# An Effective Algorithm for Computing the Numerical Range

Carl C. Cowen*
Elad Harel

August 1995

**Abstract**

In this paper, we collect a few fairly well known facts about the numerical range and assemble them into an effective algorithm for computing the numerical range of an $n \times n$ matrix. Particular attention is paid to the case in which a line segment is embedded in the boundary of the numerical range, a case in which multiplicity is present. The result is a parametrization of the curve that forms the boundary of the numerical range. A Matlab implementation of the algorithm is included.

If $A$ is an $n \times n$ complex matrix, the *numerical range of* $A$ is the subset of the complex plane given by

$$w(A) = \{\langle Av, v\rangle : v \in \mathbf{C}^n \quad \text{with} \quad \|v\| = 1\}$$

The classical Toeplitz–Hausdorff Theorem [2] asserts that the numerical range of every $n \times n$ matrix is a convex set. One approach to the proof of this theorem is to show by direct computation that the numerical range of every $2 \times 2$ matrix is an ellipse with its interior and observe that if $\langle Au, u\rangle$ and $\langle Av, v\rangle$ are two points of the numerical range of $A$, then the numerical range of the compression of $A$ to the subspace spanned by $u$ and $v$ is an ellipse that contains these two points and is contained in the numerical range of $A$. Thus, since the ellipse is a convex set, the line segment joining $\langle Au, u\rangle$

*Supported in part by National Science Foundation Grants DMS 92-06965 and DMS 95-00870.

and $\langle Av, v\rangle$ is in the numerical range of the compression and, hence, in the numerical range of $A$. This approach was applied to the computation of the numerical range by Marcus and Pesce [5] who showed that the numerical range of $A$ is the union of the numerical ranges of all the compressions to two dimensional real subspaces and computed some numerical ranges from that fact. Our approach more closely resembles the approach of Li, Sung, and Tsing [4] in their computation of $c$–numerical ranges.

**Lemma 1** *If $H$ is Hermitian, $w(H) = \{t : \lambda_1 \le t \le \lambda_n\}$ where $\lambda_1$ and $\lambda_n$ are the least and greatest eigenvalues of $H$ respectively. Moreover, if $\|v\| = 1$ and $\langle Hv, v\rangle = \lambda_n$, then $v$ is an eigenvector for the eigenvalue $\lambda_n$.*

**Proof.** Suppose $v_1, v_2, \cdots, v_n$ is an orthonormal basis for $\mathbf{C}^n$ consisting of eigenvectors of $H$ corresponding to eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$ which are arranged in increasing order.

If $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ is a vector with norm 1 so that $|\alpha_1|^2 + |\alpha_2|^2 + \cdots + |\alpha_n|^2 = 1$, then, since the eigenvalues of a Hermitian matrix are real,

$$\begin{aligned} \langle Hv, v\rangle &= \lambda_1|\alpha_1|^2 + \lambda_2|\alpha_2|^2 + \cdots + \lambda_n|\alpha_n|^2 \\ &\le \lambda_n|\alpha_1|^2 + \lambda_n|\alpha_2|^2 + \cdots + \lambda_n|\alpha_n|^2 \\ &= \lambda_n \end{aligned}$$

Note that if $\lambda_j < \lambda_n$ and $\alpha_j \ne 0$, then the inequality above is strict, so $\langle Hv, v\rangle = \lambda_n$ if and only if $v$ is a linear combination of eigenvectors of $H$ whose eigenvalue is $\lambda_n$, that is, if and only if $v$ is an eigenvector of $H$ with eigenvalue $\lambda_n$. This proves the second conclusion.

Replacing each $\lambda_j$ by $\lambda_1$ gives the inequality $\langle Hv, v\rangle \ge \lambda_1$ for all $v$ with $\|v\| = 1$. Finally, letting $v_s = \sqrt{s}v_1 + \sqrt{1-s}v_n$ for $0 \le s \le 1$ gives $\|v_s\| = 1$ and $\langle Hv_s, v_s\rangle = s\lambda_1 + (1-s)\lambda_n$. Thus, $w(H) = [\lambda_1, \lambda_n]$ ■

We need to investigate the situation in which the numerical range lies to the left of a line $\mathrm{Re}(z) = \mu$ and intersects the line. To do so, we use the canonical decomposition of a matrix into its Hermitian and skew–Hermitian parts: $A = H + iK$ where $H = H^* = (A + A^*)/2$ and $K = K^* = (A - A^*)/(2i)$.

**Lemma 2** *Suppose $A$ is an $n \times n$ matrix and $A = H + iK$ where $H = H^*$ and $K = K^*$. If $\mu$ is a real number such that $Re(\langle Av, v\rangle) \le \mu$ for every $v$ in*

$\mathbf{C}^n$ *with* $\|v\| = 1$*, then either* $w(A)$ *does not intersect the line* $Re(z) = \mu$ *or* $w(A) \cap \{z : Re(z) = \mu\}$ *is the point or line segment* $\mu + iw(PKP)$ *where* $P$ *is the orthogonal projection of* $\mathbf{C}^n$ *onto the eigenspace* $\{u \in \mathbf{C}^n : Hu = \mu u\}$*.*

**Proof.** If $v$ is in $\mathbf{C}^n$ and $\|v\| = 1$, then

$$\langle Av, v\rangle = \langle (H + iK)v, v\rangle = \langle Hv, v\rangle + i\langle Kv, v\rangle$$

Since the numerical ranges of $H$ and $K$ are both real, this means that if $w$ is in the numerical range of $A$, then the real part of $w$ is in the numerical range of $H$ and the imaginary part of $w$ is in the numerical range of $K$. In particular, if $\mu$ is the maximum of $\{\mathrm{Re}(\langle Av, v\rangle) : v \in \mathbf{C}^n \text{ with } \|v\| = 1\}$, then $\mu$ is greatest number in the numerical range of $H$, that is, $\mu = \lambda_n$, the largest eigenvalue of $H$. Moreover, Lemma 1 implies that if $v$ is a vector in $\mathbf{C}^n$ with $\|v\| = 1$ and $\mu = \mathrm{Re}(\langle Av, v\rangle) = \langle Hv, v\rangle$, then $Hv = \mu v$. That is,

$$w(A) \cap \{z : \mathrm{Re}(z) = \mu\} = \{\mu + i\langle Ku, u\rangle : Hu = \mu u\}$$

Now since $P$ is the orthogonal projection of $\mathbf{C}^n$ onto the eigenspace $\{u \in \mathbf{C}^n : Hu = \mu u\}$, $Pu = u$ for all such $u$ and $\langle Ku, u\rangle = \langle KPu, Pu\rangle = \langle PKPu, u\rangle$. Since $PKP$ is Hermitian, Lemma 1 shows its numerical range is a point or line segment. ■

We can put these results to use in describing the boundary of the numerical range and then in constructing an algorithm for computing the numerical range. The idea of the theorem is to find the points on the boundary of the numerical range that touch lines in each direction by observing that the numerical range of a multiple of a matrix is that multiple of the numerical range. To do this, we rotate the matrix and use Lemma 2 to find the point or segment in the boundary so that $\mathrm{Re}(\langle Av, v\rangle) \leq \mu$ for every $v$ in $\mathbf{C}^n$ with $\|v\| = 1$ and $w(A) \cap \{z : \mathrm{Re}(z) = \mu\}$ is non-empty.

**Theorem 3** *Let* $A$ *be an* $n \times n$ *matrix. For* $0 \leq t \leq 2\pi$*, let* $H_t$ *and* $K_t$ *be Hermitian matrices so that* $e^{-it}A = H_t + iK_t$ *and let* $P_t$ *be the projection of* $\mathbf{C}^n$ *onto the eigenspace of* $H_t$ *corresponding to the largest eigenvalue of* $H_t$*. Let* $v_t^+$ *and* $v_t^-$ *be eigenvectors of* $H_t$ *with* $\|v_t^+\| = \|v_t^-\| = 1$ *corresponding to the largest eigenvalue of* $H_t$ *such that* $v_t^+$ *and* $v_t^-$ *are eigenvectors of* $P_tK_tP_t$ *corresponding to the greatest and smallest eigenvalues, respectively, of* $P_tK_tP_t$*. Then for each* $t$*, the numbers* $\langle Av_t^+, v_t^+\rangle$ *and* $\langle Av_t^-, v_t^-\rangle$ *are in the boundary of the numerical range of* $A$ *and* $w(A)$ *is the convex hull of these numbers.*

**Proof.** The proof of Lemma 2 shows that

$$\mathrm{Re}\left(\langle e^{-it}Av, v\rangle\right) \le \mu_t$$

where $\|v\| = 1$ and $\mu_t$ is the largest eigenvalue of $H_t$. In particular, if $v$ is a point of $\mathbf{C}^n$ with $\|v\| = 1$, then $\mathrm{Re}(\langle e^{-it}Av, v\rangle) \le \mu_t$. Since there is $v$ with $\|v\| = 1$ and $\mathrm{Re}(\langle e^{-it}Av, v\rangle) = \mu_t$, any point of the numerical range $\langle Av, v\rangle$ must be on the boundary of the numerical range.

Since the numerical range is a convex set and each of the numbers above is in the numerical range, the convex hull of these numbers must be in the numerical range. Conversely, if $w$ is in the numerical range of $A$ but $w$ is not in the convex hull of the points in statement of the theorem, then there is a line separating $w$ from this convex hull. Suppose $t$ is such that $e^{-it}$ times the line is vertical and the convex hull of the numbers described in the theorem is to the left of the line. Then the choice of $w$ gives $\mathrm{Re}(e^{-it}w) > \mu_t$ which is impossible. Moreover, by Lemma 2, each point of the numerical range with $\mathrm{Re}(e^{-it}w) = \mu_t$ is on the line segment joining $\langle Av_t^+, v_t^+\rangle$ and $\langle Av_t^-, v_t^-\rangle$. Thus, every point of the numerical range is a convex combination of the points mentioned in the theorem. ■

Our algorithm for calculating the numerical range follows the construction given above. We calculate $\langle Av_t^-, v_t^-\rangle$ and $\langle Av_t^+, v_t^+\rangle$ for $t$ in $[0, 2\pi]$. Of course, for most $t$, the eigenspace of $H_t$ corresponding to the largest eigenvalue is one dimensional, so we can usually take $v_t^+ = v_t^-$. The boundary curve is now parametrized by $t$ with the understanding that the segment joining $\langle Av_t^-, v_t^-\rangle$ and $\langle Av_t^+, v_t^+\rangle$ is inserted at the appropriate point whenever these points are different.

If the eigenspace of $H_t$ corresponding to the largest eigenvalue has dimension $j$ with $j > 1$, let $u_1$, $u_2$, $\cdots$, $u_j$ be an orthonormal basis for the eigenspace. Letting $Q$ be the $n \times j$ matrix whose columns are $u_1$, $u_2$, $\cdots$, $u_j$, then $Q^*K_tQ$ is the matrix for the restriction of $P_tK_tP_t$ to the eigenspace with respect to this basis. Therefore, if $(a_1^+, a_2^+, \cdots, a_j^+)$ and $(a_1^-, a_2^-, \cdots, a_j^-)$ are eigenvectors of this matrix corresponding to the greatest and least eigenvalues, then we may take $v_t^+ = a_1^+u_1 + a_2^+u_2 + \cdots + a_j^+u_j$ and $v_t^- = a_1^-u_1 + a_2^-u_2 + \cdots + a_j^-u_j$.

Below is an implementation of these ideas in Matlab. The script finds 630 points on the boundary of the numerical range of an $n \times n$ matrix (asked for as input) corresponding to rotates by $e^{-i\theta}$ where $\theta$ ranges from 0 to $2\pi$

incremented by .01 radians and fills the resulting polygon in the complex plane: this is our approximation of the numerical range.

```
%
%  This script finds the numerical range of an n x n matrix by
%   finding the real and imaginary parts of rotates of the matrix
%   and finding the associated boundary point of that rotate by
%   finding the largest eigenvalue of the real part and using the
%   corresponding eigenvector's contribution to the numerical range.
%   Multiplicity of the largest eigenvalue, as occurs in a normal
%   matrix, is handled by plotting the end points of the corresponding
%   segment in the boundary of the numerical range.
%
A=input('For what matrix do you want the numerical range? ')
nm=ceil(norm(A));
th=[0:.01:6.29];
k=1;
w=zeros(1,630);
for j=1:630
    Ath=(exp(i*(-th(j))))*A;
    Hth=(Ath+Ath')/2;
    [r e]=eig(Hth);
    e=real(diag(e));
    m=max(e);
s=find(e==m);
 if size(s,1)==1
    w(k)=r(:,s)'*A*r(:,s);
  %
  %  This is the point of the numerical range contributed by
  %   v_t=r(:,s) when the eigenspace of Hth (H_t) is one dimensional.
  %
else
 Kth=i*(Hth-Ath);
  pKp=r(:,s)'*Kth*r(:,s);
  %
  %  The matrix Q described above is  r(:,s)
  %
[rr ee]=eig(pKp);
ee=real(diag(ee));
```

```
mm=min(ee);
sm=find(ee==mm);
w(k)=rr(:,sm(:,1))'*r(:,s)'*A*r(:,s)*rr(:,sm(:,1));
  %
  %  This is the point of the numerical range contributed by
  %    v_t^- = r(:,s)*rr(:,sm(:,1))
  %
k=k+1;
mM=max(ee);
sM=find(ee==mM);
w(k)=rr(:,sM(:,1))'*r(:,s)'*A*r(:,s)*rr(:,sM(:,1));
  %
  %  This is the point of the numerical range contributed by v_t^+
  %
 end
k=k+1;
end
fill(real(w),imag(w),'y')
axis([-nm,nm,-nm,nm])
axis('equal')
```

## References

[1] D. S. Keeler, L. Rodman, and I. M. Spitkovsky, Numerical ranges of $3 \times 3$ matrices, preprint, 1994.

[2] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[3] R. Kippenhahn, Über den Wertevorrat einer Matrix, *Math. Nachr.* **6**(1951), 193–228.

[4] C. K. Li, C. H. Sung, and N. K. Tsing $c$–convex matrices: characterizations, inclusion relations, and normality, *Linear and Multilinear Alg.* **25**(1989), 275–287.

[5] M. Marcus and C. Pesce, Computer generated numerical ranges and some resulting theorems, *Linear and Multilinear Alg.* **20**(1987), 121–157.

[6] H. Shapiro, On a conjecture of Kippenhahn about the characteristic polynomial of a pencil generated by two Hermitian matrices, I, *Linear Alg. Appl.* **43**(1982), 201–221.

[7] H. Shapiro, On a conjecture of Kippenhahn about the characteristic polynomial of a pencil generated by two Hermitian matrices, II, *Linear Alg. Appl.* **45**(1982), 97–108.

[8] H. Shapiro, Hermitian pencils with a cubic minimal polynomial, *Linear Alg. Appl.* **48**(1982), 81–103.

Department of Mathematics
Purdue University
West Lafayette, Indiana 47907-1395
e–mail: cowen@math.purdue.edu
WWW: http://www.math.purdue.edu/~cowen