ON HILBERT'S INEQUALITY IN \( n \) DIMENSIONS

BY N. G. DE BRUIJN AND HERBERT S. WILF

Communicated by I. J. Schoenberg, November 25, 1961

The celebrated inequality of Hilbert asserts that if \( a_1, \ldots, a_n \) are real and not all zero then

\[
\sum_{\mu, \nu = 1}^{n} \frac{a_\mu a_\nu}{\mu + \nu} < \pi \sum_{\nu = 1}^{n} \frac{a_\nu^2}{\nu}.
\]

Furthermore, if the constant \( \pi \) were replaced by a smaller number, then the inequality would be violated for some \( n, a_1, \ldots, a_n \). On the other hand, let us regard \( n \) as fixed, and let \( \lambda_n \) denote the best possible constant for the inequality

\[
\sum_{\mu, \nu = 1}^{n} \frac{a_\mu a_\nu}{\mu + \nu} \leq \lambda_n \sum_{\nu = 1}^{n} \frac{a_\nu^2}{\nu}.
\]

Naturally \( \lambda_n \) is just the largest eigenvalue of the section

\[
\frac{1}{\mu + \nu} \sum_{\mu, \nu = 1}^{n}
\]

of Hilbert's matrix, and we must have

\[
\lim_{n \to \infty} \left( \pi - \lambda_n \right) = \mathcal{O}(1)
\]

The question of obtaining more precise information about the term \( \mathcal{O}(1) \) in (3) seems to have been first publicly raised by W. W. Sawyer [1]. Because of the interest of this problem in numerical analysis [2] it has been investigated by several workers [3; 4; 5] with the result that various upper and lower bounds for the rate of growth of this term are known. We have been able to determine the desired rate of growth more satisfactorily by finding the first two terms of the asymptotic expansion of \( \lambda_n \). Our result is the following

\[1\] The work of the second author was supported in part by the U.S.A.E.C. under contract with the Argonne National Laboratory.
ON HILBERT'S INEQUALITY IN n DIMENSIONS

THEOREM. The best possible constant for the inequality (2) is of the form

\[ \lambda_n = \pi - \frac{1}{2} \pi^4 (\log n)^{-2} + O(\log \log n (\log n)^{-1}) \quad (n \to \infty). \]

Our method consists in relating \( \lambda_n \) to the eigenvalue \( \Lambda_n \) of the corresponding integral equation

\[ \Lambda_n f(u) = \int \frac{f(v)}{u + v} dv, \]

which seems considerably more tractable.

Indeed let us put, in (5), \( f(u) = u^{-1/2} g(u) \), \( n = e^{2A} \), \( u = e^{2A} \), \( v = e^{2A} \) and \( g(A) = \phi(x) \). Then (5) becomes

\[ \Lambda_n \phi(x) = \int_A^{\infty} K(x - y) \phi(y) dy, \]

where

\[ K(t) = \left( 2 \cosh \frac{t}{2} \right)^{-1}. \]

We will now sketch two proofs of the fact that

\[ \lambda_n = \pi - \frac{1}{2} \pi^4 (\log n)^{-2} + O(\log \log n (\log n)^{-1}) \quad (n \to \infty). \]

The first of these consists merely in observing that problems of the form (6) have been treated by Widom [6; 7] in great generality. Thus, after verifying that the Fourier transform of \( K(t) \) satisfies the hypotheses in [6] we get (8) at once from Theorem 3.1 of [6].

For a direct proof of (8), define

\[ \phi^{(1)}(x) = \cos \frac{\pi x}{2A}, \quad \Lambda_n^{(1)} = \pi \sech \frac{\pi}{2A}, \]

\[ \phi^{(2)}(x) = \cos \frac{\pi x}{2(A + \rho)}, \quad \Lambda_n^{(2)} = \pi \sech \frac{\pi}{2(A + \rho)}. \]

Then \( \phi^{(1)}(x) \) and \( \phi^{(2)}(x) \) are eigenfunctions of

\[ \int_{-\infty}^{\infty} K(x - y) \phi(y) dy = \Lambda \phi(x) \]

with eigenvalues \( \Lambda_n^{(1)} \), \( \Lambda_n^{(2)} \) respectively. Now we have

\[ \int_A^{\infty} K(x - y) \phi^{(1)}(y) dy < 0 \quad (-A \leq x \leq A), \]

and similarly for \( I_A^{(2)} \). Further, it is not difficult to show that we can fix a positive \( \rho \) such that for all sufficiently large \( A \) we have

\[ \int_{-A}^{A} K(x - y) \phi^{(2)}(y) dy > 0 \quad (-A \leq x \leq A), \]

and similarly for \( I_A^{(2)} \). Therefore

\[ \int_{-A}^{A} K(x - y) \phi^{(2)}(y) dy \gtrless \Lambda_n^{(1)} \phi^{(1)}(x), \]

\[ \int_{-A}^{A} K(x - y) \phi^{(2)}(y) dy \lbit \Lambda_n^{(2)} \phi^{(2)}(x), \]

and since \( K, \phi^{(1)} \) and \( \phi^{(2)} \) are nonnegative it can be proved in a well-known way that \( \Lambda_n^{(1)} \lbit \Lambda_n \leq \Lambda_n^{(2)} \), and (8) follows by a suitable choice of \( \rho \).

We proceed now to relate \( \lambda_n \) to \( \Lambda_n \). First, since

\[ \frac{1}{[u + v]} \leq \frac{1}{u + v} \quad (1 \leq u, v \leq n + 1), \]

the largest eigenvalue of the kernel on the left is less than \( \Lambda_n \). The former is, however, just \( \lambda_n \) as can be seen by taking for the eigenfunction a step function whose values are the components of the eigenvector corresponding to \( \lambda_n \). Hence \( \lambda_n \lbit \Lambda_n \).

On the other hand, for \( h > 0 \) we have

\[ \frac{1}{[u + v]} \leq \frac{1 + h}{u + v - 2 + 2h} \quad (1 \leq u, v \leq n + 1). \]

Thus \( \lambda_n \lbit \Lambda_n^{(1)} \) where \( \Lambda_n^{(1)} \) belongs to

\[ \frac{1 + h}{u + v - 2 + 2h} \quad \text{on} \; (1, n + 1). \]

Further, \( \lambda_n^{(0)} = \lambda_n^{(0)} \) where \( \lambda_n^{(0)} \) belongs to

\[ \frac{1 + h}{u + v} \quad \text{on} \; (h, n + h), \]

and \( \lambda_n^{(0)} = \lambda_n^{(0)} \) where \( \lambda_n^{(0)} \) belongs to
THEOREM. The best possible constant for the inequality (2) is of the form
\[ \lambda_n = \pi - \frac{1}{2} \pi^*(\log n)^{-2} + O(\log \log n (\log n)^{-1}) \quad (n \to \infty). \]

Our method consists in relating \( \lambda_n \) to the eigenvalue \( \Lambda_n \) of the corresponding integral equation
\[ \Lambda_n f(u) = \int_1^u \frac{f(v)}{u + v} dv, \]
which seems considerably more tractable.

Indeed let us put, in (5), \( f(u) = u^{-1/2} g(u) \), \( n = e^{2A} \), \( u = e^{2A} \), \( v = e^{2A} \) and \( g(e^{2A}) = \phi(x) \). Then (5) becomes
\[ \Lambda_n \phi(x) = \int_{-A}^{A} K(x - y) \phi(y) dy, \]
where
\[ K(t) = \left( 2 \cosh \frac{t}{2} \right)^{-1}. \]

We will now sketch two proofs of the fact that
\[ \lambda_n = \pi - \frac{1}{2} \pi^*(\log n)^{-2} + O(\log \log n)^{-1} \quad (n \to \infty). \]

The first of these consists merely in observing that problems of the form (6) have been treated by Widom [6, 7] in great generality. Thus, after verifying that the Fourier transform of \( K(t) \) satisfies the hypotheses in [6] we get (8) at once from Theorem 3.1 of [6].

For a direct proof of (8), define
\[ \phi^{(1)}(x) = \cos \frac{\pi x}{2A}, \quad \Lambda^{(1)}_n = \pi \sech \frac{\pi}{2A}, \]
\[ \phi^{(2)}(x) = \cos \frac{\pi x}{2(A + \rho)}, \quad \Lambda^{(2)}_n = \pi \sech \frac{\pi}{2(A + \rho)}. \]

Then \( \phi^{(1)}(x) \) and \( \phi^{(2)}(x) \) are eigenfunctions of
\[ \int_{-A}^{A} K(x - y) \phi(y) dy = \Lambda \phi(x) \]
with eigenvalues \( \Lambda^{(1)}_n \), \( \Lambda^{(2)}_n \) respectively. Now we have
\[ \int_{-A}^{A} K(x - y) \phi^{(1)}(y) dy < 0 \quad (A \leq x \leq A), \]
and similarly for \( \int_{-A}^{A} \). Further, it is not difficult to show that we can fix a positive \( \rho \) such that for all sufficiently large \( A \) we have
\[ \int_{-A}^{A} K(x - y) \phi^{(2)}(y) dy > 0 \quad (A \leq x \leq A), \]
and similarly for \( \int_{-A}^{A} \). Therefore
\[ \int_{-A}^{A} K(x - y) \phi^{(1)}(y) dy \leq \Lambda^{(1)}_n \phi^{(1)}(x), \]
\[ \int_{-A}^{A} K(x - y) \phi^{(2)}(y) dy \leq \Lambda^{(2)}_n \phi^{(2)}(x), \]
and since \( K \), \( \phi^{(1)} \) and \( \phi^{(2)} \) are nonnegative it can be proved in a well-known way that \( \Lambda^{(1)}_n \leq \Lambda_n \leq \Lambda^{(2)}_n \), and (8) follows by a suitable choice of \( \rho \).

We proceed now to relate \( \lambda_n \) to \( \Lambda_n \). First, since
\[ \frac{1}{[u] + [v]} \leq \frac{1}{u + v} \quad (1 \leq u, v \leq n + 1), \]
the largest eigenvalue of the kernel on the left is not less than \( \Lambda_n \).

The former is, however, just \( \lambda_n \) as can be seen by taking for the eigenfunction a step function whose values are the components of the eigenvector corresponding to \( \lambda_n \). Hence \( \lambda_n \leq \Lambda_n \).

On the other hand, for \( h > 0 \) we have
\[ \frac{1}{[u] + [v]} \leq \frac{1 + h}{u + v - 2 + 2h} \quad (1 \leq u, v \leq n + 1). \]

Thus \( \lambda_n \leq \Lambda^{(1)}_n \) where \( \Lambda^{(1)}_n \) belongs to
\[ \frac{1 + h}{u + v - 2 + 2h} \quad \text{on} \ (1, n + 1). \]

Further, \( \lambda_n \leq \lambda^{(2)}_n \) where \( \lambda^{(2)}_n \) belongs to
\[ \frac{1 + h}{u + v} \quad \text{on} \ (h, n + h), \]
and \( \lambda_n \leq \lambda^{(3)}_n \) where \( \lambda^{(3)}_n \) belongs to
\[
\frac{1 + h}{u + v} \quad \text{on} \quad \left(1, \frac{n + h}{h}\right).
\]

But we know that
\[
\lambda_n^{(1)} = (1 + h) \left( 1 - \frac{1}{8} \frac{n}{A_1} + O(A_1^{-1}) \right)
\]

where
\[
A_1 = \frac{1}{2} \log \left( \frac{n + h}{h} \right).
\]

Choosing \( h = \left( (1/2) \log n \right)^{-1} \), the theorem follows.

\textit{Added in proof.} In a forthcoming note in the Nederl. Akad. Wetensch., the second author observes that the methods of this paper apply generally to matrices \( K(m, n) \) if \( K(x, y) \) is symmetric, nonnegative, decreasing and homogeneous of degree \(-1\).

\section*{References}


Technische Hogeschool, Eindhoven and
University of Illinois