Mathematical
Inequalities
& Applications
Volume 1, Number 1 (1998), 69–83

WHY HÖLDER'S INEQUALITY SHOULD BE CALLED ROGERS' INEQUALITY

LECH MALIGRANDA

Dedicated to the memory of my mother
Jadwiga Maligranda (1924–1995)

(*communicated by L.-E. Persson*)

Abstract. The inequality

$$\sum_{k=1}^{n} a_k b_k \leqslant \left(\sum_{k=1}^{n} a_k^p\right)^{1/p} \left(\sum_{k=1}^{n} b_k^q\right)^{1/q} \tag{1}$$

was proved in slightly different form by Rogers in 1888 and then by Hölder in 1889 (Hölder even refered to Rogers!). Today everybody refer to (1) as the Hölder inequality. We will try to explain the history of this and closely related fundamental inequalities with the answer to the question: why the Rogers inequality is called the Hölder inequality? We claim that the Hölder inequality ought to be referred to as the Rogers inequality or at least as the Rogers-Hölder inequality.

The classical *Rogers–Hölder inequality* reads:

THEOREM 1. (Rogers – 1888, Hölder – 1889). *If* $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, *and* $a_k > 0$, $b_k > 0$, $k = 1, 2, \ldots, n$, *then*

$$\sum_{k=1}^{n} a_k b_k \leqslant \left(\sum_{k=1}^{n} a_k^p\right)^{1/p} \left(\sum_{k=1}^{n} b_k^q\right)^{1/q} \tag{1}$$

with equality if and only if $a_k^p = C b_k^q$ *for all* $k = 1, 2, \ldots, n$.

A proof as well as extensions, inverse or applications of this inequality can be found in many books about inequalities, real functions, analysis, functional analysis or L^p-spaces *but with the name Hölder's inequality* (cf. [1], [2], [11], [21]).

The special case of inequality (1) for $p = 2$ is most common and known as the *Cauchy–Bunyakovski–Schwarz inequality*

$$\sum_{k=1}^{n} a_k b_k \leqslant \left(\sum_{k=1}^{n} a_k^2\right)^{1/2} \left(\sum_{k=1}^{n} b_k^2\right)^{1/2}. \tag{CBS}$$

Mathematics subject classification (1991): 26D15, 01A50, 01A60, 01A70.

Key words and phrases: Rogers inequality, Hölder inequality, Bernoulli inequality, convex function, Minkowski inequality, biographies of L. J. Rogers and O. L. Hölder.

The (CBS) inequality goes back very far. It is a consequence of a well-known identity

$$\sum_{k=1}^{n} a_k^2 \sum_{k=1}^{n} b_k^2 - \left(\sum_{k=1}^{n} a_k b_k\right)^2 = \frac{1}{2}\sum_{k=1}^{n}\sum_{i=1}^{n}(a_i b_k - a_k b_i)^2, \tag{L}$$

due to Lagrange [17, p. 662–3] which he proved it in 1773 for $n = 3$. Cauchy [4, pp. 373–374] proved (CBS) inequality for finite sums in 1821 by showing (L) for every natural n. Bunyakovski [3, p. 4] proved it in 1859 for Riemann integrals and H. A. Schwarz [27] rediscovered it in 1885 by showing that the quadratic function

$$\sum_{k=1}^{n}(x a_k + b_k)^2 = x^2 \sum_{k=1}^{n} a_k^2 + 2x \sum_{k=1}^{n} a_k b_k + \sum_{k=1}^{n} b_k^2$$

is positive for any real x, which gives that the discriminant $\leqslant 0$ and this is the CBS-inequality.

Observe that these two simple proofs of the CBS-inequality fail for general $p > 1$.

L. J. Rogers [26] first proved in 1888 the following generalized inequality between the arithmetic and the geometric means: If $a_k > 0$, $x_k > 0$, $k = 1, 2, \ldots, n$, then

$$x_1^{a_1} x_2^{a_2} \ldots x_n^{a_n} \leqslant \left(\frac{a_1 x_1 + a_2 x_2 + \ldots + a_n x_n}{a_1 + a_2 + \ldots + a_n}\right)^{a_1 + a_2 + \ldots + a_n}, \tag{R}$$

which can be written in the following equivalent form: *If* $x_k > 0$, $\alpha_k > 0$, $k = 1, 2, \ldots, n$ *and* $\alpha_1 + \alpha_2 + \ldots + \alpha_n = 1$, *then*

$$\prod_{k=1}^{n} x_k^{\alpha_k} \leqslant \sum_{k=1}^{n} \alpha_k x_k. \tag{AG}$$

Equality in (AG) *holds if and only if* $x_1 = x_2 = \ldots = x_n$.

The inequality (AG) in the case $\alpha_1 = \alpha_2 = \ldots = \alpha_n = \frac{1}{n}$ gives a basic arithmetic-geometric means inequality

$$(x_1 x_2 \ldots x_n)^{1/n} \leqslant \frac{x_1 + x_2 + \ldots + x_n}{n}, \tag{C}$$

which was proved for the first time in 1821 by Cauchy [4, pp. 375-377]. Sometimes the inequality (C) is therefore called the *Cauchy inequality*. In the Bullen–Mitrinovic–Vasic book [2] there are given 52 proofs of this inequality and mathematicians still find new proofs (!).

Rogers main result in [26, p. 146] was to prove the inequality (R) by using the Cauchy inequality (C). We will show later on, in Theorem 3, that the inequality (C) and (R) and the classical Bernoulli inequality all are equivalent.

After that Rogers [26, p. 149] used the inequality (R) in the proof, as we can call now, *the Rogers inequality*: If $0 < r < s < t < \infty$ and $a_k > 0$, $b_k > 0$, $k = 1, 2, \ldots, n$, then

$$\left(\sum_{k=1}^{n} a_k b_k^s\right)^{t-r} \leqslant \left(\sum_{k=1}^{n} a_k b_k^r\right)^{t-s} \left(\sum_{k=1}^{n} a_k b_k^t\right)^{s-r}. \tag{2}$$

This inequality for $a_1 = a_2 = \ldots = a_n = 1$ is known in literature as Lyapunov's inequality (cf [19], [21]) and means that the norm in l^p -space is logarithmically convex as a function of p. Again, Lyapunov proved it in 1901, i.e., later than Rogers (!).

Rogers proof of the implication (R) $\implies$ (2) was the following: denote $S_r = \sum_{k=1}^{n} a_k b_k^r$. By putting $a_k b_k^s$ instead of a_k to the inequality (R) and $x_k = b_k^{t-s}$, and then also $x_k = b_k^{r-s}$ we obtain

$$\left(b_1^{a_1 b_1^s} b_2^{a_2 b_2^s} \ldots b_n^{a_n b_n^s}\right)^{t-s} \leqslant \left(S_t/S_s\right)^{S_s}$$

and

$$\left(b_1^{a_1 b_1^s} b_2^{a_2 b_2^s} \ldots b_n^{a_n b_n^s}\right)^{r-s} \leqslant \left(S_r/S_s\right)^{S_s},$$

or, equivalently

$$\left(S_s/S_r\right)^{S_s} \leqslant \left(b_1^{a_1 b_1^s} b_2^{a_2 b_2^s} \ldots b_n^{a_n b_n^s}\right)^{s-r}.$$

Combining these inequalities we get

$$\left(S_s/S_r\right)^{S_s/(s-r)} \leqslant \left(S_t/S_s\right)^{S_s/(t-s)}.$$

Taking the S_s th root and reducing we have

$$\left(S_s/S_r\right)^{t-s} \leqslant \left(S_t/S_s\right)^{s-r},$$

or

$$\left(S_s\right)^{t-r} \leqslant \left(S_r\right)^{t-s} \left(S_t\right)^{s-r},$$

which is exactly the inequality (2).

The (AG)-inequality can also be written as

$$-\ln\left(\sum_{k=1}^{n} \alpha_k x_k\right) \leqslant -\ln\left(\prod_{k=1}^{n} x_k^{\alpha_k}\right) = -\sum_{k=1}^{n} \alpha_k \ln x_k,$$

or equivalently

$$\varphi\left(\sum_{k=1}^{n} a_k x_x / \sum_{k=1}^{n} a_k\right) \leqslant \frac{\sum_{k=1}^{n} a_k \varphi(x_k)}{\sum_{k=1}^{n} a_k}, \quad \text{with } \varphi(u) = -\ln u. \tag{H}$$

In 1889 O. Hölder [14] proved that the inequality (H) is still true for any function φ which satisfies $\varphi'' \geqslant 0$ and the inequality (H) is reversed if $\varphi'' \leqslant 0$. His nice proof was the following: Let $\varphi'' \leqslant 0$ and denote $M = \sum_{k=1}^{n} a_k x_k / \sum_{k=1}^{n} a_k$. Then, according to Taylor's formula,

$$\varphi(x_k) = \varphi(M) + (x_k - M)\varphi'(M) + (x_k - M)^2 \frac{\varphi''(\xi_k)}{2}$$

and, thus,

$$\frac{\sum_{k=1}^{n} a_k \varphi(x_k)}{\sum_{k=1}^{n} a_k} = \varphi(M) + \frac{\frac{\sum_{k=1}^{n} a_k (x_k - M)^2 \varphi''(\xi_k)}{2}}{\sum_{k=1}^{n} a_k} \leqslant \varphi(M).$$

As we now know, due to Jensen [15], the inequalities $\varphi'' \geqslant 0$ or $\varphi'' \leqslant 0$ are the characterizations of convex or concave functions. Moreover, Jensen also established an integral version of inequality (H).

Hölder has clearly written that he, after Rogers, proved a more general version of the inequality (H), and then used it to get *Hölder's inequality* ($r = 0$, $s = 1$): If $1 < t < \infty$ and $a_k > 0$, $b_k > 0$, $k = 1, 2, \ldots, n$, then

$$\left(\sum_{k=1}^{n} a_k b_k\right)^t \leqslant \left(\sum_{k=1}^{n} a_k\right)^{t-1} \left(\sum_{k=1}^{n} a_k b_k^t\right). \tag{3}$$

The last inequality follows from the (H) inequality used to the function $\varphi(u) = u^t$, $t > 1$, for which we have $\varphi''(u) = t(t-1)u^{t-2} \geqslant 0$.

In 1902 A. Pringsheim [22, pp. 174–176] used the inequality (3) and referred here to Hölder [13, p. 44] but he also gave an elementary proof of (3), without using the differential calculus as Hölder did. His proof was based upon the generalized Bernoulli inequality (he gave here an elementary proof – cf. our Remark 1) and induction.

In 1906. J. L. W. V. Jensen [15] also showed that if φ is a continuous and $\frac{1}{2}$-convex function on an interval I, i.e. if $\varphi\left(\frac{1}{2}u + \frac{1}{2}v\right) \leqslant \frac{1}{2}\varphi(u) + \frac{1}{2}\varphi(v)$ for all $u, v \in I$, then

$$\varphi\left(\frac{\sum_{k=1}^{n} a_k x_k}{\sum_{k=1}^{n} a_k}\right) \leqslant \frac{\sum_{k=1}^{n} a_k \varphi(x_k)}{\sum_{k=1}^{n} a_k} \qquad \forall x_1, \ldots, x_n \in I \tag{J}$$

an $a_k > 0$, $k = 1, 2, \ldots, n$. Then, since $\varphi(u) = u^t$, $t > 1$, is a continuous and $\frac{1}{2}$-convex function on $[0, \infty)$ we obtain from (J) the inequality (3).

Jensen also observed that by taking a concave function $\varphi(u) = \ln u$ we obtain from (J) the Rogers inequality (R) and for the convex function $\varphi(u) = u \ln u$ we get the inequality:

$$\left(\frac{a_1 x_1 + a_2 x_2 + \ldots + a_n x_n}{a_1 + a_2 + \ldots + a_n}\right)^{a_1 x_1 + a_2 x_2 + \ldots + a_n x_n} \leqslant x_1^{a_1 x_1} x_2^{a_2 x_2} \ldots x_n^{a_n x_n}.$$

Jensen, instead of the above inequality, has written incorrectly inequality (12) on page 185 with the information that the special case of $a_1 = a_2 = \ldots = a_n = 1$ was done by Rogers. This is also not correct since it was done by Hölder [14, p. 45].

In the end of Jensen's paper [15] in addendum he was written "After having the above results I have observed that the fundamental formula (J) was not totally new, which I had believed. In the paper by Mr A. Pringsheim [22], I found a citation of a note by Mr O. Hölder [14] in which one could find the formula in question demonstrated. As a matter of fact Mr Hölder's hypothesis is rather different from mine in that he supposed that φ'' exists".

Before Jensen the notion of convexity was also used by J. Hadamard [7].

In 1907 E. Landau [18] proved the following equivalence: Let $p > 1$ and $a_k > 0$, $b_k > 0$ $(k = 1, 2, \ldots)$. Then $\sum_{k=1}^{\infty} a_k b_k < \infty$ for every sequence $\{a_k\}$ such that $\sum_{k=1}^{n} a_k^p < \infty$ if and only if $\sum_{k=1}^{n} b_k^{p/(p-1)} < \infty$. His proof, in one direction, contains the inequality

$$\left(\sum_{k=1}^{n} a_k b_k\right)^p \leqslant \sum_{k=1}^{n} a_k^p \left(\sum_{k=1}^{n} b_k^{p/(p-1)}\right)^{p-1}$$

and he referred here to Hölder [14, p. 44], Pringsheim [22, p. 174] and Jensen [15, p. 181].

In 1910 F. Riesz used the name Cauchy-Hölder inequality for the inequality (1) and referred to Landau's paper [18]. *F. Riesz was the first who obtained and used form* (1) *of the Rogers–Hölder inequality* (see also [29], p. 359). We should again mention here that Rogers proved inequality (2) and Hölder proved inequality (3). Then in 1913 F. Riesz, in his book [24, p. 43], again has written the inequality (1) with the references to Cauchy for $p = 2$ and to Hölder for $p > 1$.

In 1920 G. H. Hardy [8] (where it is given an inequality which is now well-known as the Hardy classical inequality) wrote "By the well known inequality

$$\sum ab \leqslant \left(\sum a^{\kappa}\right)^{1/\kappa} \left(\sum b^{\kappa/(\kappa-1)}\right)^{(\kappa-1)/\kappa}$$

which seems to be due to Hölder: see E. Landau [17]".

In 1927 Hellinger-Toeplitz [12] used the inequality (1) with the Hölder name but they referred here to Rogers [26], Hölder [14] and Jensen [15, p. 181].

In 1929 Hardy in the addenda to [9] said: "The proofs of Hölder's inequality, for sums and integrals, given on pp. 67–68 and 71, were based on proofs contained in the printed but unpublished notes of Prof. Littlewood's lectures, which have been in my possession for some years. They are, however, the same in principle as the proof of the inequality for integrals given by F. Riesz [25]".

In 1934 in the well known book of Hardy–Littlewood–Pólya [11], p. 25 it was written in an footnote that "Hölder states the theorem in a less symmetrical form given a little earlier by Rogers".

We can see here that the Hölder was lucky because Pringsheim, Landau and then Hardy–Littlewood–Pólya put the name Hölder's inequality instead Rogers' inequality and now almost everybody refer to it as Hölder's inequality. However, U. Dudley [6] in his book in Th. 5.1.2. put the name Rogers–Hölder inequality and he is probably the first who observed such priority. For such historical reasons we call the inequality (1) the *Rogers-Hölder inequality* (or the Rogers–Hölder–Riesz inequality). For the sake of clarity we prove now that all these three inequalities of Rogers, Hölder and Rogers–Hölder are equivalent.

THEOREM 2. (Equivalence Theorem). *The Rogers–Hölder inequality* (1)*, the Rogers inequality* (2) *and the Hölder inequality* (3) *are equivalent.*

Proof. (1) $\Longrightarrow$ (2). Take $p = \frac{t-r}{t-s}$. Then $q = \frac{s-r}{t-r}$, $\frac{r}{p} + \frac{t}{q} = s$, and

$$\left(\sum_{k=1}^{n} a_k b_k^s\right)^{t-r} = \left(\sum_{k=1}^{n} a_k^{1/p+1/q} b_k^{r/p+t/q}\right)^{t-r} = \left(\sum_{k=1}^{n} a_k^{1/p} b_k^{r/p} a_k^{1/q} b_k^{t/q}\right)^{t-r}$$

[by the Rogers-Hölder inequality (1)]

$$\leqslant \left(\sum_{k=1}^{n} a_k b_k^r\right)^{(t-r)/p} \left(\sum_{k=1}^{n} a_k b_k^t\right)^{(t-r)/q} = \left(\sum_{k=1}^{n} a_k b_k^r\right)^{t-s} \left(\sum_{k=1}^{n} a_k b_k^t\right)^{s-r}.$$

(2) $\Longrightarrow$ (1). Let $r = 1$, $1 < s < t < \infty$ and $p = (t-1)/(t-s)$. Then

$$\sum_{k=1}^{n} a_k b_k = \sum_{k=1}^{n} a_k^{t/(t-s)} b_k^{-1/(s-1)} \left(a_k^{t/(t-s)} a_k^{-1/(t-s)} b_k^{1/(s-1)}\right)^s$$

[by the Rogers inequality (2)]

$$\leqslant \left(\sum_{k=1}^{n} a_k^{t/(t-s)} b_k^{-1/(s-1)} a_k^{-1/(t-s)} b_k^{1/(s-1)}\right)^{(t-s)/(t-1)} \times$$

$$\times \left(\sum_{k=1}^{n} a_k^{t/(t-s)} b_k^{-1/(s-1)} \left(a_k^{-1/(t-s)} b_k^{1/(s-1)}\right)^t\right)^{(s-1)/(t-1)}$$

$$= \left(\sum_{k=1}^{n} a_k^{(t-1)/(t-s)}\right)^{(t-s)/(t-1)} \left(\sum_{k=1}^{n} b_k^{(t-1)/(s-1)}\right)^{(s-1)/(t-1)}$$

$$= \left(\sum_{k=1}^{n} a_k^p\right)^{1/p} \left(\sum_{k=1}^{n} b_k^q\right)^{1/q}.$$

(3) $\Longrightarrow$ (1). Let $p = \frac{t}{t-1}$. Then

$$\sum_{k=1}^{n} a_k b_k = \sum_{k=1}^{n} a_k^{t/(t-s)} b_k a_k^{-1/(t-1)}$$

[by the Hölder inequality (3)]

$$\leqslant \left(\sum_{k=1}^{n} a_k^{t/(t-s)}\right)^{1-1/t} \left(\sum_{k=1}^{n} a_k^{t/(t-s)} \left(b_k a_k^{-1/(t-1)}\right)^t\right)^{1/t}$$

$$= \left(\sum_{k=1}^{n} a_k^p\right)^{1/p} \left(\sum_{k=1}^{n} b_k^t\right)^{1/t} = \left(\sum_{k=1}^{n} a_k^p\right)^{1/p} \left(\sum_{k=1}^{n} b_k^q\right)^{1/q}.$$

(1) $\Longrightarrow$ (3). Let $p = \frac{t}{t-1}$. Then

$$\sum_{k=1}^{n} a_k b_k = \sum_{k=1}^{n} a_k^{1/p} a_k^{1-1/p} b_k$$

[by the Rogers-Hölder inequality (1)]

$$\leqslant \left(\sum_{k=1}^{n} a_k\right)^{1/p} \left(\sum_{k=1}^{n} (a_k^{1-1/p} b_k)^q\right)^{1/q} = \left(\sum_{k=1}^{n} a_k\right)^{1-1/t} \left(\sum_{k=1}^{n} a_k b_k^t\right)^{1/t}.$$

First proof of the Rögers-Hölder inequality (1). This proof for the first time was given by F. Riesz [25, pp. 78–79] in 1928.

First step. The essential point in this step is the AG-inequality for $n = 2$:

$$x^{\alpha} y^{1-\alpha} \leqslant \alpha x + (1-\alpha) y \quad \text{for all} \quad x \geqslant 0,\ y \geqslant 0 \quad \text{and} \quad 0 < \alpha < 1 \tag{4}$$

or equivalently

$$uv \leqslant \frac{u^p}{p} + \frac{v^q}{q} \quad \text{for all } u \geqslant 0,\ v \geqslant 0, \quad \text{and} \quad p > 1,\ \frac{1}{p} + \frac{1}{q} = 1. \tag{4'}$$

If either $x = 0$ or $y = 0$, then inequality (4) is obvious. Assume that $x > 0$ and $y > 0$. An application of the differential calculus to the function

$$f(x) = \alpha x + (1-\alpha) y - x^{\alpha} y^{1-\alpha},\ x > 0,$$

gives derivative

$$f'(x) = \alpha - \alpha x^{1-\alpha} y^{1-\alpha} = \alpha[1 - (y/x)^{1-\alpha}],$$

i.e., f is decreasing on $(0, y]$ and increasing on $[y, \infty)$ and so

$$f(x) \geqslant f(y) = \alpha y + (1-\alpha) y - y^{\alpha} y^{1-\alpha} = y - y = 0.$$

A longer but more elementary proof is the following: for

$$x_1 = x_2 = \ldots = x_m = x \quad \text{and} \quad x_{m+1} = x_{m+2} = \ldots = x_n = y$$

the Cauchy inequality (C) becomes

$$(x^m y^{n-m})^{1/n} \leqslant \frac{mx + (n-m)y}{n}$$

or

$$x^{m/n} y^{1-m/n} \leqslant \frac{m}{n} x + \left(1 - \frac{m}{n}\right) y,$$

where m, n are natural numbers and $1 \leqslant m \leqslant n-1$. Equality holds if all terms are equal, i.e., $x = y$.

Since any rational fraction α, $0 < \alpha < 1$, has a form $\frac{m}{n}$ our inequality (4) holds for all $x > 0$, $y > 0$ and any fraction α between 0 and 1.

If $0 < \alpha < 1$ is irrational, then there exists sequence $\{r_k\}$ of rational numbers in $(0,1)$ convergent to α. Then

$$x^{\alpha}y^{1-\alpha} = \lim_{k\to\infty} x^{r_k}y^{1-r_k} \leqslant \lim_{k\to\infty}[r_k x + (1-r_k)y] = \alpha x + (1-\alpha)y.$$

Second step. Put $u_k = \dfrac{a_k}{\left(\sum\limits_{k=1}^{n} a_k^p\right)^{1/p}}$ and $v_k = \dfrac{b_k}{\left(\sum\limits_{k=1}^{n} b_k^q\right)^{1/q}}$, $(k = 1,2,\ldots,n)$ into $(4')$. Then

$$u_k v_k \leqslant \frac{1}{p}\frac{a_k^p}{\left(\sum\limits_{k=1}^{n} a_k^p\right)} + \frac{1}{q}\frac{b_k^q}{\left(\sum\limits_{k=1}^{n} b_k^q\right)}$$

and summing over k from 1 to n we obtain

$$\sum_{k=1}^{n} u_k v_k \leqslant \frac{1}{p}\frac{\sum\limits_{k=1}^{n} a_k^p}{\sum\limits_{k=1}^{n} a_k^p} + \frac{1}{q}\frac{\sum\limits_{k=1}^{n} b_k^q}{\sum\limits_{k=1}^{n} b_k^q} = \frac{1}{p} + \frac{1}{q} = 1,$$

or

$$\sum_{k=1}^{n}\left[\frac{a_k}{\left(\sum\limits_{k=1}^{n} a_k^p\right)^{1/p}}\,\frac{b_k}{\left(\sum\limits_{k=1}^{n} b_k^q\right)^{1/q}}\right] \leqslant 1$$

which is the inequality (1).

REMARK 1. The inequality (4) is equivalent to the following generalized Bernoulli inequality (Jacob Bernoulli in 1689 proved in fact that $x^n \geqslant 1 + n(x-1)$ for every natural n): if $x \geqslant 0$ and $0 < \alpha < 1$, then

$$x^{\alpha} \leqslant 1 + \alpha(x-1). \tag{GB}$$

Namely, by substitution $\dfrac{x}{y}$ as x and multiplication by y, we obtain from (GB) the inequality (4). The inequality (GB) was already proved in 1900 by Stolz–Gmeiner in their book [28, pp. 202–204]. A simple proof of (GB) was given by Pringsheim [22, pp. 174–176] in 1902 and repeated by Hardy [9, pp. 68–69] (see also [1, pp. 12–13] and [11, pp. 40–41]).

REMARK 2. Inequality $(4')$ for u, $v > 0$ follows also easily from the convexity of the exponential function. Namely

$$\begin{aligned} uv = \exp[\ln(uv)] &= \exp\Big[\frac{1}{p}\ln u^p + \frac{1}{q}\ln v^q\Big] \\ &\leqslant \frac{1}{p}\exp[\ln u^p] + \frac{1}{q}\exp[\ln v^q] = \frac{1}{p}u^p + \frac{1}{q}v^q. \end{aligned}$$

Second proof of the Rogers–Hölder inequality (1). In this proof we use almost the Jensen argument in [15, pp. 181–182]. The function $\varphi(x) = x^p$, $p > 1$ is convex, i.e., the inequality (H) holds. By using this inequality with $a_k = b_k^q$ and $x_k = a_k b_k^{-q/p}$ we obtain

$$\left(\frac{\sum_{k=1}^{n} b_k^q a_k b_k^{-q/p}}{\sum_{k=1}^{n} b_k^q}\right)^p \leqslant \frac{\sum_{k=1}^{n} b_k^q (a_k b_k^{-q/p})^p}{\sum_{k=1}^{n} b_k^q},$$

which after reduction is

$$\left(\frac{\sum_{k=1}^{n} a_k b_k^{q-q/p}}{\sum_{k=1}^{n} b_k^q}\right)^p \leqslant \frac{\sum_{k=1}^{n} a_k^p}{\sum_{k=1}^{n} b_k^q},$$

or

$$\left(\sum_{k=1}^{n} a_k b_k\right)^p \leqslant \sum_{k=1}^{n} a_k^p \left(\sum_{k=1}^{n} b_k^q\right)^{p-1}.$$

Taking the p th root we get the inequality (1).

As we mentioned before the important Rogers inequality (R) is equivalent to the Cauchy inequality (C) but as we can see also to the classical Bernoulli inequality.

THEOREM 3. *The classical Bernoulli inequality*

$$x^n \geqslant 1 + n(x-1) \quad \forall x > 0,\ n \in N, \tag{B}$$

the Cauchy inequality

$$(x_1 x_2 \cdots x_n)^{1/n} \leqslant \frac{x_1 + x_2 + \ldots + x_n}{n} \quad \forall x_k > 0, \quad k = 1, 2, \ldots, n, \tag{C}$$

and the Rogers inequality

$$x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} \leqslant \left(\frac{a_1 x_1 + a_2 x_2 + \ldots + a_n x_n}{a_1 + a_2 + \ldots + a_n}\right)^{a_1 + a_2 + \ldots + a_n}$$
$$\forall a_k > 0, \quad x_k > 0, \quad k = 1, 2, \ldots, n, \tag{R}$$

are equivalent.

Proof. (B) $\implies$ (C). Denote $A_k = \dfrac{x_1 + x_2 + \ldots + x_k}{k}$. Since $\dfrac{A_k}{A_{k-1}} > 0$ for $k = 2, 3, \ldots, n$ it follows from the inequality (B) that

$$\left(\frac{A_k}{A_{k-1}}\right)^k \geqslant 1 + k\frac{A_k}{A_{k-1} - 1} = \frac{A_{k-1} + kA_k - kA_{k-1}}{A_{k-1}}$$
$$= \frac{kA_k - (k-1)A_{k-1}}{A_{k-1}} = \frac{x_k}{A_{k-1}},$$

i.e., $A_k^k \geqslant x_k A_{k-1}^{k-1}$.

Therefore

$$A_n^n \geqslant x_n A_{n-1}^{n-1} \geqslant x_n x_{n-1} A_{n-2}^{n-2} \geqslant \ldots \geqslant x_n x_{n-1} \ldots x_2 A_1^1 = x_n x_{n-1} \ldots x_2 x_1.$$

(C) $\Longrightarrow$ (B). For $n=1$ inequality (B) is obvious. For $n \geqslant 2$ and $0<x \leqslant 1-\frac{1}{n}$ we have $x^n > 0 \geqslant 1+n(x-1)$, i.e. (B) holds. Therefore assume that $n \geqslant 2$ and $x > 1-\frac{1}{n}$. Then $1+n(x-1) > 0$ and by using (C) with the positive numbers $1+n(x-1), 1, 1, \ldots, 1$ $[(n-1)$-times 1$]$ we obtain

$$\begin{aligned} x^n &= \left\{\frac{1+n(x-1)+1+1+\ldots+1}{n}\right\}^n \\ &\geqslant \left\{[1+n(x-1)]\cdot 1\cdot 1.\ldots\cdot 1\right\} = 1+n(x-1). \end{aligned}$$

(R) $\Longrightarrow$ (C). By putting either $a_1=a_2=\ldots=a_n=\frac{1}{n}$ or $a_1=a_2=\ldots=a_n=1$ into (R) we get immediately (C).

(C) $\Longrightarrow$ (R). This main implication was proved by Rogers [26, p. 146]. His proof was the following: Firstly when the numbers $a_1, a_2, \ldots, a_n$ are integers, then we have a_1 quantities, each equal to x_1, a_2 quantities, each equal to x_2, etc. up to the a_n quantities, each equal to x_n. Here we use the Cauchy inequality (C) to the whole number of them being $a_1+a_2+\ldots+a_n$ and obtain (R).

Secondly when the numbers $a_1, a_2, \ldots, a_n$ are fractions, then we consider the least common measure of their denominators by N to get

$$Na_1 = A_1,\ Na_2 = A_2,\ \ldots,\ Na_n = A_n.$$

Then, by the above proved step, we have

$$x_1^{A_1} x_2^{A_2} \ldots x_n^{A_n} \leqslant \left[\frac{A_1x_1+A_2x_2+\ldots+A_nx_n}{A_1+A_2+\ldots+A_n}\right]^{A_1+A_2+\ldots+A_n}.$$

Taking the Nth root of each side we get after reducing the bracketted fraction, the inequality

$$\begin{aligned} x_1^{a_1} x_2^{a_2} \ldots x_n^{a_n} &= \left(x_1^{A_1} x_2^{A_2} \ldots x_n^{A_n}\right)^{\frac{1}{N}} \\ &\leqslant \left[\frac{A_1x_1+A_2x_2+\ldots+A_nx_n}{A_1+A_2+\ldots+A_n}\right]^{\frac{A_1+A_2+\ldots+A_n}{N}}, \\ &= \left[\frac{a_1x_1+a_2x_2+\ldots+a_nx_n}{a_1+a_2+\ldots+a_n}\right]^{a_1+a_2+\ldots+a_n}, \end{aligned}$$

which is the inequality (R).

Thirdly for irrational numbers $a_1, a_2, \ldots, a_n$ we are using the approximation by rational numbers.

REMARK 3. The proofs of Rogers and Hölder from which we can get inequality (1) have the following structures:

Rogers arguments: inequality (C) $\Longrightarrow$ inequality (R) $\Longrightarrow$ (2),

Hölder arguments: $\varphi(u)=u^t$, $t>1$, is convex $\Longrightarrow$ inequality (H) $\Longrightarrow$ (3).

Probably the most elementary way to prove the Rogers-Hölder inequality (1) is to start with the classical Bernoulli inequality (B) (this inequality can be proved by only algebraic manipulations, cf. [22, pp. 174–175] or [11, p. 40]). Then, by Theorem

3, we get inequality (C) and then also inequality (R). Now it is enough to repeat the Rogers arguments for the implication (R) $\implies$ (2) and Theorem 2 will give the last consequence (2) $\implies$ (1).

REMARK 4. The Rogers-Hölder inequality (1) implies the next fundamental inequality, namely the so called *Minkowski inequality* [20, pp. 115–117]: if $p > 1$ and $a_k, b_k > 0$, $k = 1, 2, \ldots, n$ then

$$\left(\sum_{k=1}^{n}(a_k + b_k)^p\right)^{1/p} \leqslant \left(\sum_{k=1}^{n} a_k^p\right)^{1/p} + \left(\sum_{k=1}^{n} b_k^p\right)^{1/p}. \tag{5}$$

The idea of the most natural deduction of the Minkowski inequality from the Rogers-Hölder inequality is due to F. Riesz [24, pp. 45–46]. We have identity

$$S = \sum_{k=1}^{n}(a_k + b_k)^p = \sum_{k=1}^{n} a_k(a_k + b_k)^{p-1} + \sum_{k=1}^{n} b_k(a_k + b_k)^{p-1}.$$

Applying Rogers-Hölder inequality (1) to each sum, and observing that $q(p-1) = p$ we obtain

$$\begin{aligned}
S &\leqslant \left(\sum_{k=1}^{n} a_k^p\right)^{1/p}\left(\sum_{k=1}^{n}(a_k + b_k)^{(p-1)q}\right)^{1/q} + \left(\sum_{k=1}^{n} b_k^p\right)^{1/p}\left(\sum_{k=1}^{n}(a_k + b_k)^{(p-1)q}\right)^{1/q} \\
&= \left[\left(\sum_{k=1}^{n} a_k^p\right)^{1/p} + \left(\sum_{k=1}^{n} b_k^p\right)^{1/p}\right]\left(\sum_{k=1}^{n}(a_k + b_k\Big)^{p}\right)^{1/q} \\
&= \left[\left(\sum_{k=1}^{n} a_k^p\right)^{1/p} + \left(\sum_{k=1}^{n} b_k^p\right)^{1/p}\right] S^{1/q},
\end{aligned}$$

which is (5). There is no difficulty to find conditions giving equality in (5): $a_k = Cb_k$ for $k = 1, 2, \ldots, n$.

REMARK 5. The Minkowski inequality (5) also follows easy from the so called *quasi-linearization*, i.e., the equality

$$\max\left\{\sum_{k=1}^{n} a_k b_k : b_k \geqslant 0 \quad \text{and} \quad \sum_{k=1}^{n} b_k^q = 1\right\} = \left(\sum_{k=1}^{n} a_k^p\right)^{1/p}. \tag{QL}$$

The inequality $\leqslant$ in (QL) follows immediately from the Rogers–Hölder inequality (1) and the equality in (QL) is attained when $b_k = \dfrac{a_k^{p-1}}{\left(\sum_{k=1}^{n} a_k^p\right)^{1/q}}$, $k = 1, 2, \ldots, n$.

REMARK 6. The fundamental inequalities (1) and (5) are reversed for $p < 1$, $p \neq 0$.

REMARK 7. The inequalities proved for finite sums may often be extended to integrals by the usage of simple functions and the limiting processes. For example, is easy to extend all the inequalities (1), (2), (3) and (5) to the corresponding integral ineqaulities.

My priority discussion seems to confirm the Boyer's Law "mathematical formulas and theorems are usually not named after their original discoverers" (cf. [16]) or the L. Fejér's nice setence "the history of mathematics serves to prove that nobody has discovered anything: there was always somebody who knew it before".

I would prefer to say, in such situations, that some mathematicians are lucky and some other unlucky in getting their names to the results they proved. Hölder here was lucky and Rogers unlucky to get his name in the inequality (1). But *I hope that all my above discussion will motivate mathematicians to use at least the name Rogers–Hölder inequality for the inequality* (1).

In the end of the above commentaries I would like to say something about both authors of these inequalities since their personalities are not so well-known. My informations are taken mainly from articles of Dixon [5], Ernst Hölder [13] (the son of Otto Hölder) and the obituary by van der Waerden [30].

Leonard James Rogers was born on 30 March, 1862, in Oxford, where his father, Thorold Rogers, was Professor of Political Economy. In his childhood he had a serious illnes, and, though his recovery was complete, he was not sent to school. Mr. J. Griffith, of Jesus College, himself a well known Oxford mathematician with a strong interest in elliptic functions, noticed Rogers' marked mathematical ability, and taught him during his boyhood. In 1879 he was elected to a Scholarship in Mathematics at Balliol College, and he matriculated in October, 1880. Besides first classes in the Mathematical Schools, and the Senior and Junior Mathematical Scholarship, he took a second class in Classical Moderations in 1882, and the degree of Bachelor of Music in 1884. In the period 1888-1919 he was Professor of Mathematics at Yorkshire College, now the University of Leeds. His very serious illnes obliged him to retire in 1919. He made a remarkable recovery, however, and returned to live in Oxford, where he continued his mathematical work, did a little teaching and examining, and increased his fame as a gifted musician. He was elected a Fellow of the London Royal Society in 1924. He died on 12 September, 1933, in Oxford.

Rogers was a man of extraordinary gifts in many fields, and everything he did, he did well. Besides his mathematics and music he had many interests; he was a born linguist and phonetician, a wonderful mimic who delighted to talk broad Yorkshire, a first-class skater, and a maker of rock gardens. Hi did things well because he liked doing them. Music was the first necessity in his intellectual life, and after that came mathematics. He had very little ambition or desire for recognition.

Rogers' most important work in mathematics was done in transformation and manipulation of theta function series and products. Such names as *Rogers–Ramanujan identities, Rogers–Ramanujan continued fractions* and *Rogers transformations* are known in the theory of partitions and combinatorics.

The Rogers–Ramanujan identities are, for example,

$$1+\sum_{m=1}^{\infty}\frac{q^{m^2}}{(1-q)(1-q^2)\dots(1-q^m)}=\prod_{m=0}^{\infty}\frac{1}{(1-q^{5m+1})(1-q^{5m+4})},$$

$$1+\sum_{m=1}^{\infty}\frac{q^{m^2+m}}{(1-q)(1-q^2)\dots(1-q^m)}=\prod_{m=0}^{\infty}\frac{1}{(1-q^{5m+2})(1-q^{5m+3})},$$

$$4\int_0^{\infty}\frac{xe^{-x\sqrt{5}}}{\cosh x}\,dx=\frac{1|}{|1}+\frac{1^2|}{|1}+\frac{1^2|}{|1}+\frac{2^2|}{|1}+\frac{2^2|}{|1}+\dots.$$

They were discovered by Rogers in 1894 and discovered by Ramanujan in 1913 and I. Schur in 1917. We can quote here Hardy who has written on page 91 of [10]:

"The formulae have a very curious history. They were found first in 1894 by Rogers, a mathematician of great talent but comparatively little reputation, now remembered mainly from Ramanujan's rediscovery of his work. Rogers was a fine analyst, whose gifts were, on a smaller scale, not unlike Ramanujan's; but no one paid much attention to anything he did, and the particular paper in which he proved the formulae was quite neglected.

Ramanujan rediscovered the formulae sometime before 1913. He had then no proof (and knew that he had none), and none of the mathematicians to whom I comunicated the formulae could find one. They are therefore stated without proof in the second volume of MacMahon's Combinatory Analysis.

The mistery was solved, trebly, in 1917. In that year Ramanujan, looking through old volumes of the Proceedings of the London Mathematical Society, came accidentally across Rogers' paper. I can remember very well his suprise, and the admiration which he expressed for Rogers' work. A correspondence followed in the course of which Rogers was led to a considerable simplification of this original proof".

Rogers published more than thirty papers in mathematics.

Otto Ludwig Hölder was born on 22 December, 1859, in Stuttgart, where his father, Otto Hölder was a Professor of French language at the Polytechnikum. His mother was the former Pauline Ströbel. In Stuttgart, he attended one of the first Gymnasium devoted to science and there he studied engineering for a short time. In 1877 he came to Berlin at the University of Berlin, where he attended lectures by K. Weierstrass, L. Kronecker, and E. E. Kummer. Already in 1877 he attended the Weierstrass lecturing on the theory of functions covering the fundamentals of analysis. Weierstrass did a big impression on Hölder and this fact strongly influented Hölder's way of thinking in the future. Later he studied in Tübingen. Influenced by the rigorous foundation of analysis given by Weierstrass, he developed the continuity condition for volume density which now is well-known as the "*Hölder condition*" on a function f, i.e. the inequality $|f(x)-f(y)| \leqslant |x-y|^{\alpha}$ for all x and y from the domain of f with $0<\alpha \leqslant 1$. It appeared in his Ph. D. Dissertation "Beiträge zur Potentialtheorie", which he presented at Tübingen in 1882. His referee was Paul du Bois-Reymond.

Next he investigated analytic functions and summation procedures by arithmetic means. His generalization of the summation method of arithmetical averages introduced in 1882 is called now the *Hölder summation method*. After completing fis Ph. D. in Tübingen he went to Leipzig.

In his Habilitationsschrift submitted in 1884 at Götingen, Hölder examined the convergence of the Fourier series of a function that was not assumed to be either continuous or bounded. The notion of "improper integral" (uneigentliches Integral) appeared here. He become also a lecturer in 1884. Five years later he discovered the inequality (3) which in form (1) now is named *Hölder inequality* but as we know inequality (1) ought to have the name *Rogers–Hölder inequality*.

Hölder became interested in group theory through Kronecker and Klein. He proved the uniqueness of the factor groups in a composition series, the theorem which now is called *Jordan-Hölder theorem*. In 1892 Hölder initiated the range problem in group theory, i.e. classification of all simple groups whose orders are in a given range. Hölder proved that the only two simple groups whose orders lie between 1 and 200 are A_5 of order 60 and $\mathrm{PSL}(2,7)$ of order 168. He considered his method to be "of some interest so long as we do not possess a better one suitable for handling the problem generally". Such a general method is still lacking, despite the progress and great efforts of a recent years.

In further works Hölder treated the structure of composite groups having the orders: p^3, pq^2, pqr, p^4, where p,q,r are primes, and n, where n is square-free.

The first third of Hölder's career in research was the most fruitful. A period of depression seems to have occured at Königsberg, where he succeded Minkowski in 1894. He was happy to leave that city in 1899, when he accepted an offer from Leipzig to succeed Sophus Lie. In the same year he married Helene Lautenschläger, who also came from Stuttgart.

In 1899–1928 Hölder was a Professor of the Leipzig University.

From 1900 he became interested in the geometry of the projective line and axioms for physics (see his books [1] and [2] below). Between 1914 and 1923 these interests lead him to the lexico-philosophical studies of the foundations of mathematics which are included in his book [3] below.

In his last years one of Hölder's favorite topics was elementary number theory — his third great teacher in Berlin had been Kummer.

In 1899 Hölder was elected a memeber of the Saxon Academy of Sciences. He was active in the academy and for several years served as president. He was also a memeber of the Prince Jablonowski Society. In 1927 Hölder became a corresponding member of the Bavarian Academy of Sciences.

He was editor of the Mathematische Annalen in the period 1908–1928. He retired in 1928.

To the last years of his life he showed an indefatigable interests to mathematics. He died on 29 August, 1937, in Leipzig.

As stated by van der Waerden anybody who come in close contact with Hölder have respected his inteligence, umblemished character and nobility.

Hölder was the author over 60 papers and three books in German:

1. *Anschauungen und Denken in der Geometrie*, Teubner, Leipzig 1900, 75 pages (also Darmstad 1968),

2. *Die Arithmetik in strenger Begründung. Programmabhandlung der philosophischen Fakultät*, Teubner, Leipzig 1914, IV+74 pages (2nd ed., Springer, Berlin 1929),

3. *Die Mathematische Methode. Logisch-erkenntnistheoretische Untersuchungen im Gebiete der Mathematik, Mechanik und Physik* [Mathematical Methods. Logic Investigations in the Fields of Mathematics, Mechanics and Physics], Berlin 1924, X+563 pages.

REFERENCES

[An asterisk denotes that I have not seen this article in original form but I have found a discussion in secondary sources such as citation by other authors.]

[1] E. F. BECKENBACH AND R. BELLMAN, *Inequalities*, Springer–Verlag, 1983.
[2] P. S. BULLEN, D. S. MITRINOVIC AND P. M. VASIC, *Means and Their Inequalities*, D. Reidel Publishing Company, Dordrecht, 1988.
[3]* V. Y. BUNYAKOVSKI, *Sur quelques inequálités concernant les intégrales ordinaires et les intégrales aux différences finites*, Mémoires de l'Acad. de St. Pétersbourg **(VII)1** (1859), No. 9, 1–18.
[4]* A. L. CAUCHY, *Cours d'Analyse de l'École Royale Polytechnique*, Ire partie: *Analyse Algébrique*, Paris, Debure frères, 1821 Also in *Oeuvres Complètes d'Augustin Cauchy*, Sér II, T. 3, Gauthier–Villars, Paris 1987.
[5] A. L. DIXON, *Leonard James Rogers*, J. London Math. Soc. **9** (1934), 237–240.
[6] U. DUDLEY, *Real Analysis and Probability*, Wadsworth, 1989.
[7] J. HADAMARD, *Etude sur les propriété des fonctions entières et en particulier d'une fonction considérée par Riemann*, J. Math Pures Appl. Sér. IV **T. 9** (1983), 171–215.
[8] G. H. HARDY, *Note on a theorem of Hilbert*, Math. Zeit. **6** (1920), 314–317.
[9] ———, *Prolegomena to a chapter on inequalities*, J. London Math. Soc. **4** (1929), 61–78 and addenda **5** (1929), 80; Also in *Collected papers of G. H. Hardy*, Vol II, Oxford Univ. Press 1967, 471–489.
[10] ———, *Ramanujan*, Chelsea Publ. Co., New York, 1940.
[11] G. H. HARDY, J. E. LITTLEWOOD AND G. PÓLYA, *Inequalities*, Cambridge Univ. Press, 1934.
[12] E. HELLINGER AND O. TOEPLITZ, *Integralgleichungen und Gleichungen mit Unendlichvielen Unbekannten*, Sonderausgabe der Encyklopädie der Mathematischen Wissenschaften, 1927, Also Chelsea Publishing Comapany, New York 1953.
[13] E. HÖLDER, *Otto Ludwig Hölder*, Dictionary of Scientific Biography, Vol. 5, Charles Scribner's Sons, New York, 1981, 472–474.
[14] O. HÖLDER, *Über einen Mittelwertssatz*, Nachr. Akad. Wiss. Göttingen Math.–Phys. Kl., 1889, pp. 38–47.
[15] J. L. W. V. JENSEN, *Sur les fonctions convexes et les inégalités entre les valeurs moyennes*, Acta Math. **30** (1906), 175–193.
[16] H. C. KENNEDY, *Who discovered Boyer's law?*, Amer. Math. Monthly **79** (1972), 66–67.
[17]* J. L. LAGRANGE, *Nouv. Mém. Acad. Berlin*, 1773, Also in Oeuvres, T. 3, Gauthier–Villars, Paris 1869, p. 662f.
[18] E. LANDAU, *Über einen Konvergenzsatz*, Göttinger Nachrichten (1907), 25–27, Also in *Edmund Landau Collected Works*, Vol. 3, Thales Verlag, Essen 1985, 273–275.
[19]* A. M. LYAPUNOV (LIAPUNOFF), *Nouvelle forme du théorème sur la limite de probabilité*, Memoires de l'Acad. de St. Pétersbourg **(VIII) 12** (1901), No. 5, 1–24.

[20] H. MINKOWSKI, *Geometrie der Zahlen I,*, Leipzig, 1896, pp. 115–117.
[21] D. S. MITRINOVIC, *Analytic Inequalities*, Springer–Verlag, 1970.
[22] A. PRINGSHEIM, *Zur theorie der ganzen transzendenten Funktionen*, Sitzungsberichte der Mathematisch-Physikalischen Klasse der Königlich Bayerischen Akademie der Wissenschaften **32** (1902), 163–192.
[23] F. RIESZ, *Unterschungen über Systeme integrierbarer Funktionen*, Math. Ann. **69** (1910), 449–497.
[24] ———, *Les Systèmes D'Équations Linéaires a Une Infinite D'Inconnues*, Gauthier–Villars, Paris, 1913.
[25] ———, *Su alcune disuguaglianze*, Bollettino dell'Unione Mat. Italiana **7** (1928), 77–79.
[26] L. J. ROGERS, *An extension of a certain theorem in inequalities*, Messenger of Math. **17** (1888), 145–150.
[27]* H. A. SCHWARZ, *Über ein die Flächen kleisten Flächeninhalts betreffendes Problem der Variationsrechnung*, Acta Soc. Scient. Fenn. **15** (1885), 315–362; Also in *Gesammelte Mathematische Abhandlungen*, Band I, Springer, Berlin 1890, 224–269.
[28] O. STOLZ AND J. A. GMEÍNER, *Theoretische Arithmetik I*, Teubner, Leipzig, 1900.
[29] B. SZ.-NAGY, *Introduction to Real Functions and Orthogonal Expansions*, Oxford Univ. Press, New York, 1965.
[30] B. L. VAN DER WAERDEN, *Nachruf auf Otto Hölder*, Math. Ann. **116** (1939), 157–165.

(Received September 12, 1997)

Lech Maligranda
Department of Mathematics
Luleå University
S – 971 87 Luleå
Sweden
e-mail: lech@sm.luth.se