The Prehistory of the Hardy Inequality

Alois Kufner, Lech Maligranda, and Lars-Erik Persson

1. INTRODUCTION. The development of the famous Hardy inequality (in both its discrete and continuous forms) during the period 1906–1928 has its own history or, as we have called it, prehistory. Contributions of mathematicians other than G. H. Hardy, such as E. Landau, G. Pólya, I. Schur, and M. Riesz, are important here. In this article we describe some of those contributions. We also include and comment upon several facts and early proofs that are not available in many references on this subject.

We consider the following statement of the *Hardy inequality*: the *discrete inequality* asserts that if $p > 1$ and $\{a_k\}_1^\infty$ is a sequence of nonnegative real numbers, then

$$\sum_{n=1}^\infty \left(\frac{1}{n} \sum_{k=1}^n a_k \right)^p \le \left(\frac{p}{p-1} \right)^p \sum_{n=1}^\infty a_n^p ; \tag{1}$$

the *continuous inequality* informs us that if $p > 1$ and f is a nonnegative p-integrable function on $(0, \infty)$, then f is integrable over the interval $(0, x)$ for each positive x and

$$\int_0^\infty \left(\frac{1}{x} \int_0^x f(t)\, dt \right)^p dx \le \left(\frac{p}{p-1} \right)^p \int_0^\infty f(x)^p\, dx. \tag{2}$$

Several introductory remarks are in order:

(i) Inequalities (1) and (2) are the standard forms of the Hardy inequalities that can be found in many textbooks on analysis and were highlighted first in the famous book *Inequalities* by Hardy, Littlewood, and Pólya [16].

(ii) By restricting (2) to the class of step functions one proves easily that (2) implies (1). This important fact seems to have been mentioned first by Landau (see [11, p. 154] and section 8 of the present paper).

(iii) The constant $(p/(p-1))^p$ in both (1) and (2) is *sharp*: it cannot be replaced with a smaller number such that (1) and (2) remain true for all revelant sequences and functions, respectively.

(iv) Inequalities (1) and (2) imply the following *weak forms* of (1) and (2): if $\sum_{n=1}^\infty a_n^p < \infty$ and $a_n \ge 0$, then $\sum_{n=1}^\infty \left(\frac{1}{n} \sum_{k=1}^n a_k \right)^p < \infty$, and if $\int_0^\infty f(x)^p\, dx < \infty$ and $f(x) \ge 0$, then $\int_0^\infty \left(\frac{1}{x} \int_0^x f(t)dt \right)^p dx < \infty$, respectively. We adopt Hardy's terminology because it has been significant in the historical development that we are going to describe (see, for example, the comments at the end of Hardy's paper [9]).

(v) Inequalities (1) and (2) together with statement (iii) imply the important information that the *discrete Hardy operator h* and the *continuous Hardy operator H*, defined by

$$h(\{a_n\}) = \left\{ \frac{1}{n} \sum_{k=1}^n a_k \right\}, \quad Hf(x) = \frac{1}{x} \int_0^x f(t)\, dt,$$

map the spaces l_p into l_p and L_p into L_p, respectively ($p > 1$), and each has norm $p' = p/(p-1)$. Here, as usual, the spaces l_p and L_p are the Lebesgue spaces consisting of

all sequences $a = \{a_n\}$ of real numbers and all (equivalence classes modulo equality almost everywhere of) measurable functions f on $(0, \infty)$, respectively, such that

$$\|a\|_{l_p} := \left(\sum_{n=1}^{\infty} |a_n|^p \right)^{1/p} < \infty, \quad \|f\|_{L_p} := \left(\int_0^{\infty} |f(x)|^p \, dx \right)^{1/p} < \infty.$$

We mention that the spaces l_p and L_p were introduced and investigated in 1910 by F. Riesz [32].

(vi) Inequalities (1) and (2) have been generalized and applied in analysis and in the theory of differential equations. A lot of these developments, generalizations, and applications have been discussed in books (e.g., [16], [22], [29]) and, more recently, in the historical survey paper [21].

The aim of this paper is to complement the existing literature devoted to this subject—for instance, what is described in, say, [16], [21], [22], or [29]—by describing some important steps in the scientific developments that finally led G. H. Hardy to (2) and, subsequently, to the proof of (1) in his famous 1925 paper [11]. The story that we are going to relate is much more dramatic and intricate then we ever imagined. We hope that the reader will find it equally captivating. The tale begins over a hundred years ago and plays itself out in the period 1906–1928.

2. HARDY'S ORIGINAL MOTIVATION: THE HILBERT INEQUALITY. In this section we briefly discuss an inequality that was discovered in the early 1900s by David Hilbert (see his paper [18] from 1906) and is closely related to the discrete Hardy inequality (1). Moreover, we confirm that this inequality was, in fact, the main source of motivation for Hardy when he started the research described in this paper.

In its most basic form the *Hilbert inequality* reads: if $\sum_{m=1}^{\infty} a_m^2 < \infty$ and $\sum_{n=1}^{\infty} b_n^2 < \infty$, where $a_m \geq 0$ and $b_n \geq 0$, then the double series

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{a_m b_n}{m+n}$$

converges (the weak form). More precisely, the inequality

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{a_m b_n}{m+n} \leq \pi \left(\sum_{m=1}^{\infty} a_m^2 \right)^{1/2} \left(\sum_{n=1}^{\infty} b_n^2 \right)^{1/2} \tag{3}$$

holds, with π as the sharp constant. The determination of the sharp constant π, as well as the integral analogue of (3), are due to Schur [35] (in Hilbert's version of (3) the constant 2π appears in place of π). We remark that the following more general form of (3) is sometimes referred to in the literature as the Hilbert inequality:

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{a_m b_n}{m+n} \leq \frac{\pi}{\sin(\frac{\pi}{p})} \left(\sum_{m=1}^{\infty} a_m^p \right)^{1/p} \left(\sum_{n=1}^{\infty} b_n^{p'} \right)^{1/p'}, \tag{4}$$

where $p > 1$ and $p' = p/(p-1)$. Hilbert himself was not even close to considering this case (the l_p-spaces appeared only in 1910). It was M. Riesz and G. H. Hardy who took the first steps towards a proof of (4). In fact, Hardy acknowledged in [12] that Riesz had pointed out to him that his result in the paper [10] (see Theorem 3) actually

implied the weak form of (4). In any event, the Hilbert inequality and its evaluation into what today are called Hilbert-type inequalities have their own interesting history, of which we recall here only a few basic facts from the beginning of the story.

In his research on solutions to certain integral equations Hilbert was led to study special bilinear forms with sequences of real or complex numbers as entries. For details we refer the reader to the book [**19**] or to the section on integral equations edited by Hellinger [**17**, pp. 94–145] in Hilbert's *Collected Works* [**20**]. Moreover, several ideas from Hilbert's lectures around the period 1906–1908 can be found in Hermann Weyl's 1908 dissertation [**39**]. In particular, on page 83 Weyl presented and discussed the following remarkable formula that had been discovered by Hilbert:

$$\sum_{n=1}^{N} \sum_{m=1}^{N} \left(\frac{1}{n+m} + \frac{1}{n-m} \right) a_m b_n$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} t \left[\sum_{k=1}^{N} (-1)^k \left(a_k \sin kt - b_k \cos kt \right) \right]^2 dt \qquad (5)$$

(if $n = m$, then $1/(n-m)$ is understood to be zero). This formula implies the finite version of the Hilbert inequality

$$\sum_{n=1}^{N} \sum_{m=1}^{N} \frac{a_m b_n}{m+n} \leq \pi \left(\sum_{m=1}^{N} a_m^2 \right)^{1/2} \left(\sum_{n=1}^{N} b_n^2 \right)^{1/2}. \qquad (6)$$

The derivation of (6) from (5) is found in [**16**, pp. 235–236].

Other proofs of the Hilbert inequality were given by Wiener [**38**] and Schur [**35**] (the latter proved it with the help of the so-called Schur test). We should also mention Toeplitz's method, which is based on the identity

$$\sum_{n=1}^{N} \sum_{m=1}^{N} \frac{a_m b_n}{n+m} = \frac{i}{2\pi} \int_{0}^{2\pi} (t - \pi) \left[\sum_{n=1}^{N} a_n e^{int} \sum_{m=1}^{N} b_m e^{imt} \right] dt$$

([**37**]; see also [**36**, p. 165]). In addition, Fejér and Riesz [**5**] (see also [**16**, p. 235]) and Pólya and Szegő [**31**] gave proofs that exploited the theory of analytic functions. We are most interested, however, in Hardy's method for proving (3).

Note that (6) implies both (3) and a weaker cousin that was of special interest to Hardy: *if $\sum_{n=1}^{\infty} a_n^2 < \infty$ ($a_n \geq 0$), then the double series*

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{a_m a_n}{m+n}$$

converges. In fact, Hardy wrote in the introduction to [**10**]:

It was proved by Hilbert, in the course of his theory of integral equations, that the double series $\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{a_m a_n}{m+n}$ ($a_n \geq 0$) is convergent whenever $\sum_{n=1}^{\infty} a_n^2$ is convergent. Of this theorem, which is one of the simplest and most beautiful in the theory of double series of positive terms, at least five essentially different proofs have been published. Hilbert's own proof, which depends upon the theory of Fourier's series, is outlined by Weyl in his Inaugural-Dissertation [**39**]. Another proof was given by Wiener [**38**], and two more by Schur [**35**]; but none of these proofs is as simple and elementary as might be desired.

To these four proofs I added recently a fifth, which seemed to me to lack nothing of this simplicity. I observed first that Hilbert's theorem is an immediate corollary of another theorem which seems of some interest in itself.

(See Theorem A in [9] but note that this result was already proved in the 1915 paper [8]; compare with Theorem 2 in the next section.)

Hardy made similar pronouncements in some of the other papers that we are now going to discuss. *Therefore it seems completely clear that Hardy's original motivation when he began the research that culminated in his discovery of inequalities* (1) *and* (2) *was to prove* (*the weak form of*) *the Hilbert inequality.*

3. THE 1915 PAPER. In [8] Hardy stated and applied the following theorem, which has obvious connections to the Hilbert theorem:

Theorem 1. *Let* $a_n \geq 0$ *and* $A_n = \sum_{k=1}^{n} a_k$. *The convergence of any of the three series*

$$(i) \ \sum_{n=1}^{\infty} \frac{a_n A_n}{n} \quad (ii) \ \sum_{n=1}^{\infty} \left(\frac{A_n}{n}\right)^2 \quad (iii) \ \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{a_n a_m}{n+m}$$

implies that of the others.

Hardy actually proved the equivalence of the convergence of the three integrals

$$I_1 = \int_a^{\infty} \frac{f(x)F(x)}{x} \, dx, \quad I_2 = \int_a^{\infty} \left(\frac{F(x)}{x}\right)^2 \, dx,$$

$$I_3 = \int_a^{\infty} \int_a^{\infty} \frac{f(x)f(y)}{x+y} \, dx \, dy,$$

where $a > 0$, f is nonnegative and integrable on (a, ∞), and $F(x) = \int_a^x f(t) \, dt$, by establishing the estimates

$$\int_a^x \left(\frac{F(t)}{t}\right)^2 dt \leq 2 \int_a^x \left(\frac{f(t)F(t)}{t}\right) dt$$

$$\leq \int_a^x \left(\frac{F(t)}{t}\right)^2 dt + 4 \int_x^{2x} \left(\frac{F(t)}{t}\right)^2 dt$$

and observing that $I_1 \leq I_3 \leq 2I_1$. In the proof of the sequence case he said only that "the proof of this theorem is much like that for integrals" [8, p. 164], but he probably realized later that the sequence case is more delicate, because he considered it again in the 1919 paper [9]. However, the estimation of (iii) by (i) can be done, in a manner similar to what Hardy did in [8] for integrals, as follows:

$$\sum_{n=1}^{N} \sum_{m=1}^{N} \frac{a_m a_n}{m+n} \leq 2 \sum_{n=1}^{N} \sum_{m=1}^{n} \frac{a_m a_n}{m+n} = 2 \sum_{n=1}^{N} a_n \sum_{m=1}^{n} \frac{a_m}{m+n}$$

$$\leq 2 \sum_{n=1}^{N} a_n \frac{1}{n} \sum_{m=1}^{n} a_m = 2 \sum_{n=1}^{N} a_n \frac{A_n}{n}.$$

Moreover, Hardy also established the following result, which may be regarded as a precursor of (1) for the case $p = 2$:

Theorem 2. *The convergence of the series $\sum_{n=1}^{\infty} a_n^2$ with $a_n \geq 0$ implies that of $\sum_{n=1}^{\infty} (A_n/n)^2$, where $A_n = \sum_{k=1}^{n} a_k$.*

Proof. Let $\{a_n^*\}$ denote the nonincreasing rearrangement of the sequence $\{a_n\}$ and put $A_n^* = \sum_{k=1}^{n} a_k^*$. Since $A_n^* \geq A_n$ and $\sum_{n=1}^{\infty} a_n^2 = \sum_{n=1}^{\infty}(a_n^*)^2$, it is sufficient to prove that $\sum_{n=1}^{\infty}(A_n^*/n)^2$ is convergent. Thus, we may assume without loss of generality that the sequence $\{a_n\}$ is nonincreasing. Moreover, according to Theorem 1, it is enough to prove that $\sum_{n=1}^{\infty} a_n A_n/n$ is convergent. The last series can be written in the form $S = \sum_{n=1}^{\infty} \sum_{m=1}^{n} a_m a_n/n$, and it will be convergent if $\sum_{k=1}^{\infty} S_k < \infty$, where

$$S_k = \sum_{k \leq \frac{n}{m} < k+1} \sum \frac{a_m a_n}{n} \qquad (k = 1, 2, \dots).$$

Moreover, by making obvious estimates and using the Cauchy-Schwarz inequality, we find that

$$S_k \leq \frac{1}{k} \sum_{k \leq \frac{n}{m} < k+1} \sum \frac{a_m a_n}{m} = \frac{1}{k} \sum_{m=1}^{\infty} \frac{a_m}{m} \sum_{k \leq \frac{n}{m} < k+1} a_n$$

$$\leq \frac{1}{k} \sum_{m=1}^{\infty} \frac{a_m}{m} m a_{km} \leq \frac{1}{k} \left(\sum_{m=1}^{\infty} a_m^2 \right)^{1/2} \left(\sum_{m=1}^{\infty} a_{km}^2 \right)^{1/2}$$

$$\leq \frac{1}{k} \left(\sum_{m=1}^{\infty} a_m^2 \right)^{1/2} \left(\frac{1}{k} \sum_{m=1}^{\infty} a_m^2 \right)^{1/2} = \frac{1}{k^{3/2}} \sum_{m=1}^{\infty} a_m^2.$$

Therefore

$$S = \sum_{k=1}^{\infty} S_k \leq \sum_{k=1}^{\infty} \frac{1}{k^{3/2}} \sum_{m=1}^{\infty} a_m^2 < 3 \sum_{m=1}^{\infty} a_m^2 < \infty,$$

so $S < \infty$ and the proof is complete. ∎

Remark 1. We see that the foregoing argument is not sufficient to deliver a proof of (1) for $p = 2$ directly. It gives only the (weaker) inequality

$$\sum_{n=1}^{\infty} \frac{a_n A_n}{n} \leq C \sum_{n=1}^{\infty} a_n^2$$

with $C = \zeta(3/2) \approx 2.61$, where ζ denotes the Riemann zeta-function.

Remark 2. Hardy also pointed out the fact that had motivated him to do this work, namely, that, by virtue of the equivalence theorem (Theorem 1), Theorem 2 is essentially equivalent to (the weak form of) Hilbert's theorem in the case $p = 2$.

Remark 3. In this paper, as we have already mentioned, Hardy also stated and proved continuous versions of Theorems 1 and 2. However, he formulated his results for integrals $\int_a^{\infty}$ or $\int_a^x$ with $a > 0$ instead of in the final form $\int_0^{\infty}$ or $\int_0^x$. He continued this practice in the papers [9] and [10], but we have no clear explanation for why he first studied only this case. In any event, in the 1925 paper he formulated his result in the natural way and also explained the connection to the earlier formulation.

4. THE 1919 PAPER. Probably the most important contribution of [**9**] was the new proof of Theorem 2 that it contained. In fact, the paper also includes a first proof of inequality (1) for the case $p = 2$ and even gives the best constant 4 (see Remark 4), although Hardy did not make explicit mention of this fact.

New proof of Theorem 2. It is clear that

$$\left(\frac{A_n}{n}\right)^2 = \left(a_n + \frac{A_n}{n} - a_n\right)^2 \le 2a_n^2 + 2\left(\frac{A_n}{n} - a_n\right)^2$$

$$= 4a_n^2 + 2\left(\frac{A_n}{n}\right)^2 - 4\frac{a_n A_n}{n}.$$

Hence,

$$\sum_{n=1}^{N}\left(\frac{A_n}{n}\right)^2 \le 4\sum_{n=1}^{N} a_n^2 + 2\sum_{n=1}^{N}\left(\frac{A_n}{n}\right)^2 - 4\sum_{n=1}^{N}\frac{a_n A_n}{n} \qquad (7)$$

for each N. Moreover, $-2a_n A_n = -\left(A_n^2 - A_{n-1}^2\right) - a_n^2 \le -\left(A_n^2 - A_{n-1}^2\right)$, so

$$-2\sum_{n=1}^{N}\frac{a_n A_n}{n} \le -\sum_{n=1}^{N}\frac{\left(A_n^2 - A_{n-1}^2\right)}{n}$$

$$= -\frac{A_1^2}{1 \cdot 2} - \frac{A_2^2}{2 \cdot 3} - \cdots - \frac{A_{N-1}^2}{(N-1) \cdot N} - \frac{A_N^2}{N}$$

$$\le -\sum_{n=1}^{N}\frac{1}{n(n+1)}A_n^2.$$

By substituting this estimate into (7) we obtain

$$\sum_{n=1}^{N}\left(\frac{A_n}{n}\right)^2 \le 4\sum_{n=1}^{N} a_n^2 + 2\sum_{n=1}^{N}\left(\frac{A_n}{n}\right)^2 - 2\sum_{n=1}^{N}\frac{1}{n(n+1)}A_n^2$$

$$= 4\sum_{n=1}^{N} a_n^2 + 2\sum_{n=1}^{N}\frac{1}{n^2(n+1)}A_n^2,$$

which yields

$$\sum_{n=1}^{N}\left(1 - \frac{2}{n+1}\right)\left(\frac{A_n}{n}\right)^2 \le 4\sum_{n=1}^{N} a_n^2. \qquad (8)$$

Obviously (8) implies the statement in Theorem 2. ■

Remark 4. Hardy had plainly not yet realized that inequality (8) can be used to derive the discrete Hardy inequality (1) for $p = 2$ and do so with the sharp constant 4 (more on this in section 6).

Remark 5. In [9] Hardy also stated some results for the continuous case (e.g., a generalization of Theorem 1). However, the most important point for the history of the Hardy inequality was his claim that

$$\int_a^\infty \left(\frac{1}{x} \int_a^x f(t)\,dt \right)^2 dx \le 4 \int_a^\infty f^2(x)\,dx$$

when $a > 0$ and that 4 is the best constant. In fact, it can be shown that this estimate implies (2) for $p = 2$. Hardy did not give a proof of this assertion but only referred to the proof in the discrete case that we have just presented. This can seem a little curious if we take into account the information in Remark 4.

5. AN IMPORTANT CONTRIBUTION FROM RIESZ.
In his paper [10] from 1920 Hardy referred to a letter from Marcel Riesz and wrote [10, p. 315]:

> Dr. M. Riesz to whom I recently communicated Theorem 2 at once found another proof, which is equal to mine in simplicity and which seems to both of us more natural and therefore preferable. His proof naturally suggests an interesting generalization, viz.

Hardy proceeded to formulate the Riesz result in the following weak form:

Theorem 3 (M. Riesz). *If* $p > 1$, $a_n \ge 0$, *and* $\sum_1^\infty a_n^p$ *is convergent, then* $\sum_1^\infty (A_n/n)^p$ *is convergent, where* $A_n = \sum_{k=1}^n a_k$.

Proof. Let $\Phi_n = n^{-p} + (n+1)^{-p} + (n+2)^{-p} + \cdots$. Then, by partial summation, we have for each N (with $A_0 = 0$)

$$\sum_{n=1}^N \left(\frac{A_n}{n} \right)^p = \sum_{n=1}^N A_n^p (\Phi_n - \Phi_{n+1}) = \sum_{n=1}^N (A_n^p - A_{n-1}^p)\Phi_n - A_N^p \Phi_{N+1}$$

$$\le \sum_{n=1}^N (A_n^p - A_{n-1}^p)\Phi_n \le p \sum_{n=1}^N a_n A_n^{p-1}\Phi_n.$$

Moreover,

$$\Phi_n < n^{-p} + \int_n^\infty x^{-p}\,dx = n^{-p} + \frac{n^{-(p-1)}}{p-1} \le \frac{p}{p-1}n^{-(p-1)}.$$

From these estimates and the Hölder inequality,[1]

$$\sum_{n=1}^\infty a_n b_n \le \left(\sum_{n=1}^\infty a_n^p \right)^{1/p} \left(\sum_{n=1}^\infty b_n^q \right)^{1/q} \qquad \left(p > 1, \frac{1}{p} + \frac{1}{q} = 1 \right), \qquad (9)$$

---

[1] In [10] Hardy remarked that inequality (9) was probably due to Hölder and referred to a paper [23] by Landau. In the Hardy-Littlewood-Pólya book it is said that "Hölder states the theorem in a less symmetrical form given a little earlier by Rogers" [16, p. 25] and uses the name *Hölder inequality*. We believe that these words by Hardy could have been important when mathematicians later began to call (9) the Hölder inequality. The second author pointed out in [28] that an equivalent variant of (9) had been proved by Rogers in 1888, one year earlier than Hölder produced his version, which is again only an equivalent variant of (9) and is different from Rogers's. Inequality (9) in this precise form was proved in 1910 by F. Riesz [32]. Therefore this classical inequality could have been called the Hölder-Rogers inequality or the Hölder-Rogers-Riesz inequality.

it follows that

$$\sum_{n=1}^{N} \left(\frac{A_n}{n}\right)^p \leq \frac{p^2}{p-1} \sum_{n=1}^{N} a_n \left(\frac{A_n}{n}\right)^{p-1}$$

$$\leq \frac{p^2}{p-1} \left(\sum_{n=1}^{N} a_n^p\right)^{1/p} \left(\sum_{n=1}^{N} \left(\frac{A_n}{n}\right)^p\right)^{1-(1/p)}.$$

Hence

$$\sum_{n=1}^{N} \left(\frac{A_n}{n}\right)^p \leq \left(\frac{p^2}{p-1}\right)^p \sum_{n=1}^{N} a_n^p. \tag{10}$$

∎

Remark 6. Riesz's argument actually yields more than what Hardy formulated in Theorem 3; namely, (10) implies the correctness of (1) with the constant $(p^2/(p-1))^p$ in place of $(p/(p-1))^p$.

6. THE 1920 PAPER. In [**10**] Hardy observed that estimates in the proof of (10) were fairly rough and that the constant $C_p = \left(p^2/(p-1)\right)^p$ could be improved merely by refining the Riesz estimates. In particular, he pointed out that C_p could be replaced with the strictly smaller constant $(p\zeta(p))^p$. The argument that confirms this fact he received in a letter from Schur [**34**]. Obviously, Hardy already believed by then that $(p/(p-1))^p$ was the sharp constant, even if he did not claim so explicitly (see [**11**, p. 154]). One reason for this was surely Schur's remark in the same letter that this was at least true for $p = 2$.

Hardy himself did not directly observe that inequality (8) from his 1919 paper could actually be used to derive inequality (1) for $p = 2$ with the best constant $C = 4$. We don't know for certain what Schur's argument was, but the information in the next section convinces us that it went approximatively as follows: Let $c_n = 1 - 2/(n+1)$ and for $m = 2, 3, \ldots$ let

$$a_1 = a_2 = \cdots = a_m = b_1, \ a_{m+1} = a_{m+2} = \cdots = a_{2m} = b_2, \ldots,$$

$$a_{(N-1)m+1} = a_{(N-1)m+2} = \cdots = a_{Nm} = b_N.$$

Then from inequality (8) with Nm in place of N we obtain

$$4m \sum_{n=1}^{N} b_n^2 \geq (c_1 + \cdots + c_m) \left(\frac{B_1}{1}\right)^2 + (c_{m+1} + \cdots + c_{2m}) \left(\frac{B_2}{2}\right)^2 + \cdots$$

$$+ \left(c_{(N-1)m+1} + \cdots + c_{Nm}\right) \left(\frac{B_N}{N}\right)^2,$$

where $B_n = \sum_{k=1}^{n} b_k$. Dividing by m and letting $m \to \infty$, we find that

$$(c_1 + c_2 + \cdots + c_m)/m \to 1,$$

$$(c_{m+1} + c_{m+2} + \cdots + c_{2m})/m \to 1,$$

etc.

Accordingly,

$$\sum_{n=1}^{N} \left(\frac{B_n}{n}\right)^2 \leq 4 \sum_{n=1}^{N} b_n^2,$$

which, in particular, implies (1) for $p = 2$ with the best constant.

A significant element in the paper is the formulation of the following preliminary version of (2) (see also [**11**, p. 150]):

Theorem 4. *If $p > 1$, $a > 0$, $f(x) \geq 0$, and $\int_a^\infty f(x)^p \, dx$ is convergent, then*

$$\int_a^\infty \left(\frac{\int_a^x f(t)dt}{x}\right)^p dx \leq \left(\frac{p}{p-1}\right)^p \int_a^\infty f(x)^p \, dx, \tag{11}$$

and the constant $(p/(p-1))^p$ is sharp.

Hardy did not prove inequality (11) in the 1920 paper. However, he pointed out the fact that from consideration of the function $f(x) = x^{-(1/p)-\epsilon}$, where ϵ is a sufficiently small positive constant, it follows that the constant $(p/(p-1))^p$ is sharp. He also claimed that he could prove that the corresponding best constant in the discrete case could not be strictly less than $(p/(p-1))^p$, but he hesitated for the moment to assert that (1) held with the constant $(p/(p-1))^p$.

7. A LETTER FROM LANDAU TO HARDY. The letter [**24**] from Landau to Hardy dates from June 21, 1921. It is surprising that this letter was officially published in [**26**] five years later than the letter of Landau to Schur (see also Remark 8). The reason for this long delay is not obvious, but it is clear that the contents of the letter [**24**] are of interest, for it gives a proof of (1) with the sharp constant $(p/(p-1))^p$ and this had not been published prior to that time. The main result proved in this letter reads:

Theorem 5. *Let $p > 1$, $a_n \geq 0$, and $A_n = \sum_{k=1}^{n} a_k$. Then the inequality*

$$\sum_{n=1}^{N} \left(\frac{A_n}{n}\right)^p \leq \left(\frac{p}{p-1}\right)^p \sum_{n=1}^{N} a_n^p \tag{12}$$

holds for all N in $\mathbf{N}$ or $N = \infty$. Moreover, the constant is sharp for $N = \infty$.

Inequality (12) is sometimes called the *Hardy-Landau inequality* (see [**27**, p. 188]).

Proof. For $y \geq 0$ we have

$$y^p - py + p - 1 \geq 0.$$

This is seen to be the so-called Bernoulli inequality when we write it in the form $y^p \geq 1 + p(y - 1)$. By using this elementary inequality with $y = y_1/y_2$ we find that

$$y_1^p - py_1 y_2^{p-1} + (p-1)y_2^p \geq 0.$$

We set $y_1 = b_n$ and $y_2 = (p-1)B_n/(pn)$, where $B_n = \sum_{k=1}^{n} b_k$, and learn that

$$b_n^p - pb_n \left(\frac{p-1}{p} \frac{B_n}{n} \right)^{p-1} + (p-1) \left(\frac{p-1}{p} \frac{B_n}{n} \right)^p \geq 0,$$

so

$$\sum_{n=1}^{N} b_n^p - \left(\frac{p-1}{p} \right)^{p-1} \sum_{n=1}^{N} pb_n \left(\frac{B_n}{n} \right)^{p-1} + (p-1) \left(\frac{p-1}{p} \right)^p \sum_{n=1}^{N} \left(\frac{B_n}{n} \right)^p \geq 0.$$

Moreover, $pb_n B_n^{p-1} = pB_n^{p-1}(B_n - B_{n-1}) \geq B_n^p - B_{n-1}^p$, whence by partial summation

$$\sum_{n=1}^{N} pb_n \left(\frac{B_n}{n} \right)^{p-1} \geq \sum_{n=1}^{N} \left(B_n^p - B_{n-1}^p \right) \frac{1}{n^{p-1}}$$

$$\geq \sum_{n=1}^{N} B_n^p \left(\frac{1}{n^{p-1}} - \frac{1}{(n+1)^{p-1}} \right) \geq (p-1) \sum_{n=1}^{N} B_n^p \frac{1}{(n+1)^p}.$$

Combining the two inequalities we discover that

$$\sum_{n=1}^{N} b_n^p \geq \left(\frac{p-1}{p} \right)^p \sum_{n=1}^{N} B_n^p \left(\frac{p}{(n+1)^p} - \frac{p-1}{n^p} \right) = \left(\frac{p-1}{p} \right)^p \sum_{n=1}^{N} c_n \left(\frac{B_n}{n} \right)^p,$$

where $c_n = p(1 + \frac{1}{n})^{-p} - p + 1 \to 1$ when $n \to \infty$. Next we use the argument from the previous section, putting

$$b_1 = b_2 = \cdots = b_m = a_1, \quad b_{m+1} = b_{m+2} = \cdots = b_{2m} = a_2, \ldots,$$

$$b_{(N-1)m+1} = b_{(N-1)m+2} = \cdots = b_{Nm} = a_N$$

and replacing N with Nm to conclude that

$$m \left(\frac{p-1}{p} \right)^p \sum_{n=1}^{N} a_n^p \geq (c_1 + c_2 \cdots + c_m) \left(\frac{A_1}{1} \right)^p$$

$$+ (c_{m+1} + c_{m+2} + \cdots + c_{2m}) \left(\frac{A_2}{2} \right)^p$$

$$+ \cdots + (c_{(N-1)m+1} + c_{(N-1)m+2} + \cdots + c_{Nm}) \left(\frac{A_N}{N} \right)^p.$$

Dividing by m and then letting $m \to \infty$ we note that $(c_1 + c_2 + \cdots + c_m)/m \to 1$, $(c_{m+1} + c_{m+2} + \cdots + c_{2m})/m \to 1$, and so forth, which means that (12) holds for all finite N (hence remains valid when $N \to \infty$).

In order to prove that $(p/(p-1))^p$ is the sharp constant for $N = \infty$ we consider $a_n = n^{-1/p-\varepsilon}$ ($0 < \varepsilon < 1 - 1/p$). For this choice of a_n

$$A_n = \sum_{k=1}^{n} k^{-1/p-\varepsilon} > \int_1^n x^{-1/p-\varepsilon} \, dx$$

$$= \frac{1}{1 - 1/p - \epsilon} \left(n^{1-1/p-\varepsilon} - 1 \right) > \frac{p}{p-1} \left(n^{1-1/p-\varepsilon} - 1 \right),$$

implying that

$$\left(\frac{A_n}{n}\right)^p > \left(\frac{p}{p-1}\right)^p n^{-1-\varepsilon p}\left(1 - \frac{1}{n^{1-1/p-\varepsilon}}\right)^p$$

$$\geq \left(\frac{p}{p-1}\right)^p n^{-1-\varepsilon p}\left(1 - \frac{p}{n^{1-1/p-\varepsilon}}\right)$$

$$= \left(\frac{p}{p-1}\right)^p \left(n^{-1-\varepsilon p} - p n^{-2+1/p+\varepsilon-\varepsilon p}\right).$$

Furthermore,

$$\sum_{n=1}^{N}\left(\frac{A_n}{n}\right)^p > \left(\frac{p}{p-1}\right)^p \left(\sum_{n=1}^{N} a_n^p - p\sum_{n=1}^{N}\frac{1}{n^{2-1/p-\varepsilon+\varepsilon p}}\right)$$

$$= \left(\frac{p}{p-1}\right)^p \left(\sum_{n=1}^{N} a_n^p - p C_{N,\varepsilon}\right),$$

where $C_{N,\varepsilon} \to C$ as $N \to \infty$ for any $\varepsilon > 0$ because $2 - 1/p - \varepsilon + \varepsilon p > 1$. Thus

$$\sum_{n=1}^{N}\left(\frac{A_n}{n}\right)^p \Big/ \sum_{n=1}^{N} a_n^p > \left(\frac{p}{p-1}\right)^p \left(1 - p C_{N,\varepsilon} \Big/ \sum_{n=1}^{N} a_n^p\right) \to \left(\frac{p}{p-1}\right)^p,$$

since $\sum_{n=1}^{N} a_n^p = \sum_{n=1}^{N} n^{-1-\varepsilon p} \to \infty$ as $N \to \infty$ and $\varepsilon \to 0^+$. The sharpness assertion is thereby established. Note that the foregoing calculation still works when $\varepsilon = 0$. ∎

Remark 7. In his letter [24] Landau also mentioned that equality in (12)—and, consequently, in the discrete Hardy inequality (1)—occurs if and only if $a_n = 0$ for all n.

Remark 8. It is not evident how the letter [24] from Landau to Hardy (dated June 21, 1921) and the letter [26] from Landau to Schur (dated June 22, 1921) are related, but to judge by the information that Hardy provides in [11] and the published form that appears in [26], they must be very similar, perhaps even the same.

8. THE 1925 PAPER. In the introduction to [11] Hardy first formulated Theorem 4 and wrote the following:

> I did not give a proof, being occupied primarily with the corresponding theorem for infinite series, and did not state in what sense the integrations are effected. Prof. E. Landau has recently recalled my attention to this note and I give here a proof of the theorem in a more precise form.

(see Theorem 6). Later on in the paper he also volunteered the following information [11, p. 154]:

> In a letter [24] dated 21 June 1921, Prof. E. Landau communicated to me a direct proof of (1), which gives the correct value $(p/(p-1))^p$ of the constant. He also pointed out that if the integral theorem were extended to the case $a = 0$, then the theorem for series, with the correct value of the constant, may be deduced at once by taking $f(x) = a_1, 0 \leq x < 1, f(x) = a_2, 1 \leq x < 2, \ldots$.

Indeed, assuming, without loss of generality, that we are dealing with a decreasing function f or, equivalently, a decreasing sequence $a_1 \geq a_2 \geq \cdots$ and observing that the function

$$\frac{1}{x} \int_0^x f(t)\, dt = \frac{\sum_{k=1}^{n-1} a_k + a_n(x - n + 1)}{x}$$

is decreasing on $[n - 1, n]$, we obtain

$$\sum_{n=1}^{\infty} \left(\frac{\sum_{k=1}^{n} a_k}{n} \right)^p \leq \sum_{n=1}^{\infty} \int_{n-1}^{n} \left(\frac{\sum_{k=1}^{n-1} a_k + a_n(x - n + 1)}{x} \right)^p dx$$

$$= \int_0^{\infty} \left(\frac{1}{x} \int_0^x f(t)\, dt \right)^p dx \leq \left(\frac{p}{p - 1} \right)^p \int_0^{\infty} f(x)^p\, dx$$

$$= \left(\frac{p}{p - 1} \right)^p \sum_{n=1}^{\infty} a_n^p.$$

Hardy adds a comment concerning a different letter [**11**, p. 154]:

> In a more recent letter (13 Dec. 1924) [**25**], he shows how to deduce the integral theorem for $a = 0$ from that for $a > 0$ (by a method resembling that of Prof. Pólya) and so reduce the series theorem to dependence of the latter.

With this information in mind, Hardy formulated and proved his famous inequality (2) in the following form:

Theorem 6. *Let $p > 1$ and let $f(x) \geq 0$ be p-integrable on $(0, \infty)$. Then $F(x) = \int_0^x f(t)dt < \infty$ for every $x > 0$ and*

$$\int_0^{\infty} \left(\frac{F(x)}{x} \right)^p dx \leq \left(\frac{p}{p - 1} \right)^p \int_0^{\infty} f(x)^p dx. \tag{13}$$

The constant $(p/(p - 1))^p$ is sharp.

Hardy himself remarked in [**11**, p. 150] that if $f(x) = 0$ when $x < a$, then his previous version in [**10**] (Theorem 4) follows from Theorem 6. Hardy's original proof contained fairly many technical details and explanations, but in a postscript to the proof he pointed out an essential simplification suggested by Pólya. We present here a proof that closely follows Hardy's original ideas but avoids many technical details by appealing to Pólya's simplification.

Proof. By partial integration and the identity $d/dx\,(F(x)^p) = pF(x)^{p-1}f(x)$, which holds for almost all x in $(0, \infty)$, we obtain for arbitrary α and A with $0 < \alpha < A < \infty$:

$$\int_\alpha^A \left(\frac{F(x)}{x} \right)^p dx = -\frac{1}{p - 1} \int_\alpha^A F(x)^p \frac{d}{dx}(x^{1-p})\, dx$$

$$= \frac{\alpha^{1-p}}{p - 1} F(\alpha)^p - \frac{A^{1-p}}{p - 1} F(A)^p + \frac{1}{p - 1} \int_\alpha^A x^{1-p} \frac{d}{dx}(F(x)^p)\, dx$$

$$\leq \frac{\alpha^{1-p}}{p - 1} F(\alpha)^p + \frac{p}{p - 1} \int_\alpha^A \left(\frac{F(x)}{x} \right)^{p-1} f(x)\, dx.$$

Moreover, invoking the continuous version of the Hölder inequality (9), we see that

$$\int_\alpha^A \left(\frac{F(x)}{x}\right)^{p-1} f(x)\,dx \leq \left(\int_\alpha^A f(x)^p\,dx\right)^{1/p} \left(\int_\alpha^A \left(\frac{F(x)}{x}\right)^p\,dx\right)^{(p-1)/p}.$$

Choosing β such that $\alpha \leq \beta \leq A$ and applying the preceding two inequalities to $F(x) - F(\alpha)$ instead of $F(x)$, we find that

$$\int_\alpha^A \left(\frac{F(x) - F(\alpha)}{x}\right)^p\,dx$$

$$\leq \frac{p}{p-1}\int_\alpha^A \left(\frac{F(x) - F(\alpha)}{x}\right)^{p-1} f(x)\,dx$$

$$\leq \frac{p}{p-1}\left(\int_\alpha^A f(x)^p\,dx\right)^{\frac{1}{p}} \left(\int_\alpha^A \left(\frac{F(x) - F(\alpha)}{x}\right)^p\,dx\right)^{(p-1)/p}.$$

Hence

$$\left(\int_\alpha^A \left(\frac{F(x) - F(\alpha)}{x}\right)^p\,dx\right)^{1/p} \leq \frac{p}{p-1}\left(\int_\alpha^A f(x)^p\,dx\right)^{1/p},$$

and a fortiori

$$\left(\int_\beta^A \left(\frac{F(x) - F(\alpha)}{x}\right)^p\,dx\right)^{1/p} \leq \frac{p}{p-1}\left(\int_0^\infty f(x)^p\,dx\right)^{1/p}.$$

In this inequality we first let $\alpha \to 0^+$ and observe that $F(x) - F(\alpha)$ increases to $F(x)$. To finish the proof we let $A \to \infty$ and $\beta \to 0^+$. ∎

Remark 9. Hardy had already drawn attention to the sharpness of the constant $(p/(p-1))^p$ in his 1920 paper. In that paper he decided not to mention it as part of the theorem that he stated, but he did include the details of the proof of the sharpness (by using a modified form of the example he had considered in the 1920 paper).

9. FURTHER CONTRIBUTIONS FROM THE 1925 PAPER. In addition to the main result (Theorem 6), the 1925 paper included a number of interesting results that have exerted significant influence on research related to Hardy-type inequalities. In particular, Hardy proved that the following variant of (1) holds even if the standard arithmetic mean $(1/n)\sum_{k=1}^n a_k$ is replaced with a more general arithmetic means (with general weights):

Theorem 7. *Suppose that $a_n \geq 0$ and $\lambda_n > 0$, that $A_n = \lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_n a_n$ and $\Lambda_n = \lambda_1 + \lambda_2 + \cdots + \lambda_n$ for $n = 1, 2, \ldots$, and that $\sum_{n=1}^\infty \lambda_n a_n^p$ is convergent. Then*

$$\sum_{n=1}^\infty \lambda_n \left(\frac{A_n}{\Lambda_n}\right)^p \leq \left(\frac{p}{p-1}\right)^p \sum_{n=1}^\infty \lambda_n a_n^p. \tag{14}$$

Hardy proved this theorem by applying Theorem 6 to suitable step functions (see Landau's observation that was mentioned just prior to Theorem 6). Hardy also noticed that if one replaces a_n with $a_n^{1/p}$ in the inequality (14) and lets $p \to \infty$, then $(p/(p-1))^p \to e$ and

$$\lim_{p \to \infty} \left(\frac{\lambda_1 a_1^{1/p} + \lambda_2 a_2^{1/p} + \cdots + \lambda_n a_n^{1/p}}{\lambda_1 + \lambda_2 + \cdots + \lambda_n} \right)^p = \left(a_1^{\lambda_1} a_2^{\lambda_2} \ldots a_n^{\lambda_n} \right)^{1/\Lambda_n}.$$

The scale of power means decreases to the geometric mean. Therefore Hardy arrived at the following limit result of his Theorem 7:

Theorem 8. *If $\sum_{n=1}^{\infty} \lambda_n a_n$ is convergent, then*

$$\sum_{n=1}^{\infty} \lambda_n \left(a_1^{\lambda_1} a_2^{\lambda_2} \ldots a_n^{\lambda_n} \right)^{1/\Lambda_n} \le e \sum_{n=1}^{\infty} \lambda_n a_n, \tag{15}$$

and the constant e is sharp.

In the standard case where each $\lambda_n = 1$ (15) becomes

$$\sum_{n=1}^{\infty} (a_1 a_2 \ldots a_n)^{1/n} \le e \sum_{n=1}^{\infty} a_n, \tag{16}$$

which is the natural limit inequality of (1). Inequality (16) was first proved by Carleman in 1922 (see [1]), whence its name: the *Carleman inequality*. This inequality has been generalized and applied in several ways and has its own interesting history. Carleman's original proof, which was quite long, involved Lagrange multipliers. It must have come as a big surprise for him to see the simple proof derived from the Hardy inequalities. He no doubt learned of it quickly, because Hardy was engaged in a collaboration with Carleman at the time (see, for example, their joint paper [2]).

By carrying out a similar limiting process in (2) we obtain the following inequality:

$$\int_0^{\infty} \exp \left(\frac{1}{x} \int_0^x \ln f(t) \, dt \right) dx \le e \int_0^{\infty} f(x) \, dx, \tag{17}$$

where $f(x)$ is strictly positive and measurable on each finite interval $(0, x)$. We should mention that in the original paper [11] inequality (17) appears without the constant e and with $\exp f(x)$ in place of $f(x)$.

Hardy acknowledged that it was Pólya who made him aware of the elegant limiting argument that directly implied the inequalities (15), (16), and (17). Sometimes (17) is called the Knopp inequality, although Knopp's paper on the subject dates from 1928. The name *Pólya-Knopp inequality* seems now to have gained acceptance in the literature (see [22] and the references given there).

10. ELLIOT'S AND INGHAM'S PROOFS; THE COPSON INEQUALITY. In 1926 Elliott [4] gave a simple and very elegant proof of (1) (see also [16, p. 240] and [6, p. 50]): if we set $\alpha_n = A_n/n$ and $\alpha_0 = 0$, then from the Young inequality

$$uv \le \frac{u^{p'}}{p'} + \frac{v^p}{p}$$

we obtain

$$\alpha_n^p - \frac{p}{p-1}\alpha_n^{p-1}a_n = \alpha_n^p - \frac{p}{p-1}\left[n\alpha_n - (n-1)\alpha_{n-1}\right]\alpha_n^{p-1}$$

$$= \alpha_n^p\left(1 - \frac{np}{p-1}\right) + \frac{(n-1)p}{p-1}\alpha_n^{p-1}\alpha_{n-1}$$

$$\leq \alpha_n^p\left(1 - \frac{np}{p-1}\right) + \frac{(n-1)}{p-1}\left[(p-1)\alpha_n^p + \alpha_{n-1}^p\right]$$

$$= \frac{1}{p-1}\left[(n-1)\alpha_{n-1}^p - n\alpha_n^p\right].$$

Summing from 1 to N yields

$$\sum_{n=1}^{N}\left(\frac{A_n}{n}\right)^p - \frac{p}{p-1}\sum_{n=1}^{N}\left(\frac{A_n}{n}\right)^{p-1}a_n \leq -\frac{N\alpha_N^p}{p-1} \leq 0, \tag{18}$$

so from Hölder's inequality (9) we infer that

$$\sum_{n=1}^{N}\left(\frac{A_n}{n}\right)^p \leq \frac{p}{p-1}\sum_{n=1}^{N}\left(\frac{A_n}{n}\right)^{p-1}a_n \leq \frac{p}{p-1}\left(\sum_{n=1}^{N}a_n^p\right)^{1/p}\left(\sum_{n=1}^{N}\left(\frac{A_n}{n}\right)^p\right)^{1/p'}.$$

Division by the last factor leads to (1). ∎

Two years later Grandjot [7] derived the identity

$$\sum_{n=1}^{N}\left(\frac{A_n}{n}\right)^2 - 2\sum_{n=1}^{N}\frac{A_na_n}{n} = -\frac{A_N^2}{N} - \sum_{n=2}^{N}(n-1)\left(\frac{A_n}{n} - \frac{A_{n-1}}{n-1}\right)^2 \tag{19}$$

by observing that

$$\frac{2a_nA_n}{n} - \left(\frac{A_n}{n}\right)^2 = \frac{A_n^2}{n} - \frac{A_{n-1}^2}{n-1} + (n-1)\left(\frac{A_n}{n} - \frac{A_{n-1}}{n-1}\right)^2$$

when $n \geq 2$. This gives rise to another proof of (1) in the case $p = 2$.

Ingham also found a simple proof of (2) (see [16, p. 243]). Since

$$Hf(x) = \frac{1}{x}\int_0^x f(t)\,dt = \int_0^1 f(tx)\,dt,$$

it follows from the Minkowski integral inequality that

$$\left(\int_0^\infty (Hf(x))^p\,dx\right)^{1/p} = \|Hf\|_p = \left\|\int_0^1 f(tx)\,dt\right\|_p$$

$$\leq \int_0^1 \|f(tx)\|_p\,dt = \int_0^1\left(\int_0^\infty f(tx)^p\,dx\right)^{1/p}dt$$

$$= \int_0^1\left(\int_0^\infty f(s)^p\frac{ds}{t}\right)^{1/p}dt = \frac{p}{p-1}\left(\int_0^\infty f(s)^p\,ds\right)^{1/p}. \quad ∎$$

In 1927 Copson [**3**] proved Hardy's Theorem 7 by adapting Elliott's proof and by bringing into play the dual to Hardy's inequality, a result now known as the *Copson inequality*: if $p > 1$, $a_n \geq 0$, $\lambda_n > 0$, and $\sum_{n=1}^{\infty} \lambda_n a_n^p$ is convergent, then

$$\sum_{n=1}^{\infty} \lambda_n \left(\sum_{k=n}^{\infty} \frac{\lambda_k a_k}{\sum_{m=1}^{k} \lambda_m} \right)^p \leq p^p \sum_{n=1}^{\infty} \lambda_n a_n^p, \qquad (20)$$

and the constant p^p is best possible. The Copson inequality in the case where $\lambda_n = 1$ for each n asserts that

$$\sum_{n=1}^{\infty} \left(\sum_{k=n}^{\infty} \frac{a_k}{k} \right)^p \leq p^p \sum_{n=1}^{\infty} a_n^p \qquad (21)$$

when $p > 1$ and that the constant p^p is sharp. Hardy [**9**] had earlier stated a weak version of (21) in the case $p = 2$, and therefore (21) is sometimes called the Copson-Hardy inequality. Hardy [**13**] was the first to remark upon the duality between (20) and (14).

11. FINAL DISCUSSION.
Usually mathematicians present results in a clear and polished form that is well suited for applications and further research. Nonetheless, it is also well known that a lot of creative work, questening, collaboration, and sometimes even failure arise during the process of coming to the final formulations of results. The history that we have described of the evolution of Hardy's famous inequalities can serve as an unusually good illustration of this fact. We ourselves were quite surprised and fascinated by several of the details we have presented here. In particular, we feel that as a by-product of writing this article we have acquired improved insight into Hardy's way of thinking and acting, and we hope that we have been able to communicate to the reader at least some flavor of this experience.

Through this understanding it has become clear to us that G. H. Hardy (a) had good contacts with other mathematicians who were interested in the subject and who passed along to him significant information in various ways (e.g., via private or formal letters), (b) was a real master at developing important parts of the theory himself, (c) was very good at cleverly synthesizing knowledge gained from numerous sources, and (d) played a central role in the developments described in this paper. In particular, it is totally appropriate that his name adorns inequalities (1) and (2). However, it has to be said that other mathematicians also made very important contributions to this development (e.g., E. Landau, G. Pólya, M. Riesz, and I. Schur). For example, if the results of these individuals had been published in a different way, we might today refer to the discrete inequality (1) as the Riesz or the Landau-Riesz or the Hardy-Landau-Riesz inequality.

REFERENCES

1. T. Carleman, Sur les fonctions quasi-analytiques, in *Fifth Scandinavian Congress of Mathematicians* (Helsinki, 1922), Akadem. Buchh., Helsinki, 1923, pp. 181–196.
2. T. Carleman and and G. H. Hardy, Fourier series and analytic functions, *Proc. Royal Soc. A* **101** (1922) 124–133.

3. E. T. Copson, Note on series of positive terms, *J. London Math. Soc.* **2** (1927) 9–12.
4. E. B. Elliott, A simple exposition of some recently proved facts as to convergency, *J. London Math. Soc.* **1** (1926) 93–96.
5. L. Fejér and F. Riesz, Über einige funktionaltheoretische Ungleichungen, *Math. Z.* **11** (1925) 305–314.
6. E. C. Francis and J. E. Littlewood, *Examples in Infinite Series*, Deighton Bell, Cambridge, 1928.
7. K. Grandjot, On some identities relating to Hardy's convergence theorem, *J. London Math. Soc.* **3** (1928) 114–117.
8. G. H. Hardy, Notes on some points in the integral calculus, XLI. On the convergence of certain integrals and series, *Messenger of Math.* **45** (1915) 163–166.
9. ———, Notes on some points in the integral calculus, LI. On Hilbert's double-series theorem, and some connected theorems concerning the convergence of infinite series and integrals, *Messenger of Math.* **48** (1919) 107–112.
10. ———, Notes on a theorem of Hilbert, *Math. Z.* **6** (1920) 314–317.
11. ———, Notes on some points in the integral calculus, LX. An inequality between integrals, *Messenger of Math.* **54** (1925) 150–156.
12. ———, Note on a theorem of Hilbert concerning series of positive terms, *Proc. London Math. Soc.* **23** (1925) 45–46.
13. ———, Remarks on three recent notes in the journal, *J. London Math. Soc.* **3** (1928) 166–169.
14. ———, Prolegomena to a chapter on inequalities, *J. London Math. Soc.* **4** (1929) 61–78.
15. G. H. Hardy, J. E. Littlewood, and G. Pólya, The maximum of a certain bilinear form, *Proc. London Math. Soc.* **25** (1926) 265–282.
16. ———, *Inequalities,* 2nd ed., Cambridge University Press, Cambridge, 1967.
17. E. Hellinger, Hilberts Arbeiten über Integralgleichungssysteme und unendliche Gleichungssysteme, in *David Hilbert, Gesammelte Abhandlungen*, vol. 3, Springer-Verlag, Berlin, 1979, pp. 94–145.
18. D. Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, *Göttingen Nachr.* (1906) 157–227.
19. ———, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*, Teubner, Leipzig, 1912.
20. ———, *Gesammelte Abhandlungen*, vol. 3, Springer-Verlag, Berlin, 1979.
21. A. Kufner, L. Maligranda, and L. E. Persson, The Hardy inequality–About its history and some related results, research report, Luleå University of Technology, Luleå, 2006.
22. A. Kufner and L. E. Persson, *Weighted Inequalities of Hardy Type*, World Scientific, Singapore, 2003.
23. E. Landau, Über einen Konvergenzsatz, *Göttingen Nachr.* (1907) 25–27.
24. ———, Letter to G. H. Hardy, June 21, 1921.
25. ———, Letter to G. H. Hardy, December 13, 1924.
26. ———, A note on a theorem concerning series of positive terms: Extract from a letter of Prof. E. Landau to Prof. I. Schur, *J. London Math. Soc.* **1** (1926) 38–39.
27. B. M. Makarov, M. G. Goluzina, A. A. Lodkin, and A. N. Podkorytov, *Selected Problems in Real Analysis*, Transl. of Math. Monographs, American Mathematical Society, Providence, 1992.
28. L. Maligranda, Why Hölder's inequality should be called Rogers' inequality, *Math. Inequal. Appl.* **1** (1998) 69–83.
29. B. Opic and A. Kufner, *Hardy-Type Inequalities*, Pitman Research Notes in Mathematics, no. 219, Longman Scientific & Technical, Harlow, 1990.
30. G. Pólya, Proof of an inequality, *Proc. London Math. Soc.* (2) **24** (1926) 57.
31. G. Pólya and G. Szegő, *Aufgaben und Lehrsätze aus der Analysis*, Springer-Verlag, Berlin, 1925.
32. F. Riesz, Untersuchungen über Systeme integrierbarer Funktionen, *Math. Ann.* **69** (1910) 449–497.
33. M. Riesz, Letter to G. H. Hardy, 1919 or 1920.
34. I. Schur, Letter to G. H. Hardy, 1918 or 1919.
35. ———, Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen, *J. Math.* **140** (1911) 1–28.
36. J. M. Steele, *The Cauchy-Schwarz Master Class. An Introduction to the Art of Mathematical Inequalities*, Cambridge University Press, Cambridge, 2004.
37. O. Toeplitz, Zur Theorie der quadratischen und bilinearen Formen von unendlich vielen Veränderlichen, *Göttingen Nachr.* (1910) 489–506.
38. F. Wiener, Elementarer Beweis eines Reihensatzes von Herrn Hilbert, *Math. Ann.* **68** (1910) 361–366.
39. H. Weyl, *Singuläre Integralgleichungen mit besonderer Berücksichtigung des Fourierschen Integraltheorems*, Ph.D. dissertation, Göttingen, 1908.

ALOIS KUFNER received his Ph. D. from the Czech Academy in 1965. He has worked at the Mathematical Institute of the former Czechoslovak (now Czech) Academy of Sciences, serving as deputy director (1974–

1978) and director (1978–1990) of that institute. Simultaneously, he taught at the Charles University and the Czech Technical University, both in Prague, and at the University of West Bohemia in Pilsen, where he was head of the Department of Mathematics (1977–1990). In 1983 he received the Doctor of Sciences degree at the Academy of Sciences and became a full professor at the University of West Bohemia in 1985. In 1988, he was elected corresponding member of the Czechoslovak Academy of Sciences. He was awarded the National Prize of the Czech Republic in 1985. His main fields of interests are differential equations and theory of function spaces. He is the author or coauthor of about 120 mathematical papers and eight monographs, among them, *Hardy-Type Inequalities* (Longman, 1990) and *Weighted Inequalities of Hardy Type* (World Scientific, 2003).

Mathematical Institute, Academy of Sciences of the Czech Republic, Žitná 25, 115 67 Praha 1, Czech Republic
kufner@math.cas.cz

LECH MALIGRANDA received his Ph. D. from the Adam Mickiewicz University in Poznan (Poland) in 1979 under the supervision of Władysław Orlicz. He has held appointments at the Polish Academy of Sciences in Poznan, University Central of Venezuela in Caracas, the Venezuelan Institute for Scientific Research (IVIC) in Caracas, and Claremont McKenna College. Currently he is in the Department of Mathematics at Luleå University of Technology (Sweden), where in 1991 he completed the Doctor of Sciences degree (= Docent in Sweden) and in 2000 became professor of mathematics. He received the Stefan Banach Prize of the Polish Mathematical Society in 1982, the State Prize from CONICIT in Venezuela in 1990, and the Adélie Prize for the Best Teacher at the Luleå University of Technology in 2001. His research interests include functional analysis, real analysis, inequalities, and sometimes the history of mathematics. He is the author or coauthor of one hundred mathematical papers and the monograph *Orlicz Spaces and Interpolation* (Campinas SP Brasil, 1989).

Department of Mathematics, Luleå University of Technology, SE-971 87 Luleå, Sweden
lech@sm.luth.se

LARS-ERIK PERSSON received his Ph. D. from the Umeå University (Sweden) in 1974 under the supervision of Ingemar Wik. He has held appointments at Umeå University, the Luleå University of Technology, Narvik University, and Uppsala University. Since 1994 he has been a full professor of mathematics at the Luleå University of Technology. He was president of the Swedish Mathematical Society (1996–1998), a member of the Swedish National Committe of Mathematics at the Royal Academy of Sciences from 1995, and director of the Center of Applied Mathematics at Luleå University of Technology from 1995. His research interests include the interpolation of operators, Fourier analysis, function spaces, inequalities, and homogenization theory. He is the author or coauthor of 120 mathematical papers and three monographs: *The Homogenization Method- An Introduction* (Studentlitteratur, Lund 1993), *Weighted Inequalities of Hardy Type* (World Scientific, 2003), and *Convex Functions and their Applications* (Springer, 2006).

Department of Mathematics, Luleå University of Technology, SE-971 87 Luleå, Sweden
larserik@sm.luth.se