(\mathbf{x}_k)_{k\in\mathbb{N}_0}$ spełnia równości (14). Na mocy Twierdzenia 1, zastosowanego do średnich $N_1, \ldots, N_p$, istnieje taka średnia L na przedziale I, że

$$\mathbf{x}_k = (N_1, \ldots, N_p)^k (\mathbf{x}_0) \longrightarrow (L, \ldots, L)(\mathbf{x}_0)$$

jednostajnie na każdym zwartym podzbiorze kostki I^p ze względu na $\mathbf{x}_0$. W konsekwencji dla każdego $i = 1, \ldots, p$ ciąg $(x_{k,i})_{k\in\mathbb{N}_0}$ jest zbieżny do tej samej granicy $L(\mathbf{x}_0)$.

$\square$

Zacznijmy od następującej prostej obserwacji. Dla dowolnych liczb $x, y \in (0, +\infty)$ mamy

$$\frac{x+y}{2} \cdot \frac{2}{\frac{1}{x}+\frac{1}{y}} = (x+y)\frac{xy}{y+x} = xy,$$

skąd wynika, że

$$G\left(A(x,y), H(x,y)\right) = G(x,y).$$

Innymi słowy

$$G \circ (A, H) = G.$$

Powyższa równość wyraża powszechnie znaną własność średniej geometrycznej: jest niezmiennicza ze względu na parę (A, H) zbudowaną ze średnich arytmetycznej i geometrycznej.

Przyjmijmy teraz w Twierdzeniu 1, że $I = (0, +\infty)$, $p = 2$, $M_1 = A$ i $M_2 = G$. Wtedy, ponieważ średnie A i G są ściśle, spełnione są założenia tego twierdzenia. Korzystając z jego tezy widzimy, że istnieje taka ciągła średnia $L : I^2 \to I$, że

$$\lim_{k\to\infty} (A, G)^k = (L, L), \tag{15}$$

przy czym L jest jedyną ciągłą średnią spełniającą równość

$$L \circ (A, G) = L.$$

Z równości (6) i (15) wynika, że $L = A \otimes G$. Zatem

$$(A \otimes G) \circ (A, G) = A \otimes G,$$

a więc średnia arytmetyczno-geometryczna jest jedyną ciągłą średnią niezmienniczą ze względu na parę (A, G).

W ogólności, jeśli $M_1, \ldots, M_p : I^p \to I$ i $L : I^p \to I$ są średnimi spełniającymi równość

$$L \circ (M_1, \ldots, M_p) = L,$$

to mówimy, że średnia L jest niezmiennicza ze względu na odwzorowanie $(M_1, \ldots, M_p)$. Badanie niezmienniczości w różnych klasach średnich jest od ponad 20 lat szeroko prowadzone przez wielu matematyków. Ale pierwsze prace na ten temat zostały opublikowane ponad 100 lat temu w japońskim czasopiśmie Tôhoku Mathematical Journal.

W dalszym ciągu przydatne okaże się następujące pojęcie. Mając dany zbiór X powiemy, że funkcje $\varphi \colon X \to \mathbb{R}$ i $\psi \colon X \to \mathbb{R}$ są *równoważne* lub, że φ jest *równoważna* ψ, gdy istnieją takie liczby $a \in \mathbb{R} \setminus \{0\}$ i $b \in \mathbb{R}$, że

$$\psi(x) = a\varphi(x) + b, \quad x \in X;$$

piszemy wtedy

$$\varphi(x) \sim \psi(x), \qquad x \in X,$$

lub po prostu $\varphi \sim \psi$. Oczywiście $\sim$ jest relacją równoważności w zbiorze $\mathbb{R}^X$ funkcji rzeczywistych określonych na zbiorze X.

Twierdzenie 4

Niech $\varphi : I \to \mathbb{R}$ i $\psi : I \to \mathbb{R}$ będą ściśle monotonicznymi funkcjami analitycznymi. Para (φ, ψ) spełnia równanie

$$(16) \qquad \varphi^{-1}\left(\frac{\varphi(x)+\varphi(y)}{2}\right)+\psi^{-1}\left(\frac{\psi(x)+\psi(y)}{2}\right)=x+y$$

wtedy i tylko wtedy, gdy albo

$$\varphi(x) \sim x,\ x \in I, \quad \text{i} \quad \psi(x) \sim x,\ x \in I,$$

albo

$$\varphi(x) \sim e^{ax},\ x \in I, \quad \text{i} \quad \psi(x) \sim e^{-ax},\ x \in I,$$

z pewnym $a \in \mathbb{R} \setminus \{0\}$.

W sformułowaniu tym nie pojawiają się jawnie żadne średnie. Nie jest trudno sprawdzić, że jeśli $\varphi : I \to \mathbb{R}$ jest ściśle monotoniczną funkcją ciągłą, to wzór

$$A^{\varphi}(x,y) = \varphi^{-1}\left(\frac{\varphi(x)+\varphi(y)}{2}\right)$$

definiuje średnią A^{φ} na przedziale I. Zatem, stosując powyższy wzór i dzieląc równość (16) przez 2, Twierdzenie 4 możemy sformułować równoważnie w sposób następujący.

Twierdzenie 5

Niech $\varphi, \psi : I \to \mathbb{R}$ będą ściśle monotonicznymi funkcjami analitycznymi. Średnia arytmetyczna A jest niezmiennicza ze względu na parę $\left(A^{\varphi}, A^{\psi}\right)$:

$$A \circ \left(A^{\varphi}, A^{\psi}\right) = A, \tag{17}$$

wtedy i tylko wtedy, gdy albo $A^{\varphi} = A^{\psi} = A$, albo

$$A^{\varphi}(x,y) = \frac{1}{a}\log\left(\frac{e^{ax}+e^{ay}}{2}\right) \text{ i } A^{\psi}(x,y) = -\frac{1}{a}\log\left(\frac{e^{-ax}+e^{-ay}}{2}\right)$$

dla wszystkich $x, y \in I$, z pewnym $a \in \mathbb{R}\setminus\{0\}$.

Średnia A^{φ}, występująca powyżej, nazywa się średnią quasi-arytmetyczną o generatorze φ. Twierdzenie 5 charakteryzuje więc wszystkie pary średnich o generatorach analitycznych, względem których średnia arytmetyczna jest niezmiennicza.

Problem, rozwiązany przez Sutô w Twiedzeniu 4 w klasie funkcji analitycznych, został ponownie odkryty przez Matkowskigo. W pracy [26], nie znając jeszcze wtedy prac Sutô, udowodnił Twierdzenie 5 pod następującym słabszym założeniem: $\varphi : I \to \mathbb{R}$ i $\psi : I \to \mathbb{R}$ są ściśle monotonicznymi funkcjami ciągłymi, z których choć jedna jest dwukrotnie różniczkowalna w sposób ciągły. Kolejny krok w osłabianiu założeń regularnościowych w Twierdzeniu 5 zrobili w 2001 roku matematycy węgierscy Daróczy i Páles. W artykule [9] pokazali, że w powyższym sformułowaniu można opuścić słowo *dwukrotnie*. Badania trwały jednak dalej. Zauważmy bowiem, że żadne regularnościowe założenia o generatorach φ, ψ nie są naturalne, ponieważ sformułowanie problemu niezmienniczości (17) nie używa w ogóle pojęcia pochodnej.

Ostatecznie jednak w roku 2002 w pracy [10] Daróczy i Páles udowodnili tezę Twierdzenia 5 jedynie pod założeniem, że generatory φ i ψ są ściśle monotoniczne i ciągłe.

Opisana tu pokrótce historia badania niezmienniczości średniej arytmetycznej względem pary średnich quasi-arytmetycznych jest tylko jedną z wielu, które można by przedstawić. W ciągu ostatnich 20 lat opublikowano około 200 prac poświęconych problemowi niezmienniczości w bardzo wielu klasach średnich. Tematyce tej poświęcona jest praca przeglądowa [17] zatytułowana *Invariance of means*, którą napisałam wspólnie z Witoldem Jarczykiem.

Niewątpliwie jedną z klas średnich, w której problem niezmienniczości został najpełniej zbadany, jest klasa ważonych średnich quasi-arytmetycznych. Przejdziemy teraz do kolejnego, ostatniego już, paragrafu całkowicie poświęconego tej klasie średnich. Omówimy w nim nie tylko problem niezmienniczości, ale przedyskutujemy także własność wypukłości tych średnich.

Skupmy się na razie na średnich dwóch liczb. Mając dany przedział I, ściśle rosnącą funkcję ciągłą $\varphi : I \to \mathbb{R}$ i liczbę $p \in (0, 1)$, symbolem A_p^φ oznaczamy średnią na przediale I, określoną wzorem

$$A^\varphi(x, y)_p = \varphi^{-1}\left(p\varphi(x) + (1 - p)\varphi(y)\right).$$

Funkcję φ nazywamy generatorem średniej A_p^φ, a liczbę p jej wagą. Zauważmy, że $A_{1/2}^\varphi$ to po prostu średnia quasi-arytmetyczna. Tak więc rodzina ważonych średnich quasi-arytmetycznych zawiera klasę średnich quasi-arytmetycznych. Wśród najważniejszych przykładów średnich quasi arytmetycznych znajdziemy średnią arytmetyczną (generowaną przez funkcję identycznościową), średnia geometryczną (generowaną przez funkcję logarytmiczna), średnią harmoniczną (generowaną przez funkcję φ określoną wzorem $\varphi(x) = 1/x$). Pewne idee związane z pojęciem średniej quasi-arytmetycznej pojawiły się w roku 1928 w pracy Knoppa [21]. Ale formalna definicja została wprowadzona niezależnie przez trzech autorów: w roku 1930 przez Kołmogorowa i Nagumo, odpowiednio w pracach [22] i [28], a rok później przez de Finettiego w pracy [12]. Dziś średnie quasi-arytmetyczne bywają nazywane średnimi Kołmogorowa.

Wracając do ogólniejszej klasy ważonych średnich quasi-arytmetycznych odnotujemy następujące twierdzenie podające rozwiązanie tak zwanego *equality problem*. W poniższej ostatecznej wersji zostało udowodnione przez Maksę i Pálesa w pracy [25] w roku 2010, choć jako swoiste "folk theorem" było znane Hardy'emu, Littlewoodowi i Pólyi (zob. ich słynną książkę [14] pod tytułem *Inequalities* z 1934) oraz Aczélowi (zob. monografię [1] z roku 1966).

Theorem 6

Niech $\varphi, \psi : I \to \mathbb{R}$ będą ściśle monotonicznymi funkcjami ciągłymi i niech $p, q \in (0,1)$. Wówczas $A_p^\varphi = A_q^\psi$ wtedy i tylko wtedy, gdy $\varphi \sim \psi$ i $p = q$.

Powyższy wynik pozwala uprościć dowody szeregu lematów, które trzeba wykazać, żeby otrzymać rozwiązanie problemu niezmienniczości w klasie ważonych średnich quasi-arytmetycznych na ustalonym przedziale I. Polega on na wyznaczeniu wszystkich ściśle monotonicznych funkcji ciągłych $\varphi, \psi, \chi : I \to \mathbb{R}$ i takich liczb $p, q, r \in (0,1)$, że

$$A_r^\chi \circ \left(A_p^\varphi, A_q^\psi\right) = A_r^\chi. \tag{18}$$

Poniższe twierdzenie pochodzi z mojej pracy [15] z roku 2007.

Twierdzenie 7

Niech I będzie przedziałem, $\varphi, \psi, \chi : I \to \mathbb{R}$ ściśle monotonicznymi funkcjami ciągłymi i niech $p, q, r \in (0,1)$. Średnia A_r^χ jest niezmiennicza ze względu na parę $\left(A_p^\varphi, A_q^\psi\right)$, czyli trójki (φ, ψ, χ) i (p, q, r) spełniają równanie (18) wtedy i tylko wtedy, gdy

(i) $r = \frac{q}{1-p+q}$

oraz

(ii) albo $A_p^\varphi = A_p^\chi$ i $A_q^\psi = A_q^\chi$, albo $r = 1/2$, $p + q = 1$,

$$A_p^\varphi(x,y) = \chi^{-1}\left(\frac{1}{a}\log\left(pe^{a\chi(x)} + (1-p)e^{a\chi(y)}\right)\right),$$

$$A_q^\psi(x,y) = \chi^{-1}\left(-\frac{1}{a}\log\left(qe^{-a\chi(x)} + (1-q)e^{-a\chi(y)}\right)\right)$$

dla wszystkich $x, y \in I$, z pewnym $a \in \mathbb{R} \setminus \{0\}$.

Rezultat ten jest ukoronowaniem kilkuletnich prób rozwiązania problemu niezmienniczości ważonych średnich quasi-arytmetycznych, podejmowanych przez matematyków węgierskich i polskich. Pierwszym krokiem była tu praca [11] Daróczyego i Pálesa z roku 2003, w której autorzy, rozwiązując pewien szczególny przypadek równania (18), dodatkowo założyli, że generatory φ i ψ mają ciągłe niezerujące się pochodne. Jednym z kolejnych ważnych wyników było rozwiązanie równania (18), pod założeniem, że generatory φ, ψ i χ są dwukrotnie różniczkowalne w sposób ciągły. Rezultat ten pochodzi z pracy [20] opublikowanej w roku 2006, którą napisałam wspólnie z Matkowskim. Kilka innych prac, które przybliżyły ostateczne rozwiązanie problemu zostało napisanych przez Buraia, Daróczyego i Pálesa.

Na koniec omówię krótko najważniejsze wyniki pracy [7], która powstała w roku 2019 we współpracy z Chudziakiem, Głazowską i Witoldem Jarczykiem. Rozpocznijmy od rozszerzenia definicji ważonej średniej quasi-arytmetycznej na przypadek wielu zmiennych. Mając dany przedział I, ściśle rosnącą funkcję ciągłą $\varphi : I \to \mathbb{R}$, liczbę naturalną $n \geqslant 2$ i taki ciąg $\mathbf{p} = (p_1, \ldots, p_n) \in (0,1)^n$, że $p_1 + \ldots + p_n = 1$, symbolem $A_{\mathbf{p}}^{\varphi}$ oznaczamy średnią na przedziale I, określoną wzorem

$$A_{\mathbf{p}}^{\varphi}(\mathbf{x}) = \varphi^{-1}\left(p_1\varphi\left(x_1\right) + \ldots + p_n\varphi\left(x_n\right)\right)$$

dla każdego $\mathbf{x} = (x_1, \ldots, x_n) \in I^n$. Mówimy, że $A_{\mathbf{p}}^{\varphi}$ jest ważoną średnią quasi-arytmetyczną o generatorze φ i wagach $p_1, \ldots, p_n$.

Ustalmy przedział I i liczbę naturalną $n \geqslant 2$. Interesować nas będą średnie *wypukłe*, czyli średnie $M\colon I^n \to I$ spełniające warunek

$$M\left(t\mathbf{x} + (1-t)\mathbf{y}\right) \leqslant tM(\mathbf{x}) + (1-t)M(\mathbf{y}), \qquad \mathbf{x}, \mathbf{y} \in I^n, \tag{19}$$

dla każdego $t \in (0,1)$. Gdy liczba t jest ustalona warunek (19) definiuje średnie *t-wypukłe*. Średnie $1/2$-wypukłe nazywamy *wypukłymi w sensie Jensena*. Oczywiście wypukłość jest, na ogół, własnością silniejszą niż wypukłość w sensie Jensena. Nie mniej dla średnich mamy następujący

Twierdzenie 8

Niech M będzie średnią na przedziale otwartym I. Wtedy następujące warunki są parami równoważne:
(i) M jest wypukła,
(ii) M jest t-wypukła dla każdego $t \in (0,1)$,
(iii) M jest t-wypukła dla pewnego $t \in (0,1)$,
(iv) M jest wypukła w sensie Jensena.

Dla każdej liczby naturalnej $n \geqslant 2$ określmy zbiór Δ_n równością

$$\Delta_n = \{(p_1, \ldots, p_n) \in (0,1)^n \colon \; p_1 + \ldots + p_n = 1\}\,.$$

Twierdzenie 9

Niech $\varphi, \psi \colon I \to \mathbb{R}$ będą ściśle monotonicznymi funkcjami ciągłymi i niech $\mathbf{p}, \mathbf{q} \in \Delta_n$. Wówczas $A_{\mathbf{p}}^{\varphi} = A_{\mathbf{q}}^{\psi}$ wtedy i tylko wtedy, gdy $\psi \sim \varphi$ i $\mathbf{p} = \mathbf{q}$.

Z Twierdzenia 9 można wyprowadzić następujący rezultat.

Wniosek 10

Jeśli ważona średnia quasi-arytmetyczna jest generowana przez ściśle rosnącą funkcję wypukłą, to każdy jej ściśle rosnący generator jest wypukły.

Twierdzenie 11

Niech I będzie przedziałem otwartym, a $\varphi\colon I \to \mathbb{R}$ ściśle rosnącą funkcją ciągłą. Wtedy następujące zdania są parami równoważne:
(i) średnia $A^{\varphi}_{\mathbf{p}}$ jest wypukła dla pewnej liczby naturalnej $m \geqslant 2$ i pewnego $\mathbf{p} \in \Delta_m$,
(ii) średnia $A^{\varphi}_{\mathbf{p}}$ jest wypukła dla każdej liczby naturalnej $m \geqslant 2$ i wszystkich $\mathbf{p} \in \Delta_m$,
(iii) średnia $A^{\varphi^{-1}}_{\mathbf{q}}$ jest wklęsła dla pewnej liczby naturalnej $m \geqslant 2$ i pewnego $\mathbf{p} \in \Delta_n$,
(iv) średnia $A^{\varphi^{-1}}_{\mathbf{q}}$ jest wklęsła dla każdej liczby naturalnej $m \geqslant 2$ i wszystkich $\mathbf{p} \in \Delta_n$.

Z powyższego twierdzenia wynika, że wypukłość ważonej średniej quasi-arytmetycznej nie zależy ani od liczby zmiennych, ani od wag, zależy natomiast jedynie od generatora. Zatem badając wypukłość średniej $A^{\varphi}_{\mathbf{p}}$ z pewnym $\mathbf{p} \in \Delta_n$ i pewną liczbą $n \geqslant 2$, wystarczy zająć się najprostszym przypadkiem $n = 2$ i $\mathbf{p} = (1/2, 1/2)$, czyli średnią quasi arytmetyczną A^{φ} o wzorze $A^{\varphi}(x, y) = \varphi^{-1}\left(\frac{\varphi(x)+\varphi(y)}{2}\right)$.

Zbadamy teraz związki pomiędzy wypukłością średniej A^{φ} a wypukłością jej generatora.

Twierdzenie 12

Niech I będzie przedziałem otwartym, a $\varphi\colon I \to \mathbb{R}$ ściśle rosnącą funkcją ciągłą. Jeśli średnia $A^{\varphi}_{\mathbf{p}}$ jest wypukła to funkcja φ jest wypukła.

Poniższy przykład pokazuje, że implikacja odwrotna nie jest prawdziwa.

Przykład 13 Oczywiście wzór

$$\varphi(x) = \begin{cases} x, & x < 0, \\ 2x, & x \geqslant 0, \end{cases}$$

definiuje ściśle rosnącą funkcję wypukłą $\varphi\colon \mathbb{R} \to \mathbb{R}$, której funkcja odwrotna φ^{-1} jest dana równością

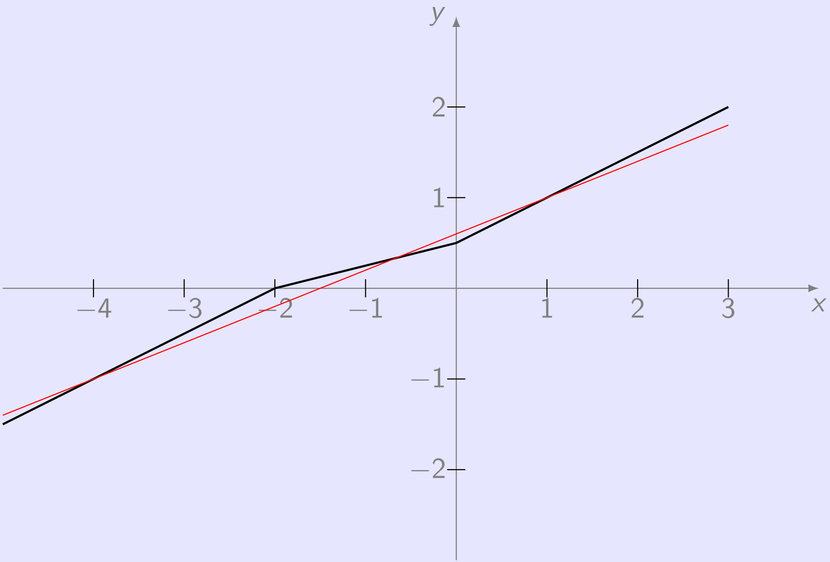
$$\varphi^{-1}(y) = \begin{cases} y, & y < 0, \\ y/2, & y \geqslant 0. \end{cases}$$

Prosty rachunek pokazuje, że średnia A^φ wyraża się wzorem

$$A^\varphi(x_1,x_2) = \begin{cases} \frac{x_1+2x_2}{2}, & \text{gdy } x_1 < 0 \leqslant x_2 \text{ i } x_1+2x_2<0, \\ \frac{x_1+2x_2}{4}, & \text{gdy } x_1 < 0 \leqslant x_2 \text{ i } x_1+2x_2 \geqslant 0, \\ \frac{x_1+x_2}{2}, & \text{gdy } x_1,x_2<0, \text{ lub } x_1,x_2 \geqslant 0, \\ \frac{2x_1+x_2}{2}, & \text{gdy } x_2 < 0 \leqslant x_1 \text{ i } 2x_1+x_2<0, \\ \frac{2x_1+x_2}{4}, & \text{gdy } x_2 < 0 \leqslant x_1 \text{ i } 2x_1+x_2 \geqslant 0. \end{cases}$$

Przyjmując tu $x_2 = 1$ otrzymujemy cięcie $A^\varphi(\cdot,1)$ funkcji A^φ:

$$A^\varphi(x_1,1) = \begin{cases} \frac{x_1+2}{2}, & \text{gdy } x_1 < -2, \\ \frac{x_1+2}{4}, & \text{gdy } -2 \leqslant x_1 < 0, \\ \frac{x_1+1}{2}, & \text{gdy } x_1 \geqslant 0. \end{cases}$$



Ponieważ cięcie $A^{\varphi}(\cdot, 1)$ nie jest funkcją wypukłą, więc funkcja A^{φ} nie jest wypukła, choć jej generator φ jest wypukły.

Przykładu takiego nie da się jednak podać, gdy generator φ jest odpowiedniej klasy regularności. Mówi o tym ostatni rezultat, który pragnę przytoczyć.

Twierdzenie 14

Niech I będzie przedziałem otwartym, a $\varphi\colon I \to \mathbb{R}$ dwukrotnie różniczkowalną w sposób ciągły funkcją o dodatniej pierwszej pochodnej. Jeśli średnia $A_{\mathbf{p}}^{\varphi}$ jest wypukła, to jej generator φ, jest albo afiniczny, albo wypukły.

[1] J. Aczél, *Lectures on functional equations and their applications*, Academic Press, New York and London, 1966.

[2] K. Baron, M. Kuczma, *Iteration of random-valued functions on the unit interval*, Colloq. Math. 37 (1977), 263–269.

[3] C.W. Borchardt, *Sur deux algorithmes analogues à celui de la moyenne arithmético-géométrique de deux éléments*, In Memoriam Dominici Chelini, Collect. Math., [etc], L. Cremona. ed. U. Hoepli, Milan, 1881, pp. 206–212; reprinted in C.W. Borchardt, *Gesammelte Werke*, Berlin, 1888, pp. 455-462.

[4] J.M. Borwein and P.B. Borwein, *Pi and the AGM. A study in analytical number theory and computational complexity*, Canadian Mathematical Society Series of Monographs and Advanced Texts, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1987.

[5] P.S. Bullen, *Handbook of Means and their Inequalities*, Kluwer Academic Publishers Group, Dordrecht, 2003.

[6] A.L. Cauchy, *Cours d'analyse de l'École Royale Polytechnique, I^{re} partie: Analyse Algébrique*, Imprimérie Royale, Paris, 1821.

[7] J. Chudziak, D. Głazowska, J. Jarczyk and W. Jarczyk, *On weighted quasi-arithmetic means which are convex*, Math. Inequal. Appl. 22 (2019), 1123-1136.

[8] D.A. Cox, *The arithmetic-geometric mean of Gauss*, Enseign. Math. 30 (1984), 275–330.

[9] Z. Daróczy and Zs. Páles, *On means that are both quasi-arithmetic and conjugate arithmetic*, Acta Math. Hungar. 90 (2001), 271–282.

[10] Z. Daróczy and Zs. Páles, *Gauss-composition of means and the solution of the Matkowski-Sutô problem*, Publ. Math. Debrecen 61 (2002), 157–218.

[11] Z. Daróczy and Zs. Páles, *The Matkowski-Sutô problem for weighted quasi-arithmetic means*, Acta Math. Hungar. 100 (2003), 237–243.

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

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

[14] G.H. Hardy, J.E. Littlewood, G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, 1934 (1st edn), 1952 (2nd edn).

[15] J. Jarczyk, *Invariance in the class of weighted quasi-arithmetic means with continuous generators*, Publ. Math. Debrecen 71 (2007), 279-294.

[16] J. Jarczyk, *Parametrized means and limit properties of their Gaussian iterations*, Appl. Math. Comput., 261 (2015), 81–89

[17] J. Jarczyk and W. Jarczyk, *Invariance of means*, Aequat. Math. 92 (2018), 801–872.

[18] J. Jarczyk and W. Jarczyk, *Gaussian iterative algorithm and integrated automorphism equation for random means*, Discrete Contin. Dyn. Syst. (2020), 6837-6844.

[19] J. Jarczyk and W. Jarczyk, *Note on generalized Archimedes-Borchardt algorithm*, w recenzji

[20] J. Jarczyk and J. Matkowski, *Invariance in the class of weighted quasi-arithmetic means*, Ann. Polon. Math. 88 (2006), 39–51.

[21] K. Knopp, *Über Reihen mit positiven Gliedern*, J. London Math. Soc. 3 (1928), 205–21.

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

[23] J.L. Lagrange, *Sur une nouvelle Méthode de Calcul Intégrale pour différentielles affectées d'un radical carre*, Mem. Acad. R. Sci. Turin II, 2 (1784-1785) 252–312.

[24] D.H. Lehmer, *On the compounding of certain means*, J. Math. Anal. Appl. 36 (1971) 183–200.

[25] Gy. Maksa and Zs. Páles, *Remarks on the comparison of weighted quasi-arithmetic means*, Colloq. Math. 120 (2010), 77–84.

[26] J. Matkowski, *Invariant and complementary quasi-arithmetic means*, Aequationes Math. 57 (1999), 87–107.

[27] J. Matkowski, *Iterations of the mean-type mappings*, In: Iteration theory (ECIT'08), A. N. Sharkovsky, I. M. Susko (eds.), Grazer Math. Ber., Bericht Nr., 354 (2009), 158–179.

[28] Mitio Nagumo, *Uber eine Klasse der Mittelwerte*, Jpn. J. Math. 7 (1930), 71–79.

[29] I.J. Schoenberg, *Mathemaical Time Exposures*, Mathematical Association of America, Washington, DC, 1982.

[30] Onosaburô Sutô, *Studies on some functional equations I*, Tôhoku Math. J. 6 (1914), 1–15.

[31] Onosaburô Sutô, *Studies on some functional equations II*, Tôhoku Math. J. 6 (1914), 82–101.