The Development of the Hardy Inequality

Tyler Perlman

May 13, 2021

Abstract

This paper is a discussion of the development of the Hardy Inequality. We detail the inequality in both the discrete and continuous cases, as well as notable work, by Hardy and other mathematicians at the time, that contributed to its development. Much of the content draws upon an article from *The American Mathematical Monthly* by Alois Kufner, Lech Maligranda and Lars-Erik Persson [8].

1 Introduction

In this paper, we will discuss the Hardy inequality (in both the continuous and discrete cases), Hardy's motivation for his research that culminated in these results, and notable intermediate results by Hardy and his contemporaries. We will then present Hardy's proofs of his inequalities and conclude with further results and proofs by other mathematicians.
The Hardy Inequality can be stated as follows:

1. Discrete case: if $p > 1$ and $\{a_k\}_{k=1}^{\infty}$ is a sequence of nonnegative real numbers, then

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

2. Continuous case: if $p > 1$ and f is a nonnegative function on $(0, \infty)$ such that $\int_0^\infty |f(x)|^p dx < \infty$, then f is integrable over the interval $(0, x)$ for all $x > 0$ and

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

In both cases, $(\frac{p}{p-1})^p$ is a sharp constant: it cannot be replaced by a smaller number such that the inequality will remain true for all relevant sequences and functions.

1.1 Lebesgue Spaces

Before going on to discuss these inequalities, we will first briefly discuss the Lebesgue Spaces $\mathcal{L}^p$ and ℓ_p and related notation and terminology. We will be using the notation used in [1].
Given a measure space $(X, \mathcal{S}, \mu)$ and $0 < p \leq \infty$, the Lebesgue Space $\mathcal{L}^p$, or $\mathcal{L}^p(X, \mathcal{S}, \mu)$, is the set of $\mathcal{S}$-measurable functions $f : X \to \mathbb{F}$ such that $||f||_p < \infty$, where $||f||_p$ is the p-norm of f, and is defined as follows:

- if $0 < p < \infty$: $||f||_p = \left(\int |f|^p d\mu\right)^{\frac{1}{p}}$
- $||f||_\infty = \inf\{t > 0 : \mu(\{x \in X : |f(x)| > t\}) = 0\}$.

Moreover, when the measure μ is the counting measure on $\mathbb{Z}^+$, and $a = (a_1, a_2, ...)$ is a sequence in $\mathbb{F}$ and $0 < p < \infty$, then

$$||a||_p = \left(\sum_{n=1}^{\infty} |a_n|^p\right)^{\frac{1}{p}} \; ; \text{ and } ||a||_\infty = \sup\{|a_n| : k \in \mathbb{Z}^+\}$$

and we write ℓ_p in place of $\mathcal{L}^p(\mu)$.

If we restrict p to the interval $[1, \infty]$ and write that $\mathcal{Z}(\mu)$ be the set of $\mathcal{S}$-measurable functions that vanish except on a set of order zero, and denote the quotient space $\mathcal{L}^p(\mu)/\mathcal{Z}(\mu)$ as $L_p(\mu)$, we construct a Banach space of p-integrable functions. In L_p we simply say that the p-norm of a function f, $||f||_p$, is equal to the norm of its representative in $\mathcal{L}^p$. In the case of the counting measure, $\mathcal{L}^p(\mu) = L_p(\mu) = \ell_p(\mu)$ as, with respect to this measure, the only set of measure zero is the empty set.

2 Motivation and Prior Results

Before detailing Hardy's proofs of the main results, we will address Hardy's motivation for beginning work toward these results, important theorems that will be of use to a reader of this paper, and Hardy's related results prior to his proof of his continuous and discrete inequalities.

2.1 Motivation: The Hilbert Inequality

In the article by Kufner, Maligranda, and Persson [8], they write that "it seems completely clear that Hardy's original motivation when he began the research that culminated in his discovery of the inequalities [(1) and (2)] was to prove (the weak form of) the Hilbert inequality." This inequality (and variants of it) pertains to sequences $\{a_m\}_{m\geq 1}$ and $\{b_n\}_{m\geq 1}$ of nonnegative real numbers such that $\sum_{m=1}^{\infty} a_m^2 < \infty$ and $\sum_{n=1}^{\infty} b_n^2 < \infty$, or in the notation we introduced earlier: $a, b \in \ell_2(\mu)$. In this context, we have the following variants of the Hilbert Inequality:

- Weak form: the double series

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{a_m b_n}{m+n} \text{ converges} \tag{3}$$

- Typical (Strong) form: the inequality

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{a_m b_n}{m+n} \leq \pi \left(\sum_{m=1}^{\infty} a_m^2\right)^{\frac{1}{2}} \left(\sum_{n=1}^{\infty} b_n^2\right)^{\frac{1}{2}} \tag{4}$$

  holds, and π is a sharp constant.

- Generalization to ℓ_p: for $p > 1$

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{a_m b_n}{m+n} \leq \frac{\pi}{\sin\left(\frac{\pi}{p}\right)} \left(\sum_{m=1}^{\infty} a_m^p\right)^{\frac{1}{p}} \left(\sum_{n=1}^{\infty} b_n^{p'}\right)^{\frac{1}{p'}} \tag{5}$$

These aforementioned inequalities, in the special case where our sequences a and b are the same sequence, imply the weaker inequality that was of interest to Hardy: if for a nonnegative sequence $\{a_m\}_{m\geq 1}$ such that $a \in \ell_2(\mu)$, then

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{a_m a_n}{m+n} \text{ converges.} \tag{6}$$

2.2 Useful Theorems and Inequalities

In the discussion and proofs that follow, there will applications of several inequalities. These inequalities are included here for ease of reading.

2.2.1 Dual Exponent

For $1 \leq p \leq \infty$, the dual exponent of p is $p' \in [1, \infty]$ such that

$$\frac{1}{p} + \frac{1}{p'} = 1$$

2.2.2 Young's Inequality

If $1 < p < \infty$, then $\forall a, b \geq 0$

$$ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'} \tag{7}$$

2.2.3 Hölder's Inequality

If $(X, \mathcal{S}, \mu)$ is a measure space, $1 \leq p \leq \infty$, and $f, h : X \to \mathbb{F}$ are $\mathcal{S}$-measurable, then

$$||fh||_1 \leq ||f||_p ||h||_{p'} \tag{8}$$

2.2.4 Minkowski's Inequality

If $(X, \mathcal{S}, \mu)$ is a measure space, $1 \leq p \leq \infty$, and $f, h \in \mathcal{L}^p(\mu)$ then

$$||f + h||_p \leq ||f||_p + ||h||_p \tag{9}$$

2.3 Intermediate Results

In the course of proving the final form of each the discrete and continuous cases of the Hardy Inequality, Hardy and his contemporaries proved various related results. Focusing primarily on the contributions from Hardy's 1915 [4], 1919 [5], and 1920 [6] articles, relevant results will be presented here.

2.3.1 1915 Article

Theorem 2.1. *let $a > 0$, $f(x)$ be non-negative and integrable on (a, ∞), and denote $F(x) = \int_a^x f(t)dt$. The convergence of the any of the integrals*

$$\int_a^\infty \frac{f(x)F(x)}{x} dx, \quad \int_a^\infty \left(\frac{F(x)}{x}\right)^2 dx, \quad \int_a^\infty \int_a^\infty \frac{f(x)f(y)}{x+y} dxdy$$

implies that of the others.

Theorem 2.2. *The convergence of $\int_a^\infty f(x)^2 dx$ for f non-negative implies that of*

$$\int_a^\infty \left(\frac{F(x)}{x}\right)^2 dx$$

Corollary 2.2.1. *The convergence of $\int_a^\infty f(x)^2 dx$ for f non-negative implies that of the integrals in Theorem 2.1.*

Theorem 2.3. *The convergence of $\sum a_n^2$ when $a_n \geq 0$ implies that of*

$$\sum_{n=0}^\infty \left(\frac{A_n}{n}\right)^2$$

where $A_n = \sum_{k=1}^n a_k$.

Theorem 2.4. *let $a > 0$, $f(x), g(x)$ be non-negative and integrable on (a, ∞), and denote $F(x) = \int_a^x f(t)dt$, $G(x) = \int_a^x g(t)dt$ The following hypotheses are equivalent:*

$$\begin{aligned}
&(1) \int_a^\infty \frac{f(x)G(x)}{x}dx \text{ and } \int_a^\infty \frac{F(x)g(x)}{x}dx \text{ are convergent,}\\
&(2) \int_a^\infty \frac{F(x)G(x)}{x^2}dx \text{ is convergent,}\\
&(3) \int_a^\infty \frac{f(x)g(y)}{x+y}dxdy \text{ is convergent.}
\end{aligned}$$

Corollary 2.4.1. *The convergence of $\int_a^\infty f(x)^2dx$ and $\int_a^\infty g(x)^2dx$ implies that of the integrals in Theorem 2.4.*

2.3.2 1919 Article

This article included a proof of the continuous case of the Hardy Inequality for $p = 2$, with the sharp constant 4, although there was no mention of it being a sharp constant.

Theorem 2.5. *When $a_n \geq 0$ and $\int_a^\infty f(x)^2dx$ is convergent,*

$$\int_a^\infty \left(\frac{F(x)}{x}\right)^2 dx = \int_a^\infty \left(\frac{1}{x}\int_a^x f(t)dt\right)^2 dx \leq 4\int_0^\infty f(x)^2dx.$$

2.3.3 1920 Article

The most relevant contributions from this article are Hardy's presentation of a proof by Frigyes Riesz of the Hardy Inequality in the discrete case with a weaker constant, as well as a subsequent improvement to said constant.

Theorem 2.6 (Proved by Riesz)**.** *If $p > 1$, $a_n \geq 0$, and $\sum a_n^x$ is convergent, then*

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

Theorem 2.7. *If $p > 1$, $a_n \geq 0$, and $\sum a_n^x$ is convergent, then*

$$\sum_{n=1}^N \left(\frac{1}{n}\sum_{k=1}^n a_k\right)^p \leq (p\ \zeta(p))^p \sum_{n=1}^N a_n^p$$

where $\zeta(p) = \sum_{n=1}^\infty n^{-p}$.

2.3.4 Lead up to the 1925 Article

In the 1925 article, Hardy first published the inequalities of interest, with the sharp constants. Prior to this paper, however, in a 1921 letter to Hardy from Edmund Landau, Landau provided a proof of the discrete inequality in the following form [8]:

Theorem 2.8 (Proved by Landau)**.** *If $p > 1$, $a_n \geq 0$, $\sum a_n^x$ is convergent, then*

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

and the constant $\left(\frac{p}{p-1}\right)^p$ is sharp when $N = \infty$.

Additionally, in another letter, Landau drew Hardy's attention to the fact that, in his 1920 paper, Hardy had remarked that $(p/(p-1))^p$ was the best constant in the continuous case, without providing a proof. This exchange was addressed in Hardy's 1925 paper alongside such a proof.

3 Proof of the Hardy Inequality

We will now present Hardy's proofs of (1) and (2) as they were originally presented in his 1925 paper [7], albeit with slight modification for notational simplicity.

3.1 Continuous Case

Theorem 3.1. *Suppose that $f(x) \geq 0$, $p > 1$, that f Lebesgue integrable over any finite interval $(0, X)$, and*

$$F(x) = \int_0^x f(t)dt;$$

and that $f \in \mathcal{L}^p(\mathbb{R}^+)$. Then

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

Proof. By applying integration by parts and making use of the chain rule $\left(\frac{d}{dx}F(x)^p = pF(x)^{p-1}f(x)\right)$:

$$\begin{aligned}
&\int_\epsilon^X \left(\frac{F(x)}{x} \right)^p dx = -\frac{1}{p-1} \int_\epsilon^X F(x)^p \frac{d}{dx}(x^{1-p})dx \\
&= \frac{\epsilon^{1-p}}{p-1} F(\epsilon)^p - \frac{X^{1-p}}{p-1} F(X)^p + \frac{1}{p-1} \int_\epsilon^X x^{1-p} \frac{d}{dx}(F(x)^p)dx \\
&= \frac{\epsilon^{1-p}}{p-1} F(\epsilon)^p - \frac{X^{1-p}}{p-1} F(X)^p + \frac{p}{p-1} \int_\epsilon^X x^{1-p} F(x)^{p-1} f(x)dx \\
&\leq \frac{\epsilon^{1-p}}{p-1} F(\epsilon)^p + \frac{p}{p-1} \int_\epsilon^X x^{1-p} F(x)^{p-1} f(x)dx.
\end{aligned}$$

When $\epsilon \to 0$, by Hölder's Inequality (8), we have that

$$F(\epsilon)^p = \left(\int_0^\epsilon f(t)dt\right)^p \leq \int_0^\epsilon f(t)^p dt \left(\int_0^\epsilon dt\right)^{p-1} = o(\epsilon^{p-1}).$$

For $\delta > 0, \exists \epsilon_\delta$ so that $\forall \epsilon \leq \epsilon_\delta$:

$$F(\epsilon)^p < (p-1)\delta ||f||_p^p \epsilon^{p-1}.$$

Also,

$$\begin{aligned}\int_\epsilon^X \frac{F(x)}{x}^{p-1} f(x)dx &\leq \left(\int_\epsilon^X f(x)^p dx\right)^{\frac{1}{p}} \left(\int_\epsilon^X \frac{F(x)}{x}^p dx\right)^{\frac{1}{p'}} \\ &\leq ||f||_p^p \left(\int_\epsilon^X \frac{F(x)}{x}^p dx\right)^{\frac{1}{p'}}\end{aligned}$$

Denoting $W = \left(\int_\epsilon^X \frac{F(x)}{x}^p dx\right)$, this yields that:

$$W < \delta ||f||_p^p + \frac{p}{p-1} ||f||_p W^{1/p'}$$

or, equivalently,

$$\left(\frac{W}{||f||_p^p}\right)^p - \frac{p}{p-1}\left(\frac{W}{||f||_p^p}\right)^{p-1} - \delta < 0.$$

As the equation $ay^p + by^{p-1} - 1 = 0$ has just one positive root, so too does this previous equation. If we denote that root by η, it is clear that $\frac{W}{||f||_p^p} < \eta$. On the other hand, however, as $\frac{p}{p-1} > 1$

$$\left(\frac{p}{p-1} + \delta\right)^p - \frac{p}{p-1}\left(\frac{p}{p-1} + \delta\right)^{p-1} - \delta = \delta\left(\left(\frac{p}{p-1} + \delta\right)^{p-1} - 1\right) > 0.$$

As such, $\eta < \frac{p}{p-1} + \delta$, $\frac{W}{||f||_p^p} < \frac{p}{p-1} + \delta$, and

$$W = \int_\epsilon^X \left(\frac{F(x)}{x}\right)^p dx < ||f||_p^p \left(\frac{p}{p-1} + \delta\right)^p.$$

As this will hold $\forall \epsilon \leq \epsilon_\delta$ and all $X > 0$, we have that

$$\int_0^\infty \left(\frac{F(x)}{x}\right)^p dx \leq \left(\frac{p}{p-1} + \delta\right)^p \int_0^\infty f(x)^p dx,$$

which is equivalent to our claim, as δ is arbitrary.

□

Later in the paper, Hardy proves that $(p/(p-1))^p$ is the best possible constant:

Theorem 3.2. *The inequality (2) is strict ($(p/(p-1))^p$ is the best possible constant).*

Proof. To show that this is the best possible constant, we proceed as follows:

$$\text{let } \delta > 0, \quad f = \begin{cases} 0 & 0 \le x < 1 \\ x^{-(\lambda+\epsilon)} & 1 < x \end{cases},$$

where $\lambda = 1/p, \quad 0 < \epsilon < \frac{1}{2}(1-\lambda) < 1-\lambda$; and choose $X = X(p,\delta)$ so that

$$X^{-1} < \delta X^{-\frac{1}{2}(1+\lambda)} < \delta X^{-}(\lambda+\epsilon).$$

This gives us that

$$\int_0^\infty f(x)^p dx = \int_1^\infty x^{-(1+p\epsilon)} dx = \frac{1}{p\epsilon}$$

$$\frac{F(x)}{x} = \frac{1}{x}\int_1^x t^{-(\lambda+\epsilon)} = \frac{1}{x}\frac{x^{1-(\lambda+\epsilon)}-1}{1-(\lambda+\epsilon)} \quad (\text{for } x > 1),$$

$$\int_0^\infty \left(\frac{F(x)}{x}\right)^p dx = \left(\frac{p}{p-1-p\epsilon}\right)^p \int_1^\infty (x^{-(\lambda+\epsilon)} - x^{-1})^p dx$$

$$> \left(\frac{p}{p-1}\right)^p \int_X^\infty x^{-(1+p\epsilon)} dx$$

$$= \left(\frac{p}{p-1-p\epsilon}\right)^p (1-\delta)^p \frac{X^{-p\epsilon}}{p\epsilon} = \left(\frac{p}{p-1-p\epsilon}\right)^p (1-\delta)^p X^{-p\epsilon} \int_0^\infty f(x)^p dx.$$

As delta is arbitrary, we can make $(1-\delta)^p$ as close to 1 as we like. Similarly, we can make $X^{-p\epsilon}$ as close to 1 as we like through our choice of ϵ. By doing so, we get the following result for this choice of $f(x)$:

$$\int_0^\infty \left(\frac{F(x)}{x}\right)^p dx > \left(\frac{p}{p-1}\right)^p \int_0^\infty f(x)^p dx$$

As such, for this choice of $f(x)$ we have equality. Thus, there can be no smaller constant in (2) that will still hold for all functions of interest.

$\square$

3.2 Discrete Case

In his 1925 paper, Hardy utilizes the continuous inequality to prove a more general version of the discrete inequality. We will present this proof here:

Theorem 3.3. *Suppose that $a_n > 0$ and $\lambda > 0$ $\forall n > 0$, that $A_n = \lambda_1 a_1 + ... + \lambda_n a_n$, that $\Lambda_n = \lambda_1 + ... + \lambda_n$, and that $\sum \lambda_n a_n^p$ is convergent. Then*

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

Proof. Denote $b_n = A_n/\Lambda_n$. First, note that we need only prove the claim for a sum from $n = 1$ up to an arbitrary, finite N, as our claim will then follow by taking the limit as $n \to \infty$. What's more, we will simply look at the special case of a monotonically decreasing sequence $(a_1 \geq ... \geq a_N)$. To justify this, we claim that, as $\sum \lambda_n a_n^p$ is convergent, $\sum \lambda_n b_n^p$ will attain a maximal value for some sub-sequence of $\{a_n\}_{n\geq 1}$.

To see this, suppose that $a_\mu < a_\mu$ for some $\mu < \nu$, and define α by

$$\alpha^p = \frac{\lambda_\mu a_\mu^p + \lambda_\nu a_\nu^p}{\lambda_\mu + \lambda_\nu}$$

so that $a_\mu < \alpha < a_\nu$. We then have that

$$\lambda_\mu a_\mu + \lambda_\nu a_\nu \leq (\lambda_\mu + \lambda_\nu)^{1/p'}(\lambda_\mu a_\mu^p + \lambda_\nu a_\nu^p)^{1/p} = (\lambda_\mu + \lambda_\nu)\alpha.$$

If A_n became A_n' when α is substituted for both a_μ and a_ν, then when $n < \mu$ $A_n = A_n'$, when $\mu \leq n < \nu$ $A_n < A_n'$ because a_μ is replaced by α, and $A_n \leq A_n'$ when $\nu < n$ because $(\lambda_\mu a_\mu + \lambda_\nu a_\nu)$ will be replaced by $(\lambda_\mu + \lambda_\nu)\alpha$. This replacement would thus increase $\sum \lambda_n b_n^p$. As such, the sub-sequence was not the maximal one. Thus, we need only work with monotonically decreasing sequences.

Next, we note that if $\{a_n\}_{n>0}$ is monotonically decreasing, then $\{b_n\}_{n>0} = \{A_n/\Lambda_n\}_{n>0}$ will also be monotonically decreasing. The condition for that is $b_{n-1} \geq b_n$ or $A_{n-1}\Lambda_n \geq A_n\Lambda_{n-1}$. Expanded, this can be written as the condition:

$$(\lambda_1 + ... + \lambda_n)(\lambda_1 a_1 + ... + \lambda_{n-1}a_{n-1}) \geq (\lambda_1 + ... + \lambda_{n-1})(\lambda_1 a_1 + ... + \lambda_n a_n)$$

Removing common factors yields:

$$\begin{aligned}\lambda_n(\lambda_1 a_1 + ... + \lambda_{n-1}a_{n-1}) &\geq (\lambda_1 + ... + \lambda_{n-1})\lambda_n a_n \\ \lambda_1 a_1 + ... + \lambda_{n-1}a_{n-1} &\geq \lambda_1 a_n + ... + \lambda_{n-1}a_n,\end{aligned}$$

which is clearly true as $a_1 \geq a_n$, ..., $a_{n-1} \geq a_n$.

Finally, we will now proceed by utilizing Theorem 3.1 and applying it to a carefully chosen step function:

$$f(x) = \begin{cases} a_1 & 0 \leq x < \Lambda_1 \\ a_n & \Lambda_{n-1} \leq x \leq \Lambda_n, \quad n = 2, 3, ..., \text{N} \\ 0 & \Lambda_N \leq x \end{cases}$$

For this function,

$$\int_0^\infty f(x)^p dx = \sum_{n=1}^\infty (\Lambda_n - \Lambda_{n-1})a_n^p = \sum_{n=1}^\infty \lambda_n a_n^p$$

If $\Lambda_{n-1} \leq x \leq \Lambda_n$, then

$$\frac{F(x)}{x} = \frac{\lambda_1 a_1 + ... + \lambda_{n-1}a_{n-1} + (x - \Lambda_{n-1})a_n}{x},$$

which will decrease monotonically from A_{n-1}/Λ_{n-1} to A_n/Λ_n as x increases from Λ_{n-1} to Λ_n. Thus

$$\frac{F(x)}{x} \geq \frac{A_n}{\Lambda_n} \quad \text{when } \Lambda_{n-1} \leq x < \Lambda_n.$$

Combining these observations with Theorem 3.1 and the earlier simplifications that we made yields the desired result:

$$\sum_{n=1}^{N} \lambda_n \left(\frac{A_n}{\Lambda_n}\right)^p \leq \int_0^\infty \left(\frac{F(x)}{x}\right)^p \leq \left(\frac{p}{p-1}\right)^p \int_0^\infty f(x)^p dx = \left(\frac{p}{p-1}\right) \sum_{n=1}^{N} \lambda_n a_n^p$$

□

4 Further Results

Having proven our main result, we will now proceed to explore alternative proofs and further results.

4.1 Pólya's Proof

In that very same article, Hardy also shared an alternative proof of Theorem 3.1 which had been pointed out to him by George Pólya in their correspondence. This proof, while beginning in the same way, makes a few notable simplifications and thus avoids some of the more technical arguments in Hardy's proof. We present this alternative proof here:

Proof. Suppose that $0 < \alpha < \beta < X$. Recall, from Hardy's proof, that:

$$\int_\epsilon^X \left(\frac{F(x)}{x}\right)^p dx \leq \frac{\epsilon^{1-p}}{p-1} F(\epsilon)^p + \frac{p}{p-1} \int_\epsilon^X x^{1-p} F(x)^{p-1} f(x) dx$$

By taking this inequality and replacing $F(x)$ with $F(x) - F(\alpha)$ and ϵ with α, dropping the first term (which is non-negative), and applying Hölder's inequality (8) in the same way we do in Hardy's proof, this yields:

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

and

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

As $f(x)$ is non-negative, $F(x)$ will be monotonically increasing. Thus, $F(x) - F(\alpha)$ increases monotonically to $F(x)$ as $\alpha \to 0$. Hence:

$$\int_\beta^X \left(\frac{F(x)}{x}\right)^p dx \le \left(\frac{p}{p-1}\right)^p \int_0^\infty f(x)^p dx,$$

which proves the theorem, as β and X are arbitrary and we can apply this inequality for $\beta \to 0$ and $X \to \infty$. $\square$

4.2 The Hardy Operator

One consequence of Hardy's Inequality, is 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 and L_p ($p > 1$) into themselves, respectively. Moreover, each of these operators have norm $p' = \frac{p}{p-1}$.

Proof. First, we recall that L_p and l_p are Banach spaces. The norm of a linear map, T between Banach spaces V and W is given by $||T|| = \sup\{||Tf||_W : f \in V \text{ and } ||f||_V \le 1\}$. Rephrasing Hardy's inequality in terms of the l_p and L_p norms and the continuous and discrete Hardy operators yields:

$$||h(\{a_n\})||_{l_p}^p \le \left(\frac{p}{p-1}\right)^p ||a||_{l_p}^p$$
$$||Hf(x)||_{L_p}^p \le \left(\frac{p}{p-1}\right)^p ||f(x)||_{L_p}^p.$$

By exponentiation of both equations by $1/p$, we arrive at:

$$||h(\{a_n\})||_{l_p} \le \frac{p}{p-1}||a||_{l_p}$$
$$||Hf(x)||_{L_p} \le \frac{p}{p-1}||f(x)||_{L_p},$$

where the constant $p' = \frac{p}{p-1}$ is the best constant (for any sequence or function). Paying special attention to the case of functions in L_p whose norms are less than or equal to 1, by the definition of the norm of a linear map and these inequalities, we immediately arrive at the desired result. $\square$

4.3 Ingham's Proof of The Hardy Inequality

By making use of the Hardy operator and Minkowski's Inequality (9), Albert Ingham was able to provide the following, much simpler, proof [8]:

Proof. By a simple change of variables, we can write that

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

By applying Minkowski's inequality, it follows that

$$\begin{aligned}\left(\int_0^\infty (Hf(x)))^p dx\right)^{1/p} &= ||Hf(x)||_p = \left|\left|\int_0^1 f(tx)dt\right|\right|_p \\ &\leq \int_0^1 ||f(tx)||_p dt = \int_0^1 \left(\int_0^\infty f(tx)^p dx\right)^{1/p} dt \\ &= \int_0^1 \left(\int_0^\infty f(s)^p \frac{ds}{t}\right)^{1/p} dt = \frac{p}{p-1}\left(\int_0^\infty f(s)^p ds\right)^{1/p}\end{aligned}$$

By exponentiation of both sides of this equation, we arrive at the desired result:

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

□

4.4 Carleman's Inequality

The 1925 paper [7] also presented the following inequality, which is derived by taking a limit of the discrete case of Hardy's inequality. This result was pointed out to Hardy by Pólya in one of the many letters that they had sent to one another, and can be stated as:

Theorem 4.1. *Suppose that $b_n > 0$ and $\lambda > 0$ $\forall n > 0$, that $\Lambda_n = \lambda_1 + ... + \lambda_n$, and that $\sum \lambda_n b_n$ is convergent. Then $\forall N$*

$$\sum_{n=1}^N \lambda_n (b_1^{\lambda_1} ... b_n^{\lambda_n})^{1/\Lambda_n} \leq e \sum_{n=1}^N \lambda_n b_n,$$

and e is the best possible constant for this inequality.

Proof. In the inequality in Theorem 3.3,

$$\sum_{n=1}^N a_n \left(\frac{\lambda_1 a_1 + ... + \lambda_n a_n}{\lambda_1 + ... + \lambda_n}\right)^p \leq \left(\frac{p}{p-1}\right)^p \sum_{n=1}^N \lambda_n a_n^p$$

we proceed by denoting a_n^p as b_n, and take the limit as $p \to \infty$. The right hand side becomes

$$\lim_{p\longrightarrow\infty}\left(\frac{p}{p-1}\right)^p\sum_{n=1}^{N}\lambda_n b_n = \lim_{p\longrightarrow\infty}\left(1+\frac{1}{p-1}\right)^p\sum_{n=1}^{N}\lambda_n b_n = e\sum_{n=1}^{N}\lambda_n b_n,$$

and on the left hand side, we will address each term of the sum,

$$\lim_{p\longrightarrow\infty}\left(\frac{\lambda_1 b_1^{1/p}+...+\lambda_n b_n^{1/p}}{\lambda_1+...+\lambda_n}\right)^p,$$

separately. For each such term, we can rewrite it as

$$\lim_{p\longrightarrow\infty}\left(\sum_{i=1}^{n} w_i b_i^{1/p}\right)^p,$$

where $w_i = \lambda_i/\Lambda_n$ and $\sum_{i=1}^n w_i = 1$. Letting $k = 1/p$, we can proceed as follows, making use of the fact that exp is a continuous function on $\mathbb{R}$:

$$\begin{aligned}\lim_{p\longrightarrow\infty}\left(\sum_{i=1}^{n} w_i b_i^{1/p}\right)^p &= \lim_{k\longrightarrow 0}\left(\sum_{i=1}^{n} w_i b_i^{k}\right)^{1/k} = \lim_{k\longrightarrow 0}\exp\left\{\log\left(\left(\sum_{i=1}^{n} w_i b_i^{k}\right)^{1/k}\right)\right\}\\ &= \lim_{k\longrightarrow 0}\exp\left\{\frac{\log\left(\sum_{i=1}^{n} w_i b_i^{k}\right)}{k}\right\} = \exp\left\{\lim_{k\longrightarrow 0}\frac{\log\left(\sum_{i=1}^{n} w_i b_i^{k}\right)}{k}\right\}.\end{aligned}$$

We will next make use of L'Hôpital's rule for limits. This is justified, as

$$\lim_{k\longrightarrow 0}\log\left(\sum_{i=1}^{n} w_i b_i^{k}\right) = \log\left(\sum_{i=1}^{n} w_i\right) = \log(1) = 0 \quad \text{and} \quad \lim_{k\longrightarrow 0} k = 0.$$

As such,

$$\begin{aligned}\lim_{k\longrightarrow 0}\frac{\log\left(\sum_{i=1}^{n} w_i b_i^{k}\right)}{k} &= \lim_{k\longrightarrow 0}\frac{\frac{d}{dk}\log\left(\sum_{i=1}^{n} w_i b_i^{k}\right)}{\frac{d}{dk}k} = \lim_{k\longrightarrow 0}\frac{\sum_{i=1}^{n} w_i b_i^{k}\log(b_i)}{\sum_{j=1}^{n} w_j b_j^{k}}\\ &= \frac{\sum_{i=1}^{n} w_i\log(b_i)}{\sum_{j=1}^{n} w_j} = \sum_{i=1}^{n} w_i\log(b_i).\end{aligned}$$

Utilizing this, our limit becomes,

$$\lim_{p\longrightarrow\infty}\left(\sum_{i=1}^{n} w_i b_i^{1/p}\right)^p = \exp\left\{\sum_{i=1}^{n} w_i\log(b_i)\right\} = \exp\left\{\log\left(\prod_{i=1}^{n} b_i^{w_i}\right)\right\} = \prod_{i=1}^{n} b_i^{w_i}.$$

Returning to our original notation, this can be written as:

$$\lim_{p\to\infty}\left(\frac{\lambda_1 b_1^{1/p}+...+\lambda_n b_n^{1/p}}{\lambda_1+...+\lambda_n}\right)^p = (b_1^{\lambda_1}...b_n^{\lambda_n})^{1/\Lambda_n}.$$

Combined with our previous result when taking the limit of the right hand side of Hardy's inequality, we arrive at our desired result:

$$\sum_{n=1}^{N}\lambda_n(b_1^{\lambda_1}...b_n^{\lambda_n})^{1/\Lambda_n} \le e\sum_{n=1}^{N}\lambda_n b_n.$$

□

By applying this result to the special case where $\forall n\ \lambda_n = 1$, and thus $\Lambda_n = n$, we can arrive at the following result:

Theorem 4.2 (Carleman's Inequality)**.** *Suppose that $\{a_n\}_{n>0}$ is a sequence of non-negative real numbers. Then*

$$\sum_{n=1}^{N}(a_1...a_n)^{1/n} \le e\sum_{n=1}^{N}a_n.$$

This result was originally proven by Torsten Carleman in 1923.

4.5 The Copson Inequality

In a 1927 article in the *Journal of the London Mathematical Society* [2], E. T. Copson presented an inequality, which I will refer to as "The Copson Inequality". The following year, Hardy explained that this inequality and his own (the discrete case) were "...'reciprocal' in the sense that either can be deduced from the other..." [3]. This inequality can be stated as:

Theorem 4.3 (Copson's Inequality)**.** *Suppose $p > 1$, $\lambda_n > 0$, $a_n > 0$ for $n = 1, 2, ...$, that $\Lambda_n = \lambda_1 + ... + \lambda_n$, and that $\sum_{n=1}^{\infty}\lambda_n a_n^p$ converges. Then*

$$\sum_{n=1}^{\infty}\lambda_n\left(\sum_{k=n}^{\infty}\frac{\lambda_k a_k}{\Lambda_k}\right)^p \le p^p\sum_{n=1}^{\infty}\lambda_n a_n^p.$$

Hardy's observation that the two inequalities are "reciprocal", or dual, can be phrased as follows:

Theorem 4.4. *Suppose $p > 1$, that $\lambda_n > 0$, $a_n \ge 0$, $\Lambda_n = \lambda_1 + ... + \lambda_n$, $A_n = \lambda_1 a_1 + ... + \lambda_n a_n$, and $A_n^\star = \frac{\lambda_n a_n}{\Lambda_n} + \frac{\lambda_{n+1}a_{n+1}}{\Lambda_{n+1}} + ...$, for $n \ge 1$. Then:*

$$\sum_{n=1}^{\infty}\lambda_n\left(\frac{A_n}{\Lambda_n}\right)^p \le \left(\frac{p}{p-1}\right)^p\sum_{n=1}^{\infty}\lambda_n a_n^p$$
$$\sum_{n=1}^{\infty}\lambda_n\ {A_n^\star}^p \le p^p\sum_{n=1}^{\infty}\lambda_n a_n^p,$$

and the constants $(\frac{p}{p-1})^p$ *and* p^p *are the best such constants in each case respectively. Moreover, each of these inequalities may be deduced from the other.*

Acknowledgements

I would like to thank Professor Dan-Andrei Geba for advising me in writing this honors thesis. I would also like to thank Professor Naomi Jochnowitz and Professor Amanda Tucker for being on the thesis committee and Professor Jonathan Pakianathan for coordinating the honors thesis program.

References

[1] Sheldon Axler. *Measure, Integration & Real Analysis.* Springer International Publishing, 2020.

[2] E. T. Copson. Note on series of positive terms. *Journal of the London Mathematical Society*, s1-2(1):9–12, 1927.

[3] G. H. Hardy. Remarks on three recent notes in the journal. *Journal of the London Mathematical Society*, s1-3(3):166–169, 1928.

[4] Godfrey Harold Hardy. Notes on some points in the integral calculus xli: On the convergence of certain integrals and series. *Messenger of Mathematics*, 44:163–166, 1915.

[5] Godfrey Harold Hardy. 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 Mathematics*, 48:107–112, 1919.

[6] Godfrey Harold Hardy. Note on a 9 theorem of hilbert. *Mathematische Zeitschrift*, 6:314–317, 1920.

[7] Godfrey Harold Hardy. Notes on some points in the integral calculus lx: An inequality between integrals. *Messenger of Mathematics*, 54:150–156, 1925.

[8] Alois Kufner, Lech Maligranda, and Lars-Erik Persson. The prehistory of the hardy inequality. *The American Mathematical Monthly*, 113(8):715–732, 2006.