NORMAL MATRICES: AN UPDATE

L. Elsner
Fakultät für Mathematik
Universität Bielefeld
Postfach 100131

33501 Bielefeld
FRG

Kh. D. Ikramov[1]
Faculty of Computational Mathematics
and Cybernetics
Moscow State University

119899 Moscow
Russia

Abstract. A list of seventy conditions on an n-by-n complex matrix A, equivalent to its being normal, published nearly ten years ago by Grone, Johnson, Sa, and Wolkowicz has proved to be very useful. Hoping that, in an extended form, it will be even more helpful, we compile here another list of about twenty conditions. They either have been overlooked by the authors of the original list or have appeared during the last decade.

[1]Research supported by Sonderforschungsbereich 343 *Diskrete Strukturen in der Mathematik*

1. Introduction.

A square complex matrix A is called normal if

$$(1) \qquad A^*A = AA^*$$

where A^* is the adjoint of A. If this relation is taken as condition 0, then the seventy conditions 1–70 equivalent to (1) constitute Section II of [6].

At the end of the introduction to their list, the authors of [6] say: "Reflecting the fact that normality arises in many ways, it is hoped that not only will it be useful now, but its utility will grow over time as conditions are added."

Nearly a decade after [6] was published, one can say that it (called henceforth the GJSW list) has indeed proved to be very useful. It is perhaps the right time to increase its utility, as the authors suggest, by supplementing the original list with additional conditions equivalent to the normality of A.

Section 2 of this paper contains nearly twenty such conditions. They either have been overlooked by the authors of [6] or have appeared during the last decade. Since our list is meant to be a continuation of that in [6], our first condition is numbered 71. The outline of proofs and/or comments to most of the conditions are given in Section 3. In nearly all cases, we point to a publication where the corresponding condition has first appeared. Our proofs are, up to minor details, the original proofs in these publications. As in [6], proofs in an obvious direction are omitted.

We conclude this section by introducing the notation, which does not in all cases coincide with that of [6]. Also, those conditions in the original GJSW list are given explicitly which are used in the body of our paper.

Our matrices belong (almost) invariably to the set $\mathbb{C}^{n,n}$ of complex $n \times n$–matrices. It is sometimes beneficial to regard $\mathbb{C}^{n,n}$ as a unitary space equipped with the inner product

$$(2) \qquad < A, B >= \text{tr } (B^*A)$$

where $\text{tr}(\cdot)$ stands for the trace of a matrix.

$\mathbb{C}^n$ denotes the set of all complex column vectors with n components. If necessary, $\mathbb{C}^n$ is also considered as a unitary space with the usual inner product $(\ ,\ )$. The symbol $||\cdot||$ is used for the 2–norm of a vector and, if not stated otherwise, for the spectral norm of a matrix. For A nonsingular, the condition number

$$(3) \qquad \text{cond } A = ||A||\ ||A^{-1}||$$

is always meant to be the spectral one. If λ is a simple eigenvalue of A, x and y corresponding unit right and left eigenvectors, then

$$(4) \qquad \kappa(\lambda) = |(x, y)|^{-1}$$

is called the condition number of λ.

We denote by $\rho(A)$ and $\sigma_1(A) \geq \ldots \geq \sigma_n(A)$ the spectral radius of A and its singular values. The commutator $[B, C] = BC - CB$ is sometimes called the self–commutator of B, if $C = B^*$. The matrices

$$|A| = (A^* A)^{1/2}, \quad |A^*| = (AA^*)^{1/2} \tag{5}$$

are called the right and left absolute value of A, respectively.

In addition to well–known matrix decompositions, such as the QR or the singular value decomposition, we also use some that are encountered less frequently. These are the Toeplitz decomposition

$$A = H + iK, \quad H = H^*, \quad K = K^* \tag{6}$$

with

$$H = \frac{1}{2}(A + A^*), \quad K = \frac{1}{2i}(A - A^*) \tag{7}$$

and the polar decompositions

$$A = U_1 \, |A| = |A^*| U_2 \tag{8}$$

with unitary matrices U_1 and U_2. Note that, for a nonsingular A, U_1 and U_2 are defined uniquely and are, in fact, the same unitary matrix. If A is singular then there exist (infinitely) many U_1, U_2 satisfying (8). However, one can choose them to be the same matrix in this case as well [11, p. 152].

For our purposes, matrices B and C are said to be simultaneously diagonalizable, if

$$B = U \, \Sigma_1 V^*, \quad C = U \, \Sigma_2 V^* \tag{9}$$

with unitary U, V and diagonal Σ_1, Σ_2. Note that, as opposed to the singular value decomposition, Σ_1 and Σ_2 are not required to be nonnegative.

The symbol $C_k(A)$ stands for the k–th compound matrix of A [14, p. 16].

Now we list some conditions from [6]:

11. There exist unitary U and diagonal D such that $A = UDU^*$.
14. There exist an orthonormal basis of $\mathbb{C}^n$ consisting of eigenvectors of A.
37. If $A = U|A|$ is the (right) polar decomposition of A, then U and $|A|$ commute.
38. A and $|A|$ commute.
58. The singular values $\sigma_1 \geq \ldots \geq \sigma_n$ of A coincide with the moduli of its eigenvalues.
59. If the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A are numbered according to

    $$|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|, \tag{10}$$

    then

    $$\sigma_1 \ldots \sigma_k = |\lambda_1 \ldots \lambda_k|, \quad k = 1, \ldots, n.$$

2. Conditions.

71 [3]. $|A| = |A^*|$.

72 [3]. A is diagonalizable:

$$A = X\Lambda X^{-1}, \quad \Lambda = \text{ diag } (\lambda_1, \ldots, \lambda_n) \tag{11}$$

with X such that cond $X = 1$. (Or: for any eigenvalue λ of A, the condition number $\kappa(\lambda) = 1$, as long as A has distinct eigenvalues).

73 [14]. A commutes with $[A, A^*]$.

74 [8]. A commutes with A^*A (or AA^*).

75 [9]. In the Toeplitz decomposition of A,

$$H^2 + K^2 = A^*A \ \ (\text{or } AA^*).$$

76 [16]. $\text{tr } (A^{*2}A^2) = \text{tr } [(A^*A)^2]$.

77 [16]. (Generalization of condition 76)

$$\text{tr } (A^{*p}A^p) = \text{tr } [(A^*A)^p] \ \text{ for some positive integer } \ p \geq 2.$$

78 [16]. (Generalization of condition 77)

$$\text{tr } [(A^{*p}A^p)^q] = \text{tr } [(A^*A)^{pq}] \ \text{ for some positive integers } \ p \geq 2, \ q \geq 1.$$

79. $e^{t_n A}$ is normal for a sequence $\{t_n\}$, converging to zero.

80 [16]. $\text{tr } (e^{A^*}e^A) = \text{tr } (e^{A^*+A})$.

81 [1,5]. The function $f_x(t) = \ln \ ||e^{tA}x||$ is convex on $\mathbb{R}$ for any vector $x \in \mathbb{C}^n$.

82 [1]. The functions $f_j(t) = \ln \ \sigma_j \ (e^{tA}), \ j = 1, \ldots, n$, are convex on $\mathbb{R}$.

83 [7]. A is diagonalizable as in (11), and a nonzero perturbation B exists such that for some eigenvalue μ_0 of the perturbed matrix $A + B$ the closest eigenvalue λ_0 of A is uniquely defined, and

$$|\lambda_0 - \mu_0| = \ ||B|| \text{ cond } X. \tag{12}$$

84 [12]. $|(Au, u)| \leq \ (|A|u, u) \quad \forall \ u \in \mathbb{C}^n$.

85 [12]. $|(Au, u)| \leq \ (|A^*|u, u) \quad \forall \ u \in \mathbb{C}^n$.

86 [12]. (Generalization of conditions 84 and 85)

$$|(Au, u)| \leq (|A|u, u)^\alpha \ (|A^*|u, u)^{1-\alpha} \quad \forall \ u \in \mathbb{C}^n \tag{13}$$

for some real $\alpha \neq \frac{1}{2}$.

87 [13]. If $A = QR$ is the QR–decomposition of A then Q^* and R are simultaneously diagonalizable.

88. For $k = 1, 2, \ldots, n-1$

$$||C_k(A)|| = \rho(C_k(A)). \tag{14}$$

89. The matrix Lyapunov operator on $\mathbb{C}^{n,n}$ defined by the formula

$$\mathcal{L}_A X = AX + XA^* \quad \forall \ X \in \mathbb{C}^{n,n} \tag{15}$$

is normal.

3. Proofs and comments.

Condition 71 is immediate from (1) and (5), and condition 72 from condition 11, if one takes into account that matrices with the spectral condition number 1 are scalar multiples of unitary matrices. The second version of condition 72 implies that A has a full orthonormal set of eigenvectors, hence condition 14 is applicable.

For the sufficiency part of condition 73, we first mention that under the unitary similarity

$$A \to U^*AU$$

the self–commutator $C = [A, A^*]$ transforms in the same way as A does:

$$C \to U^*CU.$$

Suppose C is nonzero. Then no generality is lost in assuming C to be in block diagonal form

$$C = C_{11} \oplus \ldots \oplus C_{mm} \tag{16}$$

with diagonal blocks $C_{11}, \ldots, C_{mm}$ corresponding to the distinct eigenvalues $\lambda_1(C), \ldots, \lambda_m(C)$ of C. Obviously, $m \geq 2$ because $\mathrm{tr}\ C = 0$. Since A and C commute, the former must assume the same block diagonal form as the latter:

$$A = A_{11} \oplus \ldots \oplus A_{mm}.$$

Then

$$C_{ii} = [A_{ii}, A_{ii}^*], \quad i = 1, \ldots, m,$$

which implies

$$\mathrm{tr}\ C_{ii} = 0, \quad i = 1, \ldots, m,$$

and

$$\lambda_i(C) = 0, \quad i = 1, \ldots, m.$$

This is a contradiction to the way by which the decomposition (16) has been introduced.

Condition 74 insures that A commutes with any polynomial of A^*A. Since $|A|$ is one of those polynomials, condition 38 may be employed.

Condition 75 is an immediate consequence of the identity

$$H^2 + K^2 = \frac{1}{2}\,(AA^* + A^*A)$$

which holds for any square A.

To prove the sufficiency of condition 76, one can use the identity

$$||A^*A - AA^*||_F^2 \;=\; 2\{\mathrm{tr}\ [(A^*A)^2] - \mathrm{tr}\ (A^{*2}A^2)\}$$

(this is Lemma 4.3 in [16]) that is again valid for any square A. The symbol $||\cdot||_F$ stands for the Frobenius norm of a matrix.

Assume that $p > 2$ in condition 77 (otherwise, it coincides with condition 76). The idea is to show that condition 77 with the exponent p implies the same condition with the exponent $p-1$. Then, in a finite number of steps, we arrive at condition 76 assuring the normality of A.

Let r be the rank of A, then, among its singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$, only the first r are nonzero. Letting $H = |A|$, we see that $\sigma_1, \ldots, \sigma_n$ are the eigenvalues of H. Denote the corresponding orthonormal eigenvectors by $v_1, \ldots, v_n$, and introduce

$$S(p) = \mathrm{tr}\ [(A^*A)^p] - \mathrm{tr}\ (A^{*p}A^p) \tag{17}$$

and

$$D(t,s) = \sum_{i=1}^{t} (\sigma_i^2 - \sigma_t^2)\ (\sigma_i^{2s} - ||A^sUv_i||^2).$$

Here U is the unitary factor in the polar decomposition of A.

It follows from [4] that

$$S(p) \geq 0 \ \text{ for any positive integer } p \tag{18}$$

and

$$D(t+1,s) \geq D(t,s) \ \text{ for positive integers } \ t, s (t \leq n).$$

Hence

$$D(t+1,s) \geq D(t,s) \geq D(t-1,s) \geq \ldots \geq D(1,s) = 0,$$

and $D(t,s)$ is always nonnegative. Now one can prove the inequality (this is essentially Theorem 4.4 in [16])

$$S(p) \geq \sigma_r^2\ S(p-1) + D(r,p-1).$$

We conclude that $S(p) = 0$ forces $S(p-1) = 0$.

Turning to condition 78, assume that $q > 1$ (the case $q = 1$ gives us condition 77). Denote by $\mathrm{tr}_i(H)$ the sum of the i largest eigenvalues of an Hermitian matrix H. Then an inequality stronger than (17) – (18) is proved in [4], namely

$$\mathrm{tr}_i(A^{*p}A^p) \leq \mathrm{tr}_i[(A^*A)^p] \ \text{ for any positive integer } \ p \ \text{ and } \ i = 1, 2, \ldots, n. \tag{19}$$

Inequalities (19) can be interpreted as the weak majorization relations between the vectors

$$(\sigma_1^2(A^p),\ \sigma_2^2(A^p), \ldots, \sigma_n^2(A^p))$$

and

$$(\sigma_1^{2p}, \sigma_2^{2p}, \ldots, \sigma_n^{2p}).$$

Then, using the usual majorization techniques (see, for example, [15, p. 115]), one can easily show that

$$\mathrm{tr}[(A^{*p}A^p)^q] \leq \mathrm{tr}[(A^*A)^{pq}] \tag{20}$$

for any matrix A and positive integers p, q. Moreover, equality in (20) implies that

$$\mathrm{tr}(A^{*p}A^p) = \mathrm{tr}[(A^*A)^p] \tag{21}$$

(this is Theorem 4.2 in [16]). In other words, condition 78 with $q > 1$ implies condition 77.

For condition 79, the sufficiency is obvious from the relation

$$A = \lim_{t_n \to 0} (e^{t_n A} - I)/t_n.$$

The proof of condition 80 relies on the following product exponential formula which is valid for any $n \times n$ matrices X and Y:

$$\lim_{m\to\infty} (e^{\frac{X}{m}}\, e^{\frac{Y}{m}})^m = e^{X+Y}. \tag{22}$$

This formula is proved in [2]. As a consequence of (22),

$$\lim_{m\to\infty} \mathrm{tr}[(e^{\frac{X}{m}}\, e^{\frac{Y}{m}})^m] = \mathrm{tr}(e^{X+Y}).$$

Using repeatedly (20) with $p = 2,\ q = 1, 2, 3, \ldots$ and A replaced by $e^{A/2^q}$, one can show that, for any square matrix A, the sequence

$$\mathrm{tr}\Big[\Big(e^{A^*/2^q}\, e^{A/2^q}\Big)^{2^q}\Big]$$

is monotonically increasing to $\mathrm{tr}(e^{A^*+A})$ (Theorem 4.6 in [16]).

Now condition 80 implies that

$$\mathrm{tr}(e^{A^*}e^A) = \mathrm{tr}\Big[\Big(e^{A^*/2^q}\, e^{A/2^q}\Big)^{2^q}\Big]$$

for all positive integers q. This is nothing else than equality (21) with $p = 2^q$ and A replaced by $B_q = e^{A/2^q}$. According to condition 77, B_q must be normal. It remains to apply condition 79.

For condition 81, the necessity part consists of checking, for any $t_1, t_2 \in \mathbb{R},\ t_1 < t_2$, the inequality

$$f_x\Big(\frac{t_1+t_2}{2}\Big) \leq \frac{f_x(t_1)+f_x(t_2)}{2} \tag{23}$$

that amounts to

$$(24) \qquad ||e^{\frac{t_1+t_2}{2}A}\ x||^2 \ \le\ ||e^{t_1A}x||\ ||e^{t_2A}x||.$$

Without loss of generality, one can assume the normal matrix A in (24) to be diagonal: $A = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Letting

$$x = (x_1, \dots, x_n)^T,$$

we then rewrite (24) as

$$(25) \qquad \sum_{i=1}^{n} \left|e^{\frac{t_1+t_2}{2}\lambda_i}\ x_i\right|^2 \ \le\ \Big(\sum_{i=1}^{n}\left|e^{t_1\lambda_i}\ x_i\right|^2\Big)^{1/2}\ \Big(\sum_{i=1}^{n}\left|e^{t_2\lambda_i}\ x_i\right|^2\Big)^{1/2}.$$

Since

$$|e^{\frac{t_1+t_2}{2}\lambda_i}\ x_i|^2 = |e^{(t_1+t_2)\lambda_i}\ x_i^2| = |e^{t_1\lambda_i}\ x_i|\ |e^{t_2\lambda_i}\ x_i|,$$

relation (25) is a consequence of the Cauchy–Schwarz inequality.
Thus, we proved (23).

If condition 81 holds, then the function

$$f(t) = \ \log ||e^{tA}|| = \sup_{||x||=1} f_x(t)$$

is convex on $\mathbb{R}$. Letting $t_1 = 0,\ t_2 = 2t\ (t > 0)$ in

$$f\Big(\frac{t_1+t_2}{2}\Big) \ \le\ \frac{f(t_1)+f(t_2)}{2},$$

we have

$$||e^{tA}|| \ \le\ ||e^{2tA}||^{1/2}.$$

On the other hand,

$$||e^{2tA}|| \ =\ ||e^{tA}\cdot e^{tA}|| \ \le\ ||e^{tA}||^2$$

hence

$$||e^{tA}|| \ =\ ||e^{2tA}||^{1/2}.$$

By induction, we obtain

$$||e^{tA}|| \ =\ ||e^{2^k tA}||^{2^{-k}}, \quad k = 1, 2, \dots$$

Then the spectral radius formula implies

$$(26) \qquad ||e^{tA}|| \ =\ \rho(e^{tA}).$$

It is known [11] that, for the matrix $B_t = e^{tA}$, the equality $||B_t|| = \rho(B_t)$ ensures an orthogonal decomposition of the underlying space

$$\mathbb{C}^N = H_1 \oplus H_2,$$

with the projection of B_t on H_1 being a scalar matrix. Then the argument above can be applied to H_2. Continuing in this way, we finally obtain that B_t is normal. Since this holds for any positive t, the normality of A follows from condition 79.

We mention that condition 81 has first appeared in [5]. However, the proof above is adapted from [1].

Turning to condition 82, we observe that its necessity can be established with the help of the previous condition. In fact, let A be a normal matrix with the orthonormal eigenvectors $v_1, \ldots, v_n$. Then, for any real t, the matrix e^{tA} is normal as well, and (see condition 58), with the appropriate numbering,

$$\sigma_j(e^{tA}) \ = \ |\lambda_j(e^{tA})|, \quad j = 1, \ldots, n. \tag{27}$$

Letting $x = v_j$ in the definition of $f_x(t)$, we have

$$f_{v_j}(t) = \ \ln \ ||e^{tA} v_j|| \ = \ \ln |\lambda_j(e^{tA})| \ = \ \ln \ \sigma_j(e^{tA}) = f_j(t).$$

Hence, $f_j(t)$ must be convex on $\mathbb{R}$ $\ (j = 1, \ldots, n)$.

Since

$$||B|| = \sigma_1(B)$$

for any matrix B, the proof of sufficiency can mimick that of condition 81. Another possibility chosen in [1] is to prove equalities (27) by using a result from [17]. Then, again, conditions 58 and 79 may be applied.

Condition 83 is obviously connected with the Bauer–Fike theorem. It is useful to precede the proof of Condition 83 by some discussion of this theorem.

Let A, Λ, X be as in relation (11), and B an arbitrary perturbation of A. According to the Bauer–Fike theorem, for any eigenvalue μ of the perturbed matrix $A + B$ there exists an eigenvalue λ of A such that

$$|\lambda - \mu| \ \leq \ \text{cond } X \cdot ||B||. \tag{28}$$

For A normal, X can be chosen unitary, and (28) turns into

$$|\lambda - \mu| \ \leq \ ||B||. \tag{29}$$

Now, again for a normal A, it is always possible to perturb A along its eigendirections in such a way that (29) becomes an equality for at least some μ's, and for at least one of those μ's (say μ_0) the corresponding nearest (possibly, multiple) λ_0 is defined uniquely.

Condition 83 found by O. Hald in [7] essentially says that the property above is characteristic for normal matrices. Now we turn to the proof of the sufficiency of this condition.

Let x be an eigenvector of $A+B$ corresponding to μ_0 :

$$(A+B)x = \mu_0 \; x. \tag{30}$$

Using (11), one can rewrite (30) as

$$(\Lambda + X^{-1}BX) \; T^{-1} \; x = \mu_0 \; T^{-1} \; x. \tag{31}$$

By a proper normalization of x, we can achieve that $v = T^{-1} \; x$ is a unit vector. Now (31) assumes the form

$$(\Lambda - \mu_0 \; I) \; v = -T^{-1}BT \; v. \tag{32}$$

Applying (12), we deduce from (32)

$$\text{cond } X \cdot ||B|| \; = \; \min_j |\lambda_j - \mu_0| \; \leq \; ||(\Lambda - \mu_0 \; I)v|| \; \leq \; \text{cond } X \cdot ||B||.$$

Hence

$$||(\Lambda - \mu_0 \; I)v|| \; = \; \min_j |\lambda_j - \mu_0|.$$

Since the closest λ (say, λ_0) is unique, v must be an eigenvector of $\Lambda - \mu_0 \; I$:

$$(\Lambda - \mu_0 \; I)v \; = \; (\lambda_0 - \mu_0)v$$

which, by (32), is equivalent to

$$Bx = (\mu_0 - \lambda_0)x. \tag{33}$$

Finally, using (12) and (33), we have

$$\text{cond } X \cdot ||B|| \cdot ||x|| \; = \; |\lambda_0 - \mu_0| \cdot ||x|| \; \leq \; ||B|| \cdot ||x||. \tag{34}$$

Recall that cond $X \geq 1$ for any nonsingular matrix X. Therefore, (34) implies

$$\text{cond } X = 1,$$

and the normality of A follows from condition 72.

We mention a nice corollary of condition 83 also given in [7]:
if A is not normal, then the Bauer–Fike inequality is strict for small perturbations B. Indeed, if 2 cond $X \cdot ||B||$ is smaller than the minimal distance between eigenvalues of A, then, for each μ, the closest λ (simple or multiple) is defined uniquely. Since A is nonnormal, equality in (28) is impossible.

Conditions 84, 85 are two particular cases of condition 86 found in a recent publication [12]. They have been singled out just for the reason that a much simpler proof is possible for these cases compared with that for a basic condition.

The necessity part is the same for all three conditions. Let $A = U|A|$ be the polar decomposition of A. Then, according to condition 37, U and $|A|$ commute. Since $|A|^{1/2}$ is a polynomial of $|A|$, the matrices U and $|A|^{1/2}$ commute as well. Now we have, for any $u \in \mathbb{C}^n$,

$$\begin{aligned} |(Au, u)| = \; & |(U|A|u, u)| \; = \; |(U|A|^{1/2} \, u, \; |A|^{1/2} \, u)| \\ \leq \; & ||U|A|^{1/2} \, u|| \; || \, |A|^{1/2} \, u|| \; = \; || \, |A|^{1/2} \, u||^2 \; = \; (|A|u, u). \end{aligned}$$

Suppose condition 84 holds. Without loss of generality, A may be assumed to be an upper triangular matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$. Let $(t_1, \ldots, t_n)$ be the first row of $|A|$. Then the relation $|A|^2 = A^*A$ implies

$$t_1^2 + |t_2|^2 \; + \ldots + \; |t_n|^2 = |\lambda_1|^2.$$

On the other hand, using the condition for $u = (1, 0, \ldots, 0)^T$ gives

$$|\lambda_1| \; \leq \; t_1.$$

Hence

$$t_1 = |\lambda_1|, \quad t_2 \; = \ldots = \; t_n = 0.$$

Continuing in this way, we conclude that the singular values of A coincide with the moduli of its eigenvalues. By condition 58, A is normal.

For condition 85, the sufficiency is proved in the same way, with A as a lower triangular matrix and with $|A|$ replaced by $|A^*|$.

Now we outline the proof of sufficiency for condition 86 given in [12]. The proof is preceded by the following

Lemma 1. Let $A \in \mathbb{C}^{n,n}$ have all eigenvalues equal to 1 in absolute value. If

$$|(Ay, y)| \; \leq \; (Ay, Ay)^{\alpha} \quad \text{for all unit vectors} \quad y \in \mathbb{C}^n, \tag{35}$$

then either $\alpha \; = \; \frac{1}{2}$ or A is a unitary matrix.

To prove the lemma, one may assume that A is an upper triangular matrix. Then, inspecting (35) for y with only two components nonzero, the first and the sth one, one can show that

$$a_{1s} = 0, \quad s = 2, \ldots, n,$$

if $\alpha \neq \; \frac{1}{2}$. Applying an inductive argument, one shows A to be diagonal, hence unitary.

Returning to condition 86, we first consider the case where A is nonsingular. Let $A = UDV$ be the singular value decomposition of A. For any nonzero u, set

$$y \; = \; \frac{D^{1/2} \, U^* u}{||D^{1/2} \, Uu||}.$$

Then (13) can be rewritten in the form of (35), with A replaced by

$$\widetilde{A} = D^{1/2}\ VUD^{-1/2}.$$

According to Lemma 1, $\widetilde{A}$ is unitary which gives $VUD = DVU$. ¿From the last relation, the normality of A follows easily.

In the case where A is singular, the induction on n is used. For the induction step, one may assume, without loss of generality, that

$$A = \begin{pmatrix} A_1 & b \\ 0 & 0 \end{pmatrix},$$

where A_1 is an $(n-1)$–square matrix and b is a column $(n-1)$–vector.

Assume that $b \neq 0$, then A is not normal. Take $u_1 \in \mathbb{C}^{n-1}$ such that $(b, u_1) \neq 0$ and let

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad \text{with} \quad u_2 > 0. \tag{36}$$

Examining (13) when $u_2 \to \infty$ in (36) (and u_1 is fixed), we observe that the left–hand side grows linearly in u_2, with (b, u_1) as the leading coefficient. On the other hand, the leading term in the right–hand side behaves like $\gamma\, u_2^{2\alpha}$ for a nonzero γ. For inequality (13) to be valid for all u_2, we must have $\alpha \geq \frac{1}{2}$.

Applying the same argument to A^* with (13) rewritten in the form

$$|(A^* u, u)| \ \leq \ (|A^*|u, u)^{1-\alpha}\ (|A|u, u)^{1-(1-\alpha)} \quad \forall\ u \in \mathbb{C}^n,$$

we obtain $\alpha \leq \frac{1}{2}$. Hence, $\alpha = \frac{1}{2}$ which contradicts the assumption.

Condition 87 is an immediate consequence of the following assertion that can be found on page 426 of [10]:

Lemma 2. A necessary and sufficient condition for matrices A_1 and A_2 to be simultaneously diagonalizable is that both products $A_1\, A_2^*$ and $A_2^*\, A_1$ are normal matrices.

We mention that, while the unitary character of Q is vital for the validity of condition 87, the triangular form of R has no significance at all. Therefore, similar conditions may be stated for other decompositions of A with one of the factors unitary.

Condition 88 is a slightly disguised version of condition 59. One only needs to recall that the eigenvalues of $C_k(A)$ are the $\binom{n}{k}$ products

$$\lambda_{i_1} \ldots \lambda_{i_k}$$

where $1 \le i_1 < i_2 < \ldots < i_k \le n$ [14, p. 24], and the similar expressions hold for the singular values of $C_k(A)$.

To prove condition 89, we begin with the following observation. In the orthonormal basis of the unitary space $\mathbb{C}^{n,n}$ composed of the matrices

$$E_{ij} = (\delta_{ki}\ \delta_{j\ell})_{k,\ell=1}^{n}, \quad i,j = 1,\ldots,n,$$

the operator $\mathcal{L}_A$ is described by the matrix

$$\mathcal{M}_A = I_n \otimes A + \bar{A} \otimes I_n$$

(see [14, p. 9]). Thus $\mathcal{L}_A$ is normal iff $\mathcal{M}_A$ is.

Assume that

$$A \to R = U^* A U$$

is the unitary similarity which transforms A into its (upper triangular) Schur form. Then $\mathcal{M}_A$ is unitary similar to the upper triangular matrix

$$\mathcal{M}_R = I_n \otimes R + \bar{R} \otimes I_n.$$

We conclude that $A, \mathcal{M}_A$, and $\mathcal{L}_A$ are simultaneously normal (or not normal).

References

1. B. Aupetit and J. Zemanek, *A characterization of normal matrices by their exponentials*, Linear Algebra Appl. **132** (1990), 119–121.
2. J. Cohen, S. Friedland, T. Kato, and F.P. Kelly, *Eigenvalue inequalities for products of matrix exponentials*, Linear Algebra Appl. **45** (1982), 55–95.
3. L. Elsner and M.H.C. Paardekooper, *On measures of nonnormality of matrices*, Linear Algebra Appl. **92** (1987), 107–124.
4. K. Fan, *On a theorem of Weyl concerning eigenvalues of linear transformations I*, Proc. Nat. Acad. Sci. U.S.A. **35** (1949), 652–655.
5. S. Friedland, *A characterization of normal matrices*, Israel J. Math. **42** (1982), 235–240.
6. R. Grone, C.R. Johnson, E.M. Sa, and H. Wolkowicz, *Normal matrices*, Linear Algebra Appl. **87** (1987), 213–225.
7. O.H. Hald, *A converse to the Bauer–Fike theorem*, Linear Algebra Appl. **9** (1974), 267–274.
8. R.E. Hartwig and K. Spindelböck, *Matrices for which* A^* *and* A^+ *commute*, Linear and Multilinear Algebra **14** (1984), 241–256.
9. S. Hochwald, *Linear algebra by analogy*, Amer. Math. Monthly **98** (1991), 918–926.
10. R. Horn and C. Johnson, *Matrix Analysis*, Cambridge U.P., 1985.
11. R. Horn and C. Johnson, *Topics in Matrix Analysis*, Cambridge U.P., 1991.
12. C.R. Johnson and F. Zhang, *An operator inequality and matrix normality*, Linear Algebra Appl. **240** (1996), 105–110.
13. Kh.D. Ikramov, *Unitary–triangular factorizations of a normal matrix*, Comput. Math. Math. Phys. **33** (1993), 407–410.
14. M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Dover, 1992.
15. A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic Press, 1979.
16. W. So, *Equality cases in matrix exponential inequalities*, SIAM J. Matrix Anal.Appl. **13** (1992), 1154–1158.
17. T. Yamamoto, *On the extreme values of the roots of matrices*, J. Math. Soc. Japan **19** (1967), 173–178.