ON AN INEQUALITY FOR OPERATORS
ON HILBERT SPACE

T. I. Seidman

Report 68-7

February, 1968

On an Inequality for Operators on Hilbert Space

T. I. Seidman

§1 Recently, Z. Nehari mentioned to the author the inequality (A, any bounded linear operator on a Hilbert space H)

$$(1)\quad \sup\{\|Ax\|^2 - |\langle x,Ax\rangle|^2 : x \in H,\ \|x\| = 1\} \leq \inf\{\|A-zI\|^2 : z \in C\}$$

and raised the question: Under conditions on A does equality hold? It is not difficult to obtain a necessary and sufficient condition on A for equality to hold in (1) but the condition obtained below (Theorem 6) is, unfortunately, unlikely to be useful. An independent proof is given of a sufficient condition from which it follows that equality holds in (1) if A is normal. A number of preliminary results will be obtained along the way (§2, §3); in the hope that these may be of interest in themselves, they are obtained in somewhat greater generality than would be required for their application in §4. Thanks are due to V. J. Mizel for the original proof of equality when A is an hermitian matrix and for several stimulating discussions.

For completeness we include a proof of a more general form of the inequality (1) applying to nonlinear mappings. Let A be a Lipschitzian function whose domain and codomain are the Hilbert space H. For $x,y \in X$ with $x \neq y$, let

$$\gamma(A;x,y) = [\|A(x)-A(y)\|^2 - |\langle x-y, A(x)-A(y)\rangle|^2]/\|x-y\|^4.$$

Defining A_z by $A_z(x) = A(x) + zx$ for $z \in C$, we have

$$\|x-y\|^4\gamma(A_z;x,y) = \|A(x)-A(y)+z(x-y)\|^2\|x-y\|^2 - |\langle x-y, A(x)-A(y)+z(x-y)\rangle|^2$$
$$= [\|A(x)-A(y)\|^2 + |z|^2\|x-y\|^2 + \operatorname{Re} z\langle x-y, A(x)-A(y)\rangle]\|x-y\|^2$$
$$-[|z|^2\|x-y\|^4 + |\langle x-y, A(x)-A(y)\rangle|^2 + 2\operatorname{Re} z\|x-y\|^2\langle x-y, A(x)-A(y)\rangle]$$
$$= \|x-y\|^4\gamma(A;x,y).$$

Clearly,

$$\gamma(A;x,y) = \gamma(A_z;x,y) \le \sup\{\|A_z(x)-A_z(y)\|^2/\|x-y\|^2 : x,y\in H, x\ne y\}$$

whence

$$(1') \quad \sup\{\gamma(A;x,y) : x,y\in H, x\ne y\}$$
$$\le \inf\{\sup\{\|A(x)-A(y)-z(x-y)\|^2/\|x-y\|^2 : x,y\in H, x\ne y\} : z\in C\}$$

This, when A is linear, reduces to (1).

§2. In this section A will denote a bounded operator on a Hilbert space H. Define the *minimal radius* of A (denoted by $R(A)$) by

$$R(A) = \inf\{\|A-zI\| : z \in C\}.$$

As the function $z \to \|A-zI\|$ is continuous and as we may restrict our attention, in taking the infimum, to the compact set $\{z \in C : |z| \le 2\|A\|\}$, the infimum is taken on; a complex number z such that $\|A-zI\| = R(A)$ will be called a *minimal center* for A (we show below that this is unique).

THEOREM 1: The function φ, defined by $\varphi(z) = \|A-zI\|^2$, is a strictly convex function from C to R.

Proof: Set $\varphi(z;x)=\|[A-zI]x\|^2=\langle Ax-zx,Ax-zx\rangle$ so that $\varphi(z)=\sup_x\varphi(z;x)$. Let $t,s\geq 0$, $t+s=1$, $x\in H$, $\|x\|=1$, $u,v\in C$; set $x_u=Ax-ux$, $x_v=Ax-vx$. Then

$$\begin{aligned}\varphi(tu+sv;x) &= \|tx_u+sx_v\|^2\\ &= t^2\|x_u\|^2+s^2\|x_v\|^2+2ts\,\mathrm{Re}\langle x_u,x_v\rangle\\ &= (t\|x_u\|^2+s\|x_v\|^2)-ts(\|x_u\|^2+\|x_v\|^2-2\mathrm{Re}\langle x_u,x_v\rangle).\end{aligned}$$

Since $(\|x_u\|^2+\|x_v\|^2-2\mathrm{Re}\langle x_u,x_v\rangle) = \|x_u-x_v\|^2 = |u-v|^2$, this gives

$$\varphi(tu+sv;x) = t\varphi(u;x)+s\varphi(v;x)-ts|u-v|^2.$$

Since this holds for each $x\in H$ with $\|x\|=1$, we have

$$\varphi(tu+sv) \leq t\varphi(u)+s\varphi(v)-ts|u-v|^2$$

so that φ is strictly convex.

COROLLARY: There is a unique minimal center (denoted by $z(A)$).

Proof: It has already been shown that a minimal center exists. By definition a minimal center gives a minimum for φ and, by the strict convexity of φ, this is unique.

The disk $B(A)=\{z\in C: |z-z(A)|\leq R(A)\}$ will be called the *minimal disk* of A. Denoting by $w(A)$ the *numerical range* ($w(A)=\{\langle Ax,x\rangle: x\in H, \|x\|=1)\}$, it is clear that $\sigma(A)\subseteq w(A)\subseteq B(A)$. We next see that for points of the spectrum which are boundary points of $w(A)$ the 'spectral theory' is as in the case of a normal operator.

THEOREM 2: Let $\lambda_o \epsilon \sigma(A)$ with λ_o a boundary point of $w(A)$. Then, (i) there is an 'approximate eigenvector' (for short: aev) associated with λ_o (i.e., a sequence $\{x_n : n=1,2,\ldots\}$ with $x_n \epsilon H$, $\|x_u\| = 1$ and $\|[A-\lambda_o I]x_n\| \to 0$), (ii) if $\{x_n\}$ is an aev of A associated with λ_o, it is also an aev of A^* associated with $\bar{\lambda}_o$, (iii) if $\{x_n\}$ is an aev associated with λ_o and $\{y_n\}$ is an aev associated with $\lambda \epsilon \sigma(A)$ $(\lambda \neq \lambda_o)$, then they are 'ultimately orthogonal' (i.e., $\langle x_n, y_n \rangle \to 0$), (iv) if $\|x_n\| = 1$ for $n=1,2,\ldots$ and $\|[A-\lambda_o]^m x_n\| \to 0$ for some $m \geq 1$ then $\{x_n\}$ is an aev associated with λ_o (in particular, if $x \neq 0$ and $[A-\lambda_o I]^m x = 0$ then x is an eigenvector so the index of λ_o is 1).

Proof: As $\sigma(A) \subseteq w(A)$, λ_o must be a boundary point of $\sigma(A)$ and so is in the approximate point spectrum of A; this is just property (i). As $w(A)$ is a convex set in C and as every real-linear functional on C is of the form $z \to \ell_a(z) = \mathrm{Re}\, az$ with $\|\ell_a\| = |a|$, there is a complex number a with $|a| = 1$ such that $\mathrm{Re}\, az \geq \mathrm{Re}\, a\lambda_o$ for $z \epsilon w(A)$. Replacing the operator A by $a[A-\lambda_o I]$, there is clearly no loss of generality in assuming for the remainder of this proof that $a=1$ and $\lambda_o = 0$ so that $0 \epsilon \sigma(A)$ and

$$(*) \qquad \mathrm{Re}\, \langle Ax, x \rangle \geq 0 \; .$$

If, now, $\|x\| = 1$, $\|Ax\| \leq \epsilon$, we have (setting $\alpha = 1/2\|A\|$, $y = x - \alpha A^* x$ so $\|y\| \leq 3/2$)

$$0 \le \mathrm{Re}\langle Ay,y\rangle = \mathrm{Re}\langle Ax,y\rangle - \alpha\mathrm{Re}\langle AA^*x,x\rangle + \alpha^2\mathrm{Re}\langle AA^*x,A^*x\rangle$$
$$\le \|Ax\|\|y\| - \alpha\|A^*x\|^2 + \alpha^2\|A\|\|A^*x\|$$
$$\le 3\epsilon/2 - \alpha(1-\alpha\|A\|)\|A^*x\|^2 = (3\epsilon - \alpha\|A^*x\|^2)/2 .$$

Thus,

$$(2) \qquad \|A^*x\| \le (6\|A\|\epsilon)^{1/2} \quad \text{if} \quad \|x\| = 1,\ \|Ax\| \le \epsilon .$$

Property (ii) follows immediately from this implication. If $\|x\|=\|y\|=1$ and $\|Ax\| \le \epsilon$, $\|Ay-\lambda y\| \le \epsilon$, then

$$|\lambda\langle y,x\rangle| = |\langle Ay,x\rangle - \langle[A-\lambda I]y,x\rangle|$$
$$\le \|y\|\|A^*x\| + \|Ay-\lambda y\|\|x\| = (6\|A\|\epsilon)^{1/2} + \epsilon$$

so that property (iii) follows. Finally, if $\|x\|=1$, $\|A^m x\| < \epsilon$ with $m \ge 2$ set $B = A^{m-1}$ and $y = Bx/\|Bx\|$. Then $\|By\| = \|A^{m-2}(A^m x)\|/\|Bx\| \le \|A\|^{m-2}\epsilon/\|Bx\|$ so that, by (2) $\|B^*y\|^2 \le 6\|B\|\|A\|^{m-2}\epsilon/\|Bx\|$. Thus,

$$\|Bx\| = \langle Bx,Bx\rangle^2 = \langle x,B^*y\rangle^2\|Bx\|^2$$
$$\le \|x\|\|B^*y\|^2\|Bx\|^2 \le 6\|B\|^2\|A\|^{m-2}\epsilon .$$

Hence, $\|x_n\| = 1$, $\|A^m x_n\| \to 0$ $(m\ge 2)$ implies $\|A^{m-1}x_n\| \to 0$ as $n \to \infty$ and (iv) follows by induction on m.

<u>§3</u> In this section S will denote a non-empty bounded subset of a real Banach space X. For $x_0 \in X$, define $r(S,x_0)$ by

$$r(S,x_0) = \sup\{\|x-x_0\|: x\in S\}$$
$$= \inf\{r: S \subseteq B_r(x_0)\}$$

where $B_r(x_0)$ is the closed ball with radius r centered at x_0 (i.e., $B_r(x_0) = \{x\in X: \|x-x_0\| \le r\}$). The *minimal radius* of S (denoted by $r(S)$) is given by

$$r(S) = \inf\{r(S,x) : x \in X\}.$$

A point $x \in X$ such that $r(S,x) = r(S)$ is called a minimal center of S and the corresponding ball $B_{r(S)}(x)$ is called a minimal ball of S; if the minimal center exists uniquely it is denoted by $x(S)$ and the corresponding minimal ball by $B(S)$. We denote by $\bar{S}$ and S^* the closure and closed convex hull, respectively, of $S \subset X$; as the balls $B_r(x)$ are closed and convex, any one which contains S also contains $\bar{S}$ and S^* so $r(S^*) = r(\bar{S}) = r(S)$ and, if it is defined, $x(S^*) = x(\bar{S}) = x(S)$.

THEOREM 3: Let $S \subset X$ as above with X uniformly convex. Then $x(S)$ (and so, too, $B(S)$) is well-defined.

REMARK: Under the weaker condition that X is reflexive it can be shown that the map $x \to r(S,x)$ is weakly lower semi-continuous. As, in taking the infimum, attention can be restricted to the weakly sequentially compact set $\{x \in X : \|x\| \le 2 \sup\{\|y\| : y \in S\}\}$, the infimum must be attained and a minimal center must exist. In this case the set of minimal centers is a non-empty bounded convex set - but need not be a singleton.

Proof (of Theorem 3): Let $r = r(S)$ be the minimal radius. Then there must exist a sequence $\{x_n\}$ of points $x_n \in X$ such that $r_n = r(S,x_n) \to r$. The condition of uniform convexity asserts that: for every $\epsilon > 0$ there is a $\delta = \delta(\epsilon)$ with $0 < \delta \le 1/2$ such that, for $u,x \in X$ with $\|u\|, \|v\| \le 1$ and $\|(u+v)/2\| \ge 1-\delta$, one has $\|u-v\| \le \epsilon$. Given $\epsilon > 0$, let n,m be large enough that

$r \leq r_n, r_m \leq r' < r/(1-\delta)$. By the minimality of r there exists $x \in S$ such that $\|x-a\| \geq r$ where $a = (x_n+x_m)/2$; take $u = (x-x_n)/r'$, $v = (x-x_m)/r'$ so $\|u\| = \|x-x_n\|/r' \leq r_n/r' \leq 1$ and, similarly, $\|v\| \leq 1$. Then $\|(u+v)/2\| = \|x-a\|/r' \geq r/r' > 1-\delta$ so $\|x_n-x_m\| = r'\|u-v\| < r'\epsilon \leq 2r\epsilon$. Thus $\{x_n\}$ is a Cauchy sequence and $x_o = \lim_n x_n$ is defined. By the continuity of the map $x \to r(S,x)$, $r(S,x_o) = \lim_n r_n = r$ so x_o is a minimal center of S. That <u>every</u> sequence $\{x_n\}$ for which $r(S,x_n) \to r$ must be a Cauchy sequence proves that this limit x_o is uniquely determined so $x_o = x(S)$.

THEOREM 4: Let S,X be such that $B(S)$ is well-defined and such that, for every subspace $Y \subset X$ with codim $Y = 1$, there is a projection P onto Y with $\|P\| = 1$. Then $x(S) \in S^*$.

REMARK: If X is 3-space with a regular octohedron as unit ball, then taking S to be a face of this octahedron gives an example such that $x(S)$ is well-defined but $0 = x(S) \notin S^*$. Perturbing this slightly, 'puffing out' the octahedron a very little, provides an example such that X is uniformly convex but $x(S) \notin S^*$.

<u>Proof (of Theorem 4)</u>: There is no loss of generality in assuming that $x(S) = 0$, $r(S) = 1$ so $B(S)$ is the unit ball of X. If $0 = x(S) \notin S^*$, then, as S^* is closed and convex, there exists $\xi \in X^*$ with $\|\xi\| = 1$ such that $\langle x,\xi\rangle \geq d > 0$ for all $x \in S^*$. Let Y be the nullspace of ξ, so codim $Y = 1$, and let P be a projection onto Y with $\|P\| = 1$; let a be in the nullspace of P

with $\langle a,\xi\rangle = \alpha$. Now, for $x \in S$ we have $\|x\| \le 1$ so $\|Px\| \le 1$; as $x-Px = ta$ with $t = \langle x,\xi\rangle/\alpha \ge 1$ we have $x-a$ on the segment joining x and $x-Px$. Thus, for $x \in S$ we have $\|x-a\| \le 1$ so that $r(S,a) \le 1$, contradicting the assumption that $B(S)$ is the unit ball of X.

COROLLARY: If X is a Hilbert space or $\dim X = 2$ then, if $x(S)$ is defined, $x(S) \in S^*$.

<u>Proof</u>: If X is a Hilbert space then, for any subspace Y, the orthogonal projection onto Y has norm one. If $\dim X = 2$ then $\operatorname{codim} Y = 1$ implies $\dim Y = 1$ so that, by a theorem of Kakutani, there is a projection of norm one onto Y. (Indeed, if Y is a 1-dimensional subspace of any Banach space we have $Y = \operatorname{sp}\{y\}$ with $\|y\| = 1$; by the Hahn-Banach Theorem there exists $\xi \in X^*$ such that $\langle y,\xi\rangle = \|\xi\| = 1$, in which case $x \to \langle x,\xi\rangle y$ is the desired projection.)

If $B(S)$ is well-defined, let $S^o = \{x \in \overline{S} : \|x-x(S)\| = r(S)\}$. We wish to show that $B(S)$ is determined by the 'extreme points', in this sense, of S. For general S this need not be true even if X is quite 'nice'. For example, if $X = \ell_2$ and $S = \{\pm(1-1/n)e_n : n = 1,2,\ldots\}$ where $\{e_n : n=1,2,\ldots\}$ is an orthonormal sequence, then $S = \overline{S}$, $B(S)$ is the unit ball, but S^o is empty. We have, however, the following.

THEOREM 5: Let S be compact and X uniformly convex. Then $B(S) = B(S^o)$.

__Proof__: There is no loss of generality in assuming $x(S) = 0$, $r(S) = 1$ so $B(S)$ is the unit ball of X. By the compactness of $\bar{S}$, S^o is non-empty; by its definition, $r(S^o,0) = 1$. Thus, if $B(S^o) \neq B(S)$ we must have $r(S^o) < 1$ and $x(S^o) = a \neq 0$; hence, $\|x-a\| < 1$ for $x \in S^o$.

For $x \in \bar{S}$ let $t(x) = \sup\{t: \|x-ta\| \leq 1\}$. For $x \in \bar{S} \setminus S^o$, $\|x\| < 1$ and $t(x) \geq (1-\|x\|)/\|a\|$ while, for $x \in S^o$, $t(x) \geq 1$ by the assumption above; thus, $t(x) > 0$ for all $x \in \bar{S}$. Let $x_o \in \bar{S}$, $t = t(x_o)$ so $\|x_o - ta\| = 1$. For $0 < \epsilon < t$, set $x_\epsilon = x_o - (t-\epsilon)a$; then, by the uniform convexity of X, $\|x_\epsilon\| < 1$ and we set $\delta(\epsilon) = 1 - \|x_\epsilon\| > 0$. Now, for $x \in \bar{S}$ with $\|x - x_o\| < \delta(\epsilon)$, $x-(t-\epsilon)a = (x-x_o) + x_\epsilon$ so $\|x-(t-\epsilon)a\| \leq 1$ and $t(x) \geq t-\epsilon$. Thus, the map $x \to t(x)$ is lower semi-continuous on the compact set $\bar{S}$ and so attains its minimum t_o on $\bar{S}$; as $t(x) > 0$ for $x \in \bar{S}$, $t_o > 0$. It follows that $\|x-t_o a\| \leq 1$ for $x \in \bar{S}$ so $r(S, t_o a) \leq 1$ which contradicts the uniqueness of $B(S)$.

__§4__ In this section A is a bounded operator on a Hilbert space H. Let $A_o = A - z(A)I$; then $z(A_o) = 0$ and $R(A_o) = R(A) = \|A_o\| = \alpha$. We may write A_o in polar form: $A_o = UP$ with P non-negative and U unitary. For $x \in H$ set

$$\gamma(A;x) = \|Ax\|^2 - |\langle x, Ax\rangle|^2$$

and let $\gamma(A) = \sup\{\gamma(A;x): \|x\| = 1\}$ (clearly $\gamma(A) \leq \|A\|^2$). As in 1, we have

$$\begin{aligned}\gamma(A+zI;x) &= \langle Ax+zx, Ax+zx\rangle - |\langle x, Ax+zx\rangle|^2 \\ &= [\|Ax\|^2 + |z|^2\|x\|^2 + 2\mathrm{Re}\ z\langle x,Ax\rangle] \\ &- [|\langle x,Ax\rangle|^2 + |z|^2\|x\|^2 + 2\mathrm{Re}\ z\langle x,Ax\rangle] \\ &= \gamma(A;x)\end{aligned}$$

so that $\gamma(A+zI) = \gamma(A)$. The inequality (1) now takes the form

$$\gamma(A) \le R(A)^2 \tag{3}$$

which follows from

$$\gamma(A) = \gamma(A-z(A)I) \le \|A-z(A)I\|^2 = R(A)^2.$$

THEOREM 6: Equality holds in (3) (i.e., $\gamma(A) = R(A)^2$) if and only if there is an approximate eigenvector $\{x_n\}$ of P associated with α (as P is non-negative with $\|P\| = \|A_o\| = \alpha$, α is in the approximate point spectrum of P) such that $\langle x_n, Ux_n\rangle \to 0$.

Proof: Suppose, first, that there is such an aev of P. Then

$$\begin{aligned}\gamma(A;x_n) &= \gamma(A_o;x_n) = \|A_o x_n\|^2 - |\langle x_n, A_o x_n\rangle|^2 \\ &= \|Px_n\|^2 - |\langle x_n, UPx_n\rangle|^2 \\ &\le \alpha^2\|x_n\|^2 + 2\alpha\|Px_n - \alpha x_n\| - \alpha^2|\langle x_n, Ux_n\rangle|^2 + \|x_n\|^2\|Px_n - \alpha x_n\|^2.\end{aligned}$$

Since $\|x_n\| = 1$ and $\|Px_n - \alpha x_n\| \to 0$, $\gamma(A;x_n) \to \alpha^2$; thus $\gamma(A) \ge R(A)^2$ which, with (3) gives the desired equality.

Conversely, suppose equality holds in (3). Then there exists, for any $\epsilon > 0$, $x_\epsilon \in H$ with $\|x_\epsilon\| = 1$ such that $\gamma(A;x_\epsilon) \ge \alpha^2 - \epsilon^2$; i.e.,

$$\begin{aligned}\gamma(A;x_\epsilon) &= \gamma(A_o;x_\epsilon) \\ &= \|Px_\epsilon\|^2 - |\langle x_\epsilon, UPx_\epsilon\rangle|^2 \ge \alpha^2 - \epsilon^2 .\end{aligned}$$

It follows that

$$\epsilon^2 \geq \alpha^2 - \|Px_\epsilon\|^2 = \int_0^\alpha (\alpha^2-\lambda^2)\,\|dE_\lambda x_\epsilon\|^2$$

$$= \int_0^\alpha (\alpha-\lambda)^2 \|dE_\lambda x_\epsilon\|^2 = \|\alpha x_\epsilon - Px_\epsilon\|^2$$

where $\{E_\lambda\}$ is the spectral resolution of P; thus $\|Px_\epsilon - \alpha x_\epsilon\| \leq \epsilon$. Now,

$$\epsilon \geq |<x_\epsilon, UPx_\epsilon>| = |\bar{\alpha}<x_\epsilon . Ux_\epsilon> + <x_\epsilon, U[Px_\epsilon - \alpha x_\epsilon]>|$$
$$\geq |\alpha|\,|<x_\epsilon, Ux_\epsilon>| - \|Px_\epsilon - \alpha x_\epsilon\|$$

so that $|<x_\epsilon, Ux_\epsilon>| \leq 2\epsilon/|\alpha|$ (if $\alpha=0$, then A_0 is the 0 operator and the result is trivial). The sequence $\{x_{1/n}: n=1,2,\ldots\}$ is thus the desired aev of P.

This condition for equality is not very helpful. We give an independent proof of a sufficient condition which seems handier. Considering the spectrum $\sigma(A)$ as a subset of C (considered as real 2-space) we may apply the results of 3; it seems more natural to refer to the *minimal disc*, rather than to the minimal ball, of $\sigma(A)$. It is clear that $\sigma(A) \subseteq B(A)$, that $r(\sigma(A)) \leq r(\sigma(A), z(A))$ which is the spectral radius of A_0, and that $r(\sigma(A)) \leq R(A) = \|A_0\|$.

THEOREM 7: If $r(\sigma(A)) = R(A)$ (so the minimal disk of A is just the minimal disk of $\sigma(A)$), then equality holds in (3).

Proof: Without loss of generality we may assume $z(A) = 0$ and $r(\sigma(A)) = R(A) = 1$; since $\sigma(A) \subseteq B(A) = B_1(0)$, this implies that $B(\sigma(A)) = B(A)$. As $\sigma(A)$ is compact, we may apply Theorem 5 to show that $B(\sigma_o) = B_1(0)$ where $\sigma_o = \{z\epsilon\sigma(A): |z|=1\}$. By Theorem 4, then, 0 is in the convex hull of σ_o (which is already closed) so that

$$0 = \Sigma_k c_k^2 z_k, \ \Sigma_k c_k^2 = 1, \ |z_k| = 1$$

where $\{z_k\}$ is a finite subset of σ_o and the $\{c_k\}$ are non-negative reals. Clearly, each z_k is not only a boundary point of $\sigma(A)$ but, as $w(A) \subseteq B(A) = B_1(0)$, each z_k is a boundary point of $w(A)$. It now follows from Theorem 2 that, for any $\epsilon > 0$, there are elements $x_k \epsilon H$ such that

$$\|x_k\| = 1, \ \|Ax_k - z_k x_k\| \leq \epsilon, \ |\langle x_j, x_k\rangle| \leq \epsilon \quad \text{for} \quad j \neq k .$$

Now set

$$x = \Sigma_k c_k x_k, \quad y = \Sigma_k c_k z_k x_k .$$

Then

$$\begin{aligned}
\|Ax-y\| &\leq \Sigma\, c_k \|Ax_k - z_k x_k\| \leq \epsilon \\
\|y\|^2 &= \Sigma_{j,k}\, c_j c_k z_j \overline{z}_k \langle x_j, x_k\rangle \\
&\geq \Sigma_k\, c_k^2 |z_k|^2 \|x_k\|^2 - \Sigma_{j\neq k}\, c_j c_k |z_j \overline{z}_k| \, |\langle x_j, x_k\rangle| \\
&\geq 1 - \epsilon\, \Sigma_{j\neq k}\, c_j c_k \geq 1 - K(K-1)\epsilon
\end{aligned}$$

where K is the cardinality of $\{z_k\}$ (as we are working, here, in 2-space, it can be shown that we may take $K \leq 3$); thus, as $\|Ax\| \geq \|y\| - \|Ax-y\|$, we have $\|Ax\|^2$ arbitrarily close to 1 as ϵ gets small. At the same time,

$$\begin{aligned}
|\langle y,x\rangle| &= |\Sigma_{j,k}\, c_j c_k \overline{z}_k \langle x_j, x_k\rangle| \\
&\leq |\Sigma_k\, c_k^2 z_k \|x_k\|^2| + \Sigma_{j\neq k}\, c_j c_k |z_k| \, |\langle x_j, x_k\rangle| \\
&= 0 + \Sigma_{j\neq k}\, c_j c_k |\langle x_j, x_k\rangle| \leq K(K-1)\epsilon
\end{aligned}$$

so that $|\langle x,Ax\rangle| \leq |\langle y,x\rangle| + |\langle x,Ax-y\rangle| \leq (K^2-K+1)\epsilon$ which goes to 0 as ϵ does. Combining these shows that $\gamma(A;x)$ may be made arbitrarily close to 1 by taking ϵ small so that $\gamma(A) = 1 = R(A)^2$.

COROLLARY: If A is normal, then

$$\sup\{\|Ax\|^2 - |\langle x,Ax\rangle|^2 : \|x\| = 1\} = \inf_z\{\|A-zI\|^2\}.$$

<u>Proof</u>: If A is normal then

$$\|A-zI\| = \sup\{|\lambda-z| : \lambda\in\sigma(A)\}$$

(i.e., $r(\sigma(A),z) = \|A-zI\|$) so $r(\sigma(A)) = R(A)$ and Theorem 7 applies.