Antieigenvalues: Method of Estimation and Calculation

Boris Mirman
Digital Equipment Corporation
Maynard, Massachusetts 01754

Submitted by Chandler Davis

ABSTRACT

Let T be a linear operator on a Hilbert space $\mathcal{H}$. A method of estimation and calculation of $\inf\{\operatorname{Re}(Tx,x)/\|Tx\| : x \in \mathcal{H}, \|x\| = 1\}$ as well as some other expressions containing two quadratic forms is proposed. This method is based on the Toeplitz-Hausdorff theorem on the convexity of the numerical range of any operator on $\mathcal{H}$.

Throughout the paper, $\mathcal{H}$ is a Hilbert space with inner product (x, y), A, B, S, T denote linear operators in $\mathcal{H}$, and $A > 0$ implies that A defines a strictly positive hermitian form. An operator T is strictly accretive if $\operatorname{Re} T = \frac{1}{2}(T^* + T) > 0$.

M. G. Krein [1] and K. Gustafson [2] have pointed out the importance of evaluating

$$\mu_1(T) = \inf_{x \neq 0} \frac{\operatorname{Re}(Tx, x)}{\|Tx\|\|x\|} \tag{1}$$

for determining some properties of strictly accretive T. K. Gustafson [2] formulated the problem of finding $\mu_1(T)$, which he called the first antieigenvalue of T. He also considered higher antieigenvalues

$$\mu_n(T) = \inf\left\{ \frac{\operatorname{Re}(Tx, x)}{\|Tx\|\|x\|} : x \neq 0, x \perp x^{(1)}, \dots, x^{(n-1)} \right\}, \tag{2}$$

where it was assumed that $\mu_k(T)$ was actually attained by a kth *antieigenvector* $x^{(k)}$ for $k = 1, \dots, n-1$.

In this paper, a method of computation of $\mu_k(T)$ is proposed. The structure of a set of antieigenvectors is found. Upper and lower estimates of $\mu_1(T)$ are given in the following, including some known results (L. V. Kantorovich [3], Ch. Davis [4]). This work constitutes a new approach to the

52 Vanderbilt Ave., New York, NY 10017 0024-3795/83/010247-09$3.00

results in this area by K. Gustafson and his collaborators [2, 5, 6]. Finally, the method considered here allows one to find extrema of expressions of the form $f((Ax,x),(Bx,x))$, where $A=A^*$, $B=B^*$, and $f(\xi,\eta)$ is some function of real ξ and η.

This method is based on consideration of the numerical range $\Omega(S)$ of the operator $S=A+iB$. $\Omega(S)$ is the set of values (Sx,x) on the unit sphere of $\mathcal{H}$:

$$\Omega(S)=\{(Sx,x):\|x\|=1\}.$$

To simplify matters, let's mention the following well-known properties of numerical ranges of operators:

1. $\Omega(S)$ is convex.
2. The closure $\overline{\Omega}(S)$ of the numerical range of S contains its spectrum $\sigma(S)$.
3. If S is normal, then $\overline{\Omega}(S)$ is the convex hull $\operatorname{Co}\sigma(S)$ of $\sigma(S)$.
4. Consider $\Omega(S)$ on the plane of complex numbers $\xi+i\eta$. Let l be a straight line of support of $\Omega(S)$, with $\Omega(S)$ to the right of l. Moreover, let the angle between l and the ξ axis be denoted by ϕ, and the intersection of l and the ξ axis be denoted by λ (see Figure 1). Then λ is the lower bound of $\sigma(A-B\cot\phi)$, where $A=\operatorname{Re}S=\frac{1}{2}(S+S^*)$, $B=\operatorname{Im}S=(1/2i)(S-S^*)$.
5. If l has a common point $[\xi;\eta]$ with $\Omega(S)$, then this point is generated by some eigenvectors of operator $A-B\cot\phi$; namely, if $\|x\|=1$, $(Sx,x)=\xi+i\eta$, then $(A-B\cot\phi)x=\lambda x$. In the case of $\dim\mathcal{H}<\infty$ it is possible to

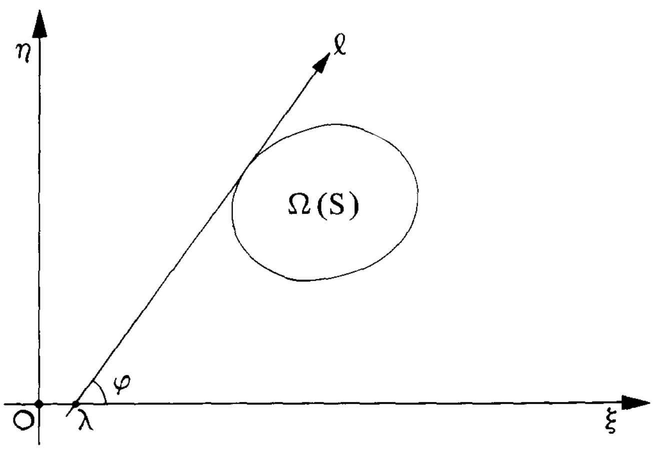


FIG. 1.

state more: let $[\xi;\eta]\in l$ and λ be a k-fold eigenvalue of $A-B\cot\phi$; if $[\xi;\eta]$ is an extreme point of $\Omega(S)$, then $\{x:(Sx,x)=(\xi+i\eta)\|x\|^2\}$ is a subspace; if $[\xi;\eta]$ is the only common point of l and $\Omega(S)$ or if $[\xi;\eta]$ is not an extreme point of $\Omega(S)$, then the number of independent unit vectors which generate the point $[\xi;\eta]$ is equal to k.

6. If $\xi+i\eta$ is an angle point of $\Omega(S)$, then $\xi+i\eta\in\sigma(S)$.
7. The numerical range of a two-dimensional operator

$$\begin{bmatrix}\alpha_1 & \lambda\\ 0 & \alpha_2\end{bmatrix}$$

is an ellipse with foci α_1, α_2 and minor axis $|\lambda/2|$.

THEOREM 1. *Let T be strictly accretive. If $\|T\|<\infty$, then*

$$\mu_1^2(T)=4\max_t(\lambda_t t)>0, \tag{3}$$

where λ_t is the lower bound of the spectrum of the operator $S_t=\operatorname{Re}T-tT^*T$.

Proof. Let $S=\operatorname{Re}T+iT^*T$. According to (1), $\mu_1^2(T)=\inf\{\xi^2/\eta:\ \xi+i\eta\in\Omega(S)\}$, i.e., $\mu_1^2(T)$ is equal to the inf of values of c such that the parabola $\mathcal{P}_c$: $\eta=\xi^2/c$ has at least one common point with $\Omega(S)$. By the conditions of the Theorem, $\Omega(S)$ is a bounded convex set in the positive quadrant of the $\xi O\eta$ plane (see property 1 and Figure 2). Let line l satisfy the conditions of property 4. It is easy to check that l is tangent to the parabola $\mathcal{P}_{c=4t\lambda_t}$ and separates it from $\Omega(S)$. So

$$\mu_1^2(T)\geqslant 4t\lambda_t. \tag{4}$$

We find that $[2\lambda_t,\lambda_t/t]$ is the point of tangency of l with ρ_c.

In Figure 3 are indicated the possible positions of line l in relation to $\Omega(S)$. The bold line represents a part of $L=\smile\zeta_1\zeta_2$ on the northwest boundary of $\Omega(S)$. The points $\zeta_1=[\xi_1;\eta_1]$ and $\zeta_2=[\xi_2;\eta_2]$ are such that the lines $\xi=\xi_1$ and $O\zeta_2$ are support lines of $\Omega(S)$. It follows from $\operatorname{Re}T>0$ that $0<\xi_1<\xi_2$. We must consider the support lines l tangent to $\Omega(S)$ at points $\zeta=[\xi;\eta]\in L$, making an angle ϕ with the axis $O\xi$ at λ_t, with $t=\cot\phi$. From the convexity of $\Omega(S)$, we see that as t increases from 0 to ξ_2/η_2, λ_t must decrease from ξ_1 to 0, and ξ must increase from ξ_1 to ξ_2. Hence the ratio $2\lambda_t/\xi$ decreases (strictly) monotonically from 2 to 0. Thus there is only one point $[\xi_0;\eta_0]\in L$

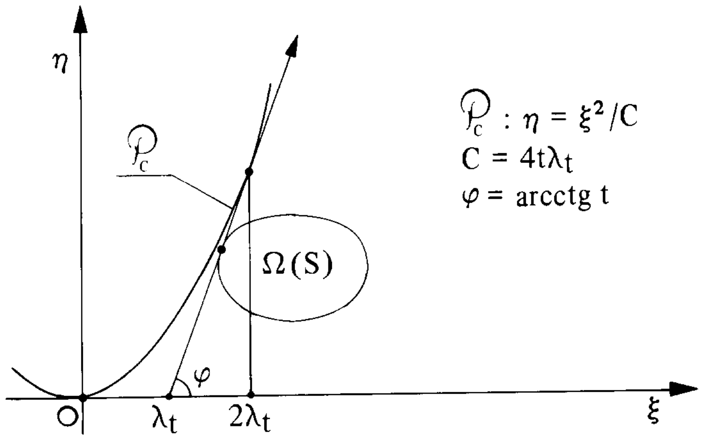

η
$\wp_c$
$\wp_c$: $\eta = \xi^2/C$
$C = 4t\lambda_t$
φ = arcctg t
Ω(S)
φ
O
λ_t
$2\lambda_t$
ξ

FIG. 2.

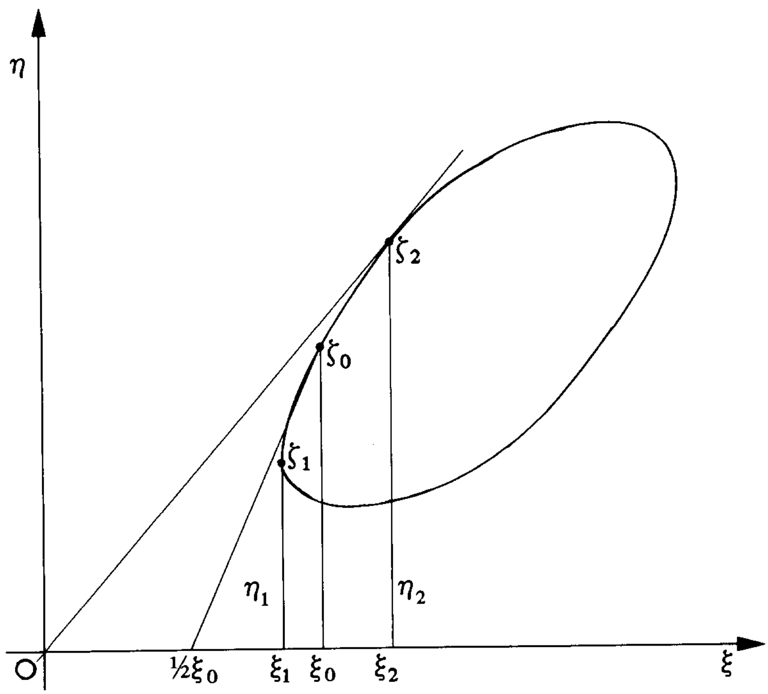

η
ζ_2
ζ_0
ζ_1
η_1
η_2
O
$\frac{1}{2}\xi_0$
ξ_1
ξ_0
ξ_2
ξ

FIG. 3.

such that the tangent l_0 passing through this point has intercepted $\xi_0/2$ on the ξ axis. Hence (4) turns into equality if $t = \cot\phi_0$, where ϕ_0 is the angle between this l_0 and $O\xi$.

The theorem is proved. ■

Note. If $\|T\| = \infty$, then also $\|S\| = \infty$ and $\Omega(S)$ is an unbounded convex set within the parabola $\mathcal{P}_{c=1}$ [because $\operatorname{Re}(Tx, x) \leqslant \|Tx\|$ for any unit x]. As may be easily seen, this means that $\Omega(S)$ contains a vertical ray. So for any $c > 0$, we have $\mathcal{P}_c \cap \Omega(S) \neq 0$, and

$$\mu_1(T) = 0$$

(K. Gustafson and B. Zwallen [5]).

COROLLARY 1. *Let T be normal and strictly accretive. Then*

$$\mu_1^2(T) = 4 \max_t \min_{\lambda \in \sigma(T)} \left(t \operatorname{Re}\lambda - t^2|\lambda|^2\right). \tag{5}$$

Indeed, if T is normal, then $\sigma(S_t) = \{\operatorname{Re}\lambda - t|\lambda|^2 \colon \lambda \in \sigma(T)\}$, so

$$\lambda_t = \min_{\lambda \in \sigma(T)} \left(\operatorname{Re}\lambda - t|\lambda|^2\right).$$

According to (5), if $\lambda_1, \lambda_2 \in \sigma(T)$, then $\mu_1^2(T) \geqslant 4\max_t \min_{1,2}(t\operatorname{Re}\lambda_j - t^2|\lambda_j|^2)$; this inequality leads to the inequalities of Ch. Davis [4]. These inequalities can also be deduced from Theorem 3 below.

COROLLARY 2 (L. V. Kantorovich [3]). *Let $T = T^* > 0$, and m and M be respectively the lower and the upper bounds of $\sigma(T)$. Then*

$$\mu_1(T) = \frac{2\sqrt{mM}}{m + M}.$$

Indeed, $S_t = T - tT^2$ in this case. So $\lambda_t = \min\,(m - tm^2;\, M - tM^2)$, and by Theorem 1, $\mu_1^2(T) = 4\max_t \min(tm(1 - tm),\ tM(1 - tM)) = 4mM/(m + M)^2$.

COROLLARY 3 (K. Gustafson [2]). *If $\operatorname{Re} T > 0$, then*

$$\gamma(T) = \inf_{\varepsilon > 0} \|\varepsilon T - I\| = \sqrt{1 - \mu_1^2(T)}$$

Indeed,

$$\gamma^2(T) = \inf_{\varepsilon>0} \sup_{\|x\|=1} ((\varepsilon T^* - I)(\varepsilon T - I)x, x)$$

$$= 1 + \inf_{\varepsilon>0}\left[-2\varepsilon \inf\left(\left(\operatorname{Re} T - \frac{\varepsilon}{2}T^*T\right)x, x\right)\right]$$

$$= 1 - 4\sup_{t>0}\left[t \inf_{\|x\|=1} ((\operatorname{Re} T - tT^*T)x, x)\right] = 1 - \mu_1^2(T).$$

The fact of attainment of the max in (3) can be expressed in terms of eigenvectors of the operator S_t. This will be done in Theorem 2 below.

Let $\dim H < \infty$ and $\delta_t = \{\xi: \xi = \operatorname{Re}(Tx, x), \|x\| = 1, S_t x = \lambda_t x\}$. δ_t is not empty, by the definition of λ_t, and is convex because $\Omega(S)$ is convex. Generally speaking, δ_t is some segment of $\{a_t, b_t\}$, which can degenerate into a point $a_t = b_t$. The condition $\dim H < \infty$ is not essential for the determination of δ_t. In general case $\dim H \leqslant \infty$,

$$\delta_t = \left\{\xi: \xi = \lim_n \operatorname{Re}(TX_n, X_n), \|X_n\| = 1, (S_t - \lambda_t)X_n \to 0\right\}.$$

THEOREM 2. *Let T be strictly accretive and $\delta_t = \{a_t, b_t\}$ be the above-determined segment for arbitrary $t \geqslant 0$. If $t = t_0$ is such that $a_{t_0} \leqslant 2\lambda_{t_0} \leqslant b_{t_0}$, then $\mu_1(T) = 2\sqrt{t_0\lambda_{t_0}}$. If t is such that $2\lambda_t < a_t$, then $t > t_0$. If t is such that $2\lambda_t > b_t$, then $t < t_0$.*

For the proof, we simply retrace the proof of Theorem 1, keeping track of the direction of the inequalities between the slopes of the support lines.

This theorem allows us to develop a computational procedure for the determination of $\mu_1(T)$ (and also of $\mu_k(T)$, $K > 1$). For example, taking $\{0, \tilde{t}\}$ as initial segment, t_0 can be found by the bisection method. Here $\tilde{t}$ is an arbitrary positive value such that $\lambda_{\tilde{t}} \leqslant 0$.

The inequality (4) gives lower bounds on $\mu_1(T)$. Upper bounds on $\mu_1(T)$ can be found by considering two-dimensional compressions of T, i.e. $S = \operatorname{Re} T + iT^*T$. According to 7^0, for two-dimensional T it is easy to find a point $[\xi_0 : \eta_0]$ of the ellipse $\Omega(S)$ such that the subtangent is equal to $\xi_0/2$.

Another way to find upper bounds on $\mu_1(T)$ is the following. Let $\mu(\Omega)$ denote $\inf\{\xi^2/\eta: [\xi; \eta] \in \Omega\}$ for Ω any bounded convex set in the positive quadrant of the $\xi O\eta$ plane. It is obvious that $\mu(\Omega_2) \leqq \mu(\Omega_1)$ follows from $\Omega_1 \leqslant \Omega_2$. Hence any simple subsets of $\overline{\Omega}(S)$ provide upper estimates of $\mu_1(T)$. To avoid huge calculations we consider here only the case of a segment in $\overline{\Omega}(S)$.

THEOREM 3. *Let T be strictly accretive, $S = Re\,T + iT^*T$, $[\xi_1; \eta_1]$, $[\xi_2; \eta_2] \in \bar{\Omega}(S)$, $\xi_1 \leqq \xi_2$; define*

$$\tilde{\xi} = \frac{\xi_1\eta_2 - \xi_2\eta_1}{\eta_2 - \eta_1}.$$

Then $\mu_1(T) \leqq \tilde{\mu}$, where

(a) *if $\eta_1 < \eta_2$, $\xi_1/2 < \tilde{\xi} < \xi_2/2$, then*

$$\tilde{\mu}^2 = 4\tilde{\xi}\frac{\xi_2 - \xi_1}{\eta_2 - \eta_1};$$

(b) $\tilde{\mu}^2 = \xi_1^2/\eta_1$ *if $\eta_1 \geqq \eta_2$ or if $\eta_1 < \eta_2$ and $\tilde{\xi} \leqq \xi_1/2$;*
(c) $\tilde{\mu}^2 = \xi_2^2/\eta_2$ *if $\eta_1 < \eta_2$ and $\tilde{\xi} \geqq \xi_2/2$.*

Indeed, as $[\xi, \eta]$ runs over the segment from $[\xi_1; \eta_1]$ to $[\xi_2; \eta_2]$, the inf of $\{c : [\xi; \eta] \in \mathcal{P}_c\}$ is attained at an endpoint except in case (a), where $\eta_1 < \eta_2$ and some $\mathcal{P}_c$ is tangent to the segment at an internal point (see Figure 4). The rest follows by elementary computations.

It is easy to find the structure of a set of antieigenvectors in the case $\dim H < \infty$.

THEOREM 4. *Let* $\dim H = n < \infty$ *and T be strictly accretive on H. Then:*

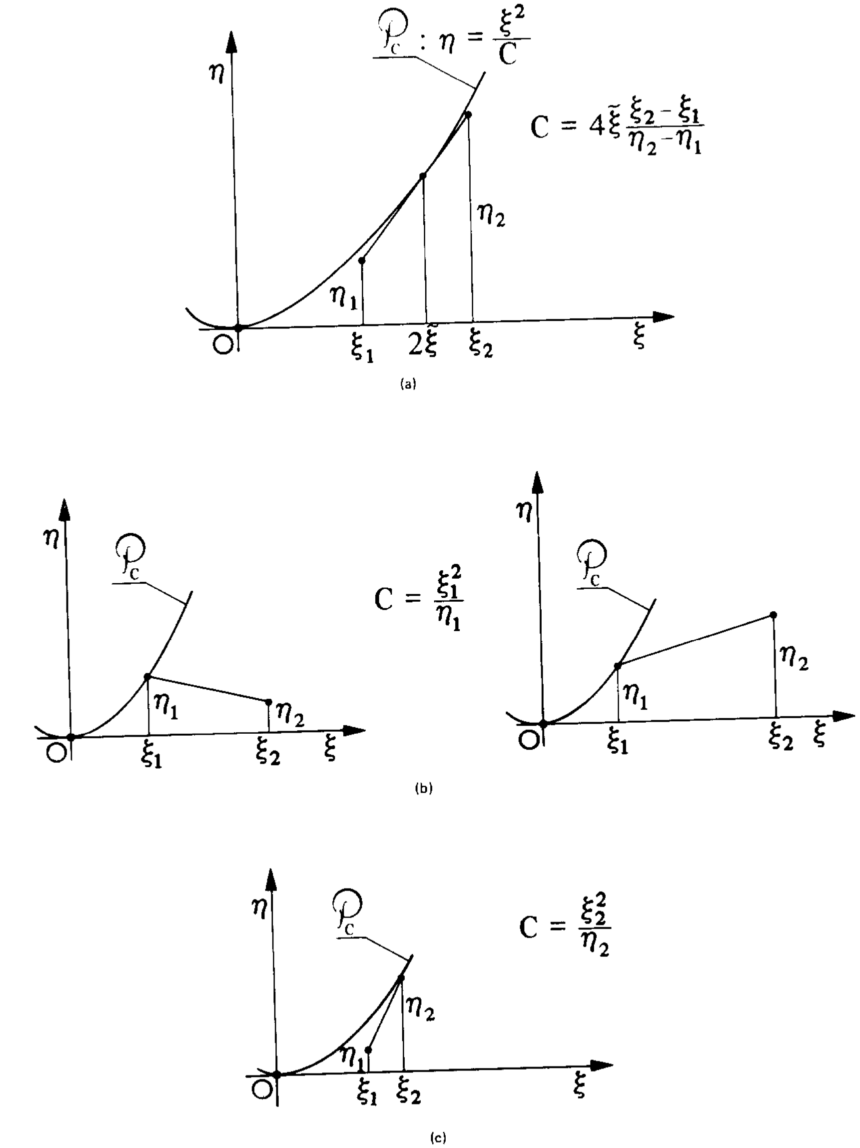
1. *It is possible to choose unit vectors $x^{(k)}$ ($k = 1, \dots, n$) such that they are antieigenvectors corresponding to antieigenvalues $\mu_k(T)$ and the system $\{x^{(k)}\}_1^n$ is an orthonormal basis in H.*

2. *The set of all antieigenvectors which correspond to one k-fold antieigenvalue ($\mu_1(T) = \cdots = \mu_k(T) < \mu_{k+1}(T)$) is a k-dimensional subspace if and only if the point $[\xi_0; \eta_0]$ in the proof of Theorem 1 is an extreme point of $\Omega(S)$.*

3. *Let $T = T^* > 0$. All antieigenvalues are equal ($\mu_1(T) = \cdots = \mu_n(T)$) if and only if it is possible to find an orthogonal projector P such that $T = mP + M(I - P)$, where $M \geqslant m > 0$.*

The proof of these statements follows from properties 5 and 6, of numerical ranges and Theorems 1, 2.

Since the proofs of Theorems 1, 2 use mainly one property of the curve $f(\xi, \eta) = \xi^2/\eta = c$, its convexity, one can in the same manner consider other families of convex curves $f(\xi, \eta) = c$. Using the same way of determining lower or upper bounds of $\sigma(A - tB)$, one can find extrema of expressions of



$\mathcal{P}_C : \eta = \frac{\xi^2}{C}$
$C = 4\tilde{\xi}\frac{\xi_2 - \xi_1}{\eta_2 - \eta_1}$
η
η_1
η_2
O
ξ_1
$2\tilde{\xi}$
ξ_2
ξ
(a)
$C = \frac{\xi_1^2}{\eta_1}$
$\mathcal{P}_C$
(b)
$C = \frac{\xi_2^2}{\eta_2}$
(c)

FIG. 4.

the form $\{f((Ax, x),(Bx, x)):\|x\|=1\}$ where $A=A^*$, $B=B^*$. If A or B is unbounded, the absence of a common asymptote of the curves $f(\xi, \eta)=c$ is additionally required.

I would like to use this occasion to express my gratitude to Chandler Davis, whose contribution was vital to this work.

REFERENCES

1 M. G. Krein, Angular localization of the spectrum of a multiplicative integral in a Hilbert space, *Funkcional. Anal. i Prilozhen.* 3(1):89–90 (1969).
2 K. Gustafson, Antieigenvalue inequalities in operator theory, in *Inequalities*, Vol. III, pp. 115–119, Academic, New York, 1972.
3 L. V. Kantorovich, *Uspehi Mat. Nauk*, 3(6):89–185 (1978).
4 Ch. Davis, Extending the Kantorovic inequality to normal matrices, *Linear Algebra Appl.* 31:173–177 (1980).
5 K. Gustafson and B. Zwallen, On the cosine of unbounded operators, *Acta Sci. Math. (Szeged)* 30:33–34 (1969).
6 K. Gustafson and D. Rao, Numerical range and accretivity of operator products, *J. Math. Anal. Appl.* 60:693–702 (1977).

Received 1 July 1981; revised 11 May 1982