ON A GENERALIZATION OF AN INEQUALITY OF
L. V. KANTOROVICH

WERNER GREUB[1] AND WERNER RHEINBOLDT

1. **Introduction.** In his study on the applications of functional analysis to numerical analysis L. V. Kantorovich [1] proves the following inequality:

If the sequence $\{\gamma_k\}$ $(k=1, 2, \cdots)$ of real numbers has the property

$$(1) \qquad 0 < m \leq \gamma_k \leq M$$

and $\{\xi_k\}$ $(k=1, 2, \cdots)$ denotes another sequence with $\sum_{k=1}^{\infty} \xi_k^2 < \infty$ then the inequality

$$(2) \qquad \sum_{k=1}^{\infty} \gamma_k \xi_k^2 \sum_{k=1}^{\infty} \frac{1}{\gamma_k} \xi_k^2 \leq \frac{(M+m)^2}{4Mm} \left[\sum_{k=1}^{\infty} \xi_k^2 \right]^2$$

holds.

In a footnote Kantorovich states that (2) is a special case of the following inequality enunciated by G. Pólya and G. Szegö in their book [2]:

If the real numbers a_k and b_k $(k=1, 2, \cdots, n)$ fulfill the conditions

$$(3) \qquad 0 < m_1 \leq a_k \leq M_1; \qquad 0 < m_2 \leq b_k \leq M_2$$

then

$$(4) \qquad 1 \leq \frac{\sum_{k=1}^{n} a_k^2 \sum_{k=1}^{n} b_k^2}{\left[\sum_{k=1}^{n} a_k b_k \right]^2} \leq \frac{(M_1 M_2 + m_1 m_2)^2}{4 m_1 m_2 M_1 M_2}.$$

G. E. Forsythe, who edited the translation of Kantorovich's paper, included the following remark about this footnote: "It is not clear to me that Kantorovich's inequality really is a special case of that of Pólya and Szegö." Examining the relation between the two inequalities more closely we found that this remark is well justified and can be made even more specific in that the inequality of Pólya and Szegö in the form (4) is a special case of the Kantorovich inequality

(2). This is rather obvious. With given a_k, b_k $(k=1, 2, \cdots, n)$ subject to (3) one only needs to set

$$\gamma_k = \frac{a_k}{b_k}; \qquad \xi_k = (a_k b_k)^{1/2} \qquad (k = 1, \cdots, n)$$

$$\xi_{n+1} = \xi_{n+2} = \cdots = 0$$

while γ_{n+1}, γ_{n+2}, $\cdots$ can be chosen arbitrarily within the interval $[m_1/M_2, M_1/m_2]$.

However, this remark about (2) including (4) does not give a complete picture about the relationship between the two inequalities. The Pólya-Szegö inequality is a direct specialization of the following more general inequality which in turn is equivalent to the Kantorovich inequality:

If $\{a_k\}$ and $\{b_k\}$ $(k=1, 2, \cdots)$ are two sequences of real numbers subject to the condition (3) and $\{\xi_k\}$ denotes another sequence with $\sum_{k=1}^{\infty} \xi_k^2 < \infty$ then

(5) $$\sum_{k=1}^{\infty} a_k^2 \xi_k^2 \sum_{k=1}^{\infty} b_k^2 \xi_k^2 \leqq \frac{(M_1 M_2 + m_1 m_2)^2}{4 m_1 m_2 M_1 M_2} \left[\sum_{k=1}^{\infty} a_k b_k \xi_k^2 \right]^2.$$

One realizes that the specialization

$$\xi_1 = \cdots = \xi_n = 1; \qquad \xi_k = 0 \text{ for } k > n$$

which leads from (5) to (4) is very restrictive and results in (4) not having a geometrical interpretation like (5). Furthermore the inequality (4) ceases to have a meaning for $n \to \infty$. In fact, for $n \to \infty$ both the numerator and denominator of the center term of (4) tend to infinity (while the whole term of course stays finite).

We shall not show at this point that the inequalities (5) and (2) are equivalent. This will be an easy consequence of a proof given later.

The subject of this paper is the proof of a generalized form of the inequality (2) for linear, bounded and selfadjoint operators in Hilbert space. This generalized Kantorovich inequality proves to be equivalent to a similarly generalized form of the inequality (5) which we shall call the generalized Pólya-Szegö inequality. Our generalized Kantorovich inequality is already implicitly contained in the paper of L. V. Kantorovich. However, its proof there involves the use of the theory of spectral decomposition for the operators in question. The proof we shall present here will proceed in a considerably simpler way.

2. The generalized Kantorovich inequality.

THEOREM 1. *Given a linear and selfadjoint operator A of the Hilbert space H. If the operator A fulfills the condition[2]*

(6) $$0 < mE \leqq A \leqq ME$$

then

(7) $$(x, x)^2 \leqq (Ax, x)(A^{-1}x, x) \leqq \frac{(M + m)^2}{4Mm} (x, x)^2$$

for all $x \in H$.

PROOF. The left hand side of the inequality follows directly from Schwarz's inequality[3]

$$(x, x)^2 = (A^{1/2}x, (A^{-1})^{1/2}x)^2 \leqq (A^{1/2}x, A^{1/2}x)((A^{-1})^{1/2}x, (A^{-1})^{1/2}x)$$
$$= (Ax\ x)(A^{-1}x, x).$$

We shall first prove the right hand side of (7) for finite dimensional spaces H. Then we will show that the proof for the general case can be reduced to that of the finite dimensional case.

1. Let H be a finite dimensional space. Then the unit sphere $S \subset H$ is compact. Hence, considered on S, the continuous functional

$$f(x) = \frac{(Ax, x)(A^{-1}x, x)}{(x, x)^2}$$

attains its maximum at a certain point, say $x_0 \in S$, i.e.,

$$f(x_0) = \max_{x \in S} f(x) = (Ax_0, x_0)(A^{-1}x_0, x_0).$$

With a fixed vector $y \in H$ and the real parameter t ($|t| < 1$) we consider the real valued function

$$g(t) = f(x_0 + ty).$$

This function $g(t)$ has a relative maximum at $t = 0$ and therefore we must necessarily have $g'(0) = 0$. Using the selfadjointness of A and A^{-1} we find

$$g'(0) = 2(Ax_0, y)(A^{-1}x_0, x_0) + 2(A^{-1}x_0, y)(Ax_0, x_0) - 4f(x_0)(x_0, y) = 0$$

and thus

---

[2] We use here the usual order relation among selfadjoint operators by writing $A \leqq B$ when $(Ax, x) \leqq (Bx, x)$ for all $x \in H$. E denotes the identity operator in H.

[3] $A^{1/2}$ and $(A^{-1})^{1/2}$ are the positive square roots of the operators A and A^{-1}, respectively. (See e.g. F. Riesz-B.Sz.-Nagy [3, p. 263].)

(8) $$(\gamma A x_0 + \mu A^{-1} x_0 - x_0, y) = 0$$

with

(9) $$\gamma = \frac{1}{2(A x_0, x_0)}; \quad \mu = \frac{1}{2(A^{-1} x_0, x_0)}.$$

(8) must hold for all $y \in H$. Consequently

$$x_0 = \gamma A x_0 + \mu A^{-1} x_0.$$

Applying A and A^{-1} successively to this equation we find that

$$A x_0 = \gamma A^2 x_0 + \mu x_0 \quad \text{and} \quad A^{-1} x_0 = \gamma x_0 + \mu A^{-2} x_0$$

or

$$\left(A - \frac{E}{2\gamma} \right)^2 x_0 = \frac{1 - 4\gamma\mu}{4\gamma^2} x_0 \quad \text{and} \quad \left(A^{-1} - \frac{E}{2\mu} \right)^2 x_0 = \frac{1 - 4\gamma\mu}{4\mu^2} x_0.$$

This shows that

(10a) $\dfrac{1 - 4\gamma\mu}{4\gamma^2}$ is an eigenvalue of $\left(A - \dfrac{E}{2\gamma} \right)^2$,

(10b) $\dfrac{1 - 4\gamma\mu}{4\mu^2}$ is an eigenvalue of $\left(A^{-1} - \dfrac{E}{2\mu} \right)^2$.

Going back to Equation (6) we deduce

$$\left(m - \frac{1}{2\gamma} \right)^2 E \leqq \left(A - \frac{E}{2\gamma} \right)^2 \leqq \left(M - \frac{1}{2\gamma} \right)^2 E$$

and thus for the eigenvalue (10a)

(11a) $$\left(m - \frac{1}{2\gamma} \right)^2 \leqq \frac{1 - 4\gamma\mu}{4\gamma^2} \leqq \left(M - \frac{1}{2\gamma} \right)^2.$$

Similarly we find

$$\left(\frac{1}{M} - \frac{1}{2\mu} \right)^2 E \leqq \left(A^{-1} - \frac{E}{2\mu} \right)^2 \leqq \left(\frac{1}{m} - \frac{1}{2\mu} \right)^2 E$$

and

(11b) $$\left(\frac{1}{M} - \frac{1}{2\mu} \right)^2 \leqq \frac{1 - 4\gamma\mu}{4\mu^2} \leqq \left(\frac{1}{m} - \frac{1}{2\mu} \right)^2.$$

By taking the positive square root of the inequalities (11a/b) we obtain

$$2\gamma m \leqq 1 + (1 - 4\gamma\mu)^{1/2} \leqq 2\gamma M$$

and

$$2\mu \frac{1}{M} \leqq 1 + (1 - 4\gamma\mu)^{1/2} \leqq 2\mu \frac{1}{m} \cdot$$

When combined these inequalities yield

$$4\gamma\mu \frac{m}{M} \leqq (1 + (1 - 4\gamma\mu)^{1/2})^2 \leqq 4\gamma\mu \frac{M}{m} \cdot$$

From this follows

$$\left[4\gamma\mu \left(\frac{m}{M} + 1 \right) - 2 \right]^2 \leqq 4(1 - 4\gamma\mu) \leqq \left[4\gamma\mu \left(\frac{M}{m} + 1 \right) - 2 \right]^2$$

or

$$\frac{4\gamma\mu}{M^2} [4\gamma\mu(M + m)^2 - 4mM] \leqq 0 \leqq \frac{4\gamma\mu}{m^2} [4\gamma\mu(M + m)^2 - 4mM]$$

and therefore

$$4\gamma\mu(M + m)^2 - 4mM = 0.$$

Hence, recalling from (9) that

$$4\gamma\mu = \frac{1}{(Ax_0, x_0)(A^{-1}x_0, x_0)}$$

we finally find

(12) $$(Ax_0, x_0)(A^{-1}x_0, x_0) = \frac{(M + m)^2}{4mM}$$

which was to be proved. (12) shows furthermore that (at least in the finite dimensional case) the upper bound in (7) can not be improved.

2. We now remove the restriction of the finite-dimensionality of H. Let x_0 be a fixed vector of H and let $H_0 \subset H$ be a finite dimensional subspace of H which contains the three vectors x_0, Ax_0 and $A^{-1}x_0$. We denote by P the projection of H onto H_0. For the operator $B = PA$ we have

$$B(H_0) \subset H_0$$

and

$$(Bx, y) = (PAx, y) = (PAPx, y) = (x, PAPy) = (x, By)$$

for all x, $y \in H_0$. Hence, B is a selfadjoint operator in the space H_0. Furthermore, we find for $x \in H_0$

$$(Bx, x) = (PAx, x) = (Ax, Px) = (Ax, x)$$

and therefore in H_0

(13) $$0 < mE \leqq m'E \leqq B \leqq M'E \leqq ME$$

where

$$m' = \inf_{x \in H_0} \frac{(Bx, x)}{(x, x)} ; \qquad M' = \sup_{x \in H_0} \frac{(Bx, x)}{(x, x)} .$$

Hence, we can apply the first part of the proof to the operator B in the finite dimensional space H_0. In doing this we obtain for all $x \in H_0$

(14) $$\frac{(Bx, x)(B^{-1}x, x)}{(x, x)^2} \leqq \frac{(M' + m')^2}{4m'M'} = \frac{1}{4}\left(\frac{M'}{m'} + \frac{m'}{M'}\right) + \frac{1}{2} .$$

From (13) we conclude that

$$1 \leqq \frac{M'}{m'} \leqq \frac{M}{m}$$

and

(15) $$\frac{M'}{m'} + \frac{m'}{M'} \leqq \frac{M}{m} + \frac{m}{M} .$$

This last inequality is a result of the fact that for $u \geqq 1$ the function $f(u) = u + 1/u$ is monotonically increasing. (14) and (15) together yield

(16) $$\frac{(Bx, x)(B^{-1}x, x)}{(x, x)^2} \leqq \frac{1}{4}\left(\frac{M}{m} + \frac{m}{M}\right) + \frac{1}{2} = \frac{(M + m)^2}{4mM}$$

for all $x \in H_0$. Since H_0 contains x_0, Ax_0 and $A^{-1}x_0$ we find

(17) $$Bx_0 = PAx_0 = Ax_0$$

and

$$x_0 = Px_0 = PAA^{-1}x_0 = BA^{-1}x_0.$$

The last relation implies

(18) $$B^{-1}x_0 = A^{-1}x_0$$

when one considers that the existence of B^{-1} in H_0 is a direct conse-

quence of (13). Substituting (17) and (18) in (16) we obtain finally

$$(Ax_0, x_0)(A^{-1}x_0, x_0) \leqq \frac{(M + m)^2}{4mM} (x_0, x_0)^2.$$

Since x_0 was arbitrary the theorem is hereby completely proved.

3. **The generalized Pólya-Szegö inequality.** Theorem 1 is equivalent to the following:

THEOREM 2. *Given two permutable, linear and selfadjoint operators A and B of the Hilbert space H which fulfill the conditions*

(21)　　　$0 < m_1E \leqq A \leqq M_1E; \quad 0 < m_2E \leqq B \leqq M_2E$

then

$$(Ax, Ax)(Bx, Bx) \leqq \frac{(M_1M_2 + m_1m_2)^2}{4m_1m_2M_1M_2} (Ax, Bx)^2$$

for all $x \in H$.

PROOF. (1) It is rather obvious that Theorem 1 is contained in Theorem 2. In fact, let C be any given selfadjoint operator H with

$$0 < ME \leqq C \leqq ME.$$

We set $A = C^{1/2}$ and $B = (C^{-1})^{1/2}$. Since

$$0 < m^{1/2}E \leqq A \leqq M^{1/2}E; \quad 0 < (M^{-1})^{1/2}E \leqq B \leqq (m^{-1})^{1/2}E$$

it follows immediately from Theorem 2 that

$$\frac{(Cx, x)(C^{-1}x, x)}{(x, x)^2} = \frac{(Ax, Ax)(Bx, Bx)}{(Ax, Bx)^2} \leqq \frac{(M + m)^2}{4mM}$$

for all $x \in H$ and this is the statement of Theorem 1.

(2) We will now show that the reverse is also true, i.e., that Theorem 2 is a consequence of Theorem 1.

For the selfadjoint operator $C = AB^{-1}$ we have from (21)

$$0 < \frac{m_1}{M_2} E \leqq C \leqq \frac{M_1}{m_2} E.$$

Therefore, we find from Theorem 1 that

(22)　　$\dfrac{(Cx, x)(C^{-1}x, x)}{(x, x)^2} \leqq \dfrac{(M_1M_2 + m_1m_2)^2}{4m_1m_2M_1M_2}$

for all $x \in H$.

From the permutability of A and B it follows that this will also hold true[4] for the operators $A^{1/2}$ and $B^{1/2}$. Using this together with the selfadjointness of all operators[5] involved we obtain for $x = (AB)^{1/2}y$:

$$(Cx, x) = (AB^{-1}(AB)^{1/2}y, (AB)^{1/2}y) = (Ay, Ay),$$

$$(C^{-1}x, x) = (BA^{-1}(AB)^{1/2}y, (AB)^{1/2}y) = (By, By),$$

$$(x, x) = ((AB)^{1/2}y, (AB)^{1/2}y) = (Ay, By).$$

Substituting these relations in (22) we get the statement of Theorem 1.

4. **Special cases.** We shall now show that the Kantorovich inequality is a direct specialization of Theorem 1, while the same holds true for inequality (5) in regard to Theorem 2. At the same time this will prove that the inequalities (5) and (2) are equivalent.

We consider the Hilbert space l_2 which consists of all sequences of real numbers $x = \{\xi_1, \xi_2, \cdots\}$ with $\sum_{k=1}^{\infty} \xi_k^2 < \infty$. In l_2 we select a complete, orthonormal system of elements $z_1, z_2, \cdots$.

Let now $\{a_k\}$ and $\{b_k\}$ be two arbitrary sequences of real numbers subject to the condition

$$0 < m_1 \leqq a_k \leqq M_1; \qquad 0 < m_2 \leqq b_k \leqq M_2 \qquad (k = 1, 2, \cdots).$$

We consider in l_2 the two operators

$$Ax = \sum_{k=1}^{\infty} a_k \xi_k z_k; \qquad Bx = \sum_{k=1}^{\infty} b_k \xi_k z_k$$

where $\xi_k = (x, z_k)$ are the components of x.

These two operators in l_2 obviously fulfill all conditions of Theorem 2. Hence, we obtain

$$\sum_{k=1}^{\infty} a_k^2 \xi_k^2 \sum_{k=1}^{\infty} b_k^2 \xi_k^2 \leqq \frac{(M_1 M_2 + m_1 m_2)^2}{4 m_1 m_2 M_1 M_2} \left[\sum_{k=1}^{\infty} a_k b_k \xi_k^2 \right]^2$$

for all $x \in l_2$. This is inequality (5).

Since the inverse of the operator Ax can be written in the form

$$A^{-1}x = \sum_{k=1}^{\infty} \frac{1}{a_k} \xi_k z_k$$

we obtain similarly from Theorem 1

---

[4] See e.g. F. Riesz-B. Sz.-Nagy [3, p. 264].
[5] Here we make use of the fact that $(AB)^{1/2}$ is a transformation of H onto H.

$$\sum_{k=1}^{\infty} a_k \xi_k^2 \sum_{k=1}^{\infty} \frac{1}{a_k} \xi_k^2 \leqq \frac{(M+m)^2}{4mM} \left[\sum_{k=1}^{\infty} \xi_k^2 \right]^2$$

for all $x \in l_2$. This is inequality (2).

REFERENCES

1. L. V. Kantorovich, *Functional analysis and applied mathematics*, Uspehi Mat. Nauk, vol. 3 (1948) pp. 89–185. Translated from the Russian by Curtis D. Benster, National Bureau of Standards, Report No. 1509, March 7, 1952.

2. G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, Springer, Berlin, 2d ed., 1954, p. 57.

3. F. Riesz and B. Sz.-Nagy, *Lectures on functional analysis*, New York, F. Ungar Publishing Company, 1955.

Mathematisches Institut, Universität Zürich, Switzerland and
National Bureau of Standards

ON A PROBLEM OF LOHWATER

G. PIRANIAN

Bagemihl [1] has shown that, for each complex-valued function f defined in the unit disk D, there exist at most countably many points $e^{i\theta}$ that are endpoints of two Jordan arcs $\gamma_1(\theta)$ and $\gamma_2(\theta)$ such that the corresponding cluster sets $C(f, e^{i\theta}; \gamma_1(\theta))$ and $C(f, e^{i\theta}; \gamma_2(\theta))$ are disjoint. Lohwater [2, p. 173] has recently asked whether there exists a function f for which uncountably many points $e^{i\theta}$ are endpoints of three Jordan arcs $\gamma_1(\theta)$, $\gamma_2(\theta)$, $\gamma_3(\theta)$ such that the intersection

$$C(f, e^{i\theta}; \gamma_1(\theta)) \cap C(f, e^{i\theta}; \gamma_2(\theta)) \cap C(f, e^{i\theta}; \gamma_3(\theta))$$

is empty. The property of points just described will be called the *three-path property*.

THEOREM. *There exists a complex-valued function, continuous in the unit disk, for which each point $e^{i\theta}$ has the three-path property.*

My proof is based on a slight modification of a technique recently used in connection with a problem on ambiguous points of a function defined in the unit sphere [3]. Let T_1, T_2, T_3 be three trees in D, with the property that each point $e^{i\theta}$ can be approached along three Jordan arcs which lie on T_1, T_2, T_3, respectively, except for their common endpoint $e^{i\theta}$. Let T_1, T_2, T_3 have the further property that no point

---