Moving zeros among matrices

by John Holbrook and Jean–Pierre Schoch, [1]

with an appendix by T. Košir and B. A. Sethuraman.

Dedicated to Roger Horn.

Abstract: We investigate the zero–patterns that can be created by unitary similarity in a given matrix, and the zero–patterns that can be created by simultaneous unitary similarity in a given sequence of matrices. The latter framework allows a "simultaneous Hessenberg" formulation of Pati's tridiagonal result for 4×4 matrices. This formulation appears to be a strengthening of Pati's theorem. Our work depends at several points on the simplified proof of Pati's result by Davidson and Djoković. The Hessenberg approach allows us to work with ordinary similarity and suggests an extension from the complex to arbitrary algebraically closed fields. This extension is achieved and related results for 5×5 and larger matrices are formulated and proved.

AMS classification: 15A21; 15A22

Keywords: tridiagonal matrices; Hessenberg forms; unitary similarity; algebraically closed fields

---

1. Introduction What patterns of zeros can be created in an arbitrary $n \times n$ matrix by means of unitary similarity? This question is perhaps too general to have a satisfactory solution. Nevertheless something of interest can be said and the present note deals with results, conjectures (supported to various extents by computer experiment), and more specific problems suggested by the general question.

This work was inspired in part by the recent paper of Davidson and Djoković

[1]Addresses: John Holbrook, Math/Stat Dept, University of Guelph, Guelph, Ontario N1G 2W1, Canada
Jean–Pierre Schoch, 839 Watson Rd. S., Arkell, Ontario N0B 1C0, Canada
e-mail: jholbroo@uoguelph.ca, jp_schoch@yahoo.co.uk

[DD]; they provide a new and simpler proof of the theorem of Pati [P]: every 4×4 complex matrix is unitarily similar to a tridiagonal matrix. Note that a 4×4 tridiagonal matrix has a particular pattern of six zeros, the same number as in the more familiar Schur upper–triangular form. We may say, in this case, that the zeros of the upper–triangular form can be *moved* (by a unitary similarity) into the positions corresponding to tridiagonal form. See section 2 for a discussion of this and related results such as the "simultaneous Hessenberg" conjecture.

In the upper–triangular form of an $n \times n$ matrix we have the triangular number $\Delta_n = n(n-1)/2$ of zeros below the diagonal. This form and the upper-Hessenberg form are examples of "upper–forms", ie forms in which the required zeros are in positions $\{(i,j) : i \geq z(j)\}$ and $z(j) \leq z(j+1)$. Thus for upper–triangular form $z(j) = j+1$ and for upper–Hessenberg form $z(j) = j+2$. Equivalently, a pattern of zeros defines an upper–form if, whenever (i,j) is the position of a required zero, so is (i',j') provided $i' \geq i$ and $j' \leq j$. We shall say a pattern of zeros in a sequence of 0–1 $n \times n$ matrices $A_1, A_2, \ldots, A_m$ is *feasible* if every sequence $B_1, B_2, \ldots, B_m$ of $n \times n$ complex matrices can be transformed simultaneously via similarity by a unitary U into the given pattern: $(\forall k) \quad (A_k)_{ij} = 0 \Rightarrow (UB_kU^*)_{ij} = 0$. In this note we focus on zero patterns where each A_k is an upper–form (which may depend on k).

Remark 1: For upper–forms *any* similarity is as good as unitary similarity: if S is invertible and SBS^{-1} has a given upper–form, then UBU^* also has that form, where $S = UT$ factors S into a unitary U and an upper–triangular T (U is obtained by applying the Gram–Schmidt process to the columns of S). Note that $S^{-1} = T^{-1}U^*$, that T^{-1} is also upper–triangular, and that an upper–form is preserved by right or left multiplication by upper–triangular matrices. This remark means that we need only deal with ordinary similarity and that the problems of feasible forms become simply a matter of complex algebraic geometry.

In section 3 we consider some feasibility problems that may be answered by means of some standard results on the dimensions of complex projective varieties. In section 4 we survey some more challenging problems, including certain conjectures about feasible zero patterns and the experimental evidence that supports them. We also mention problems that arise when commutativity of the matrix sequence is assumed.

An appendix by Tomaž Košir and B. A. Sethuraman establishes several of our conjectures, and greatly extends the range of such results.

Much of this work stems from discussions during the Conference in Honor of Heydar Radjavi's 70th Birthday and the 4th Linear Algebra Workshop held in Bled, Slovenia in May 2005. We are grateful to all the participants for their input but particularly to Rajendra Bhatia, Charles Johnson, and Roy Meshulam for very helpful comments. This work was supported in part by NSERC of Canada.

2. Tridiagonal and simultaneous Hessenberg forms

It appears that Longstaff [L] first introduced the problem of tridiagonalizing arbitrary matrices via unitary similarities. While this is easy for any 3×3 matrix, Longstaff and Sturmfels [St] used a dimensional argument to show that among matrices 6×6 and larger there must be some that are *not* unitarily similar to tridiagonals. Fong and Wu [FW] modified this approach to show that not all 5×5 matrices are unitarily similar to tridiagonal forms. This left the apparently challenging question of whether the tridiagonal form was feasible (in our sense) for 4×4 matrices. Fong and Wu guessed that it was not, but computer experiments designed by Holbrook and Schoch [Sch] strongly suggested that it *was* feasible. Finally Pati proved this via a rather formidable argument in algebraic geometry [P].

Holbrook and Schoch had also considered an (apparently) stronger conjecture which they called the "simultaneous Hessenberg" conjecture. This claims that the zero pattern consisting of two upper–Hessenberg 4×4 matrices is feasible, ie that given any pair B_1, B_2 of 4×4 complex matrices there is an invertible S such that both SB_1S^{-1} and SB_2S^{-1} are upper–Hessenberg. In view of Remark 1, this would immediately imply the tridiagonal result for any 4×4 B by applying it to the pair B, B^*. While numerical experiments reported in [Sch] provided considerable support for the simultaneous Hessenberg conjecture, it did not seem clear how to apply Pati's technique to settle that conjecture. Recently it was observed by Košir and Sethuraman that the new approach to Pati's result developed in [DD] *can* be adapted to establish the "stronger" conjecture.

Proposition 2: The zero pattern determined by two 4×4 upper–Hessenberg 0–1 matrices is feasible. A fortiori, any 4×4 B can be (unitarily) tridiagonalized.

Proof: Most of the arguments in section 2 of [DD], which deal with a pair A, B of Hermitian 4×4s, apply without change to arbitrary 4×4 matrices B_1, B_2. Of course, the subspace chains $V_1 \subset V_2 \subset V_3$ with $\dim(V_k) = k$ and $B_j V_k \subset V_{k+1}$ imply upper–Hessenberg forms in this general case, rather than tridiagonal forms. Note that the matrix pairs that are simultaneously similar to upper–Hessenbergs form a closed space; here Remark 1 plays a role: we may assume the similarities involved are unitary, and the unitary 4×4s form a compact space. Thus we need only prove the result for a dense subset of pairs B_1, B_2; we may thus assume that B_1 has distinct eigenvalues $\alpha_1, \alpha_2, \alpha_3, \alpha_4$. Since we can deal with ordinary similarity (Remark 1 again), we assume that B_1 is diagonal, with $B_1 e_i = \alpha_i e_i$. We also assume the analogue of condition (ii) in the proof of [DD, Theorem 2.1]:

$$\{e_i, B_2 e_i, B_2^2 e_i, B_1 B_2 e_i\} \text{ are linearly independent } (1 \leq i \leq 4). \qquad (1)$$

We observe, as in Remark 2.3 of [DD], that if one of the conditions (1) fails we immediately obtain a chain $V_1 \subset V_2 \subset V_3$ with $\dim(V_k) = k$ and $B_j V_k \subset V_{k+1}$. For example, if $V = \operatorname{span}\{e_i, B_2 e_i, B_2^2 e_i, B_1 B_2 e_i\}$ has dimension 2, then V is invariant for B_1 and B_2, so that the chain defined by $V_1 = \mathbb{C} e_i$, $V_2 = V$, and any three–dimensional V_3 (containing V_2) allows us to put B_1, B_2 in upper–Hessenberg form (with an extra zeros in the (3,2) position). The rest of the argument is just as in [DD] (with B_1 in place of A, B_2 in place of B). In particular, (1) allows us to check the multiplicities of the extraneous solutions as in Lemma 2.2 of [DD]. QED

Expressed visually, this proposition says that, along with the upper–triangular zero pattern

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

the following pattern is also feasible:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Note that the simultaneous Hessenberg formulation of the result may suggest further simplifications in the arguments: since ordinary similarity is all we require the problem becomes purely algebraic and we can set aside the more troublesome unitary (or Hermitian) conditions.

Note that a version of Proposition 2 may be given for matrices over any algebraically closed field; the appendix by Košir and Sethuraman takes this point of view and includes extensions to 5×5 and larger matrices. See also section 4.

In [FW] Fong and Wu show that a 4×4 matrix B can be unitarily tridiagonalized iff there exists a nonzero vector v such that

$$\text{rank}[v, Bv, B^*v, B^2v, (B^*)^2v, BB^*v, B^*Bv] \leq 3,$$

an observation that the authors attribute to Heydar Radjavi. In [Sch] the corresponding result for a pair B, C of 4×4 matrices is established: B and C can be put simultaneously into upper–Hessenberg form by a similarity iff the "joint Radjavi condition" is satisfied, namely, there exists a nonzero vector v such that

$$\text{rank}[v, Bv, Cv, B^2v, C^2v, BCv, CBv] \leq 3.$$

Thus another way of expressing Proposition 2 is to say that every pair of 4×4 matrices satisfies the joint Radjavi condition.

3. First–column zero patterns

We are indebted to Roy Meshulam for pointing out that certain zero–pattern problems can be resolved using standard results from algebraic geometry. The key result (see for example Proposition 12.2 of [H]) is that the complex projective variety $M(n, m, k)$ consisting of $n \times m$ nonzero complex matrices of rank at most k has codimension $(n-k)(m-k)$, assuming that $k \leq \max(n, m)$. Thus any linear subspace of the $n \times m$ matrices having dimension greater than $(n-k)(m-k)$ contains a (nonzero) matrix in $M(n,m,k)$, whereas if $d \le (n-k)(m-k)$ there is a subspace of dimension d that does not intersect $M(n,m,k)$. The following two examples illustrate these ideas.

EXAMPLE 3: The $\Delta_3 = 3$ zeros in the upper–triangular form of a 3×3 matrix can be moved to the lower left corners of any three 3×3 matrices B_1, B_2, B_3, ie there exists nonsingular S such that each of SB_1S^{-1}, SB_2S^{-1}, SB_3S^{-1} has a zero in position (3,1). Using our notion of feasible zero patterns, this says that the pattern determined by three upper–Hessenberg 3×3 matrices A_1, A_2, A_3 is feasible. We need only find $v \ne \vec{0}$ such that $\operatorname{rank}[v, B_1v, B_2v, B_3v] \le 2$; then choose a basis $\{v, u, t\}$ such that each $B_kv \in \operatorname{span}\{v, u\}$ so that with respect to this basis the matrix of each B_k is upper–Hessenberg. Consider the map $\varphi : \mathbb{C}^3 \to M(3,4)$ (where $M(n,m)$ denotes the space of $n\times m$ complex matrices) defined by

$$\varphi(v) = [v, B_1v, B_2v, B_3v].$$

Since φ is linear (and evidently injective) the dimension of $\varphi(\mathbb{C}^3)$ is 3. This exceeds the codimension of $M(3,4,2)$: $(3-2)(4-2) = 2$. Hence there exists $v \ne \vec{0}$ such that $\varphi(v) \in M(3,4,2)$. Expressed visually, the zero pattern

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

is feasible, as $\Delta_3 = 3$ might suggest.

EXAMPLE 4: The $\Delta_4 = 6$ zeros in the upper–triangular form of a 4×4 matrix *cannot* in general be moved to the 6 positions (2,1), (3,1) of B_1, B_2, B_3. Expressing this visually: although the upper–triangular zero pattern

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

is, as always, feasible, the pattern

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \tag{2}$$

is not feasible. To see this, note that if $\{v, u, t, s\}$ is a basis with respect to which B_1, B_2, B_3 assume the proposed zero pattern we have $B_1v, B_2v, B_3v \in \text{span}\{v, u\}$. Defining $\varphi : \mathbb{C}^4 \to M(4,4)$ by $\varphi(x) = [x, B_1x, B_2x, B_3x]$, we see that v is such that $\varphi(v) \in M(4,4,2)$. Since the codimension of $M(4,4,2)$ is $(4-2)(4-2) = 4$ there exists a 4–dimensional linear subspace L of $M(4,4)$ such that $L \cap M(4,4,2) = \emptyset$. Let C_1, C_2, C_3, C_4 be any basis for L and let $D_0, D_1, D_2, D_3 \in M(4,4)$ be defined by setting $D_je_k = C_ke_{j+1}$, where e_k is the k–th standard basis vector in $\mathbb{C}^4$; that is, the k–th column of D_j is column $j+1$ of C_k. Let $\psi : \mathbb{C}^4 \to M(4,4)$ be defined by $\psi(x) = [D_0x, D_1x, D_2x, D_3x]$. We see that $\psi(e_k) = C_k$ so that $\psi(\mathbb{C}^4) = L$. Since $M(4,4,2)$ is closed we can "wiggle" D_0, if necessary, so that D_0 is invertible and we still have $\psi(\mathbb{C}^4) \cap M(4,4,2) = \emptyset$. Let $B_k = D_0^{-1}D_k$ $(k = 1,2,3)$. Now we cannot have $\varphi(v) \in M(4,4,2)$ for then also $D_0\varphi(v) \in M(4,4,2)$, yet $D_0\varphi(v) = D_0[v, D_0^{-1}D_1v, D_0^{-1}D_2v, D_0^{-1}D_3v] = \psi(v)$.

Experiment 5: The existence of L, as in Example 4, follows on general principles, but in this setting we can (as Roy Meshulam pointed out) construct an appropriate L explicitly. It will be convenient to do so as follows. Let L consist of matrices of the form

$$\begin{bmatrix} a & c & 0 & 0 \\ b & a+d & c & 0 \\ c & b & a+2d & c \\ d & 0 & b & d \end{bmatrix},$$

where $a, b, c, d \in \mathbb{C}$. It is easy to see that such a matrix has rank greater than 2 unless it is 0_4, ie $a = b = c = d = 0$. Indeed, looking at the upper right 3×3 submatrix we see it has rank 3 unless $c = 0$; if $c = 0$ then the matrix has rank at least 3 unless unless some two of the diagonal elements are 0, implying $a = d = 0$; but the remaining matrix has rank 3 unless $b = 0$. Thus $L \cap M(4,4,2) = \emptyset$. Furthermore, using again the notation of Example 4, we have $L = \varphi(\mathbb{C}^4)$ where

$$B_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, B_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \end{bmatrix}, B_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Thus B_1, B_2, B_3 are three explicit matrices that cannot be transformed by a similarity to the zero pattern (2); this may seem unlikely in view of the fact that these matrices already have zeros in all but one of the required positions!

A computational experiment may be performed to test this explicit example and to determine, in a certain sense, how close we can come to the zero pattern (2). The techniques described here may also be instructive because they amount to toy versions of the techniques that earlier led to our belief in the tridiagonal result (proved, in time, by Pati) and in the simultaneous Hessenberg conjecture, as well as those that support the conjectures we shall discuss in section 4.

We claim that if a unitary U is such that UB_1U^* and UB_2U^* have zeros in positions (3,1) and (4,2), and UB_3U^* has zero in position (4,1), then the minimum possible value of $|(UB_3U^*)_{31}|$ is about 0.1975. An orthonormal basis v, u, t, s with respect to which B_1 and B_2 have the required zeros must be such that $B_1v, B_2v \in \text{span}\{v, u\}$, so that v, B_1v, B_2v are linearly dependent. Thus v must be a unit eigenvector of $B(z) = zB_1 + (1 - |z|)B_2$ for some $z \in \mathbb{D}$, the closed unit disc in $\mathbb{C}$. The corresponding u is then the (essentially unique) vector such that v, u are orthonormal in $\text{span}\{v, B_1v, B_2v\}$. A simple MATLAB program searches over a fine grid of z–values in $\mathbb{D}$ to determine the possible pairs v, u. If, in addition, B_3 has 0 in position (4,1) with respect to v, u, t, s we must have $B_3v \in \text{span}\{v, u, t\}$, ie t is obtained by orthonormalizing v, u, B_3v. For B_3, the entry in position (3,1) becomes (B_3v, t) and must have modulus $\sqrt{\|B_3v\|^2 - |(B_3v, v)|^2 - |(B_3v, u)|^2}$. The minimum of this value over a fine grid on $\mathbb{D}$ turns out to be about 0.1975. Perversely, if we interchange the roles of B_2 and B_3 so that we *start* with the required zeros in B_1 and B_2, the minimum modulus of the (3,1) entry in UB_3U^* is *larger*, about 0.68.

Based on the ideas of Examples 3 and 4, we may establish the following more general result.

PROPOSITION 6: A zero pattern of $n \times n$ matrices $A_1, A_2, \ldots, A_q$ requiring zeros (for each matrix) in the first column positions (i,1) for $i > k$ is feasible iff $n > (q + 1 - k)(n - k)$, assuming $k \leq q + 1, n$.

PROOF: We must have n greater than the codimension of $M(n, q + 1, k)$. QED

Remark 7: It follows easily that the total number of required zeros in such a pattern, namely $(n-k)q$, cannot exceed Δ_n, the number in the upper–triangular form of a single $n \times n$ matrix. In fact the inequality of Proposition 6 shows that the total number of zeros cannot even attain Δ_n for $n \geq 4$. We have seen an instance of this in Example 4; similarly, we cannot move $6 = \Delta_4$ zeros into the lower left corners of 6 arbitrary 4×4 matrices $(n = 4, q = 6, k = 3)$.

From Charles Johnson we learned of a related issue: it seems that among matrix analysts there is a folklore problem that asks whether there is some n such that every $n \times n$ matrix is unitarily similar to one with *more* than Δ_n zeros, where the pattern is allowed to change with the matrix. This is easy to rule out directly for very small n but otherwise requires an indirect, dimension–counting argument. Of course, Proposition 6 deals with *fixed* (and very special) zero patterns.

4. Further results, experiments, conjectures, and questions

Computer experiments (similar to that outlined in Experiment 5) with pairs of 5×5 matrices, attempting to put them in something like simultaneous Hessenberg form, supported conjectures that certain upper–forms are feasible for pairs of 5×5 matrices but that others are not. There is a connection with the examples in [DD] of 5×5 matrices that cannot be tridiagonalized (and with examples of that phenomenon found computationally in [Sch], and earlier suggestions in [Cam]). For example, computations suggested that the zero pattern

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \tag{3}$$

is feasible. Note the $10 = \Delta_5$ zeros in total. This conjecture is verified in the appendix below, and in fact there is a nice $n \times n$ generalization of (the simultaneous Hessenberg version of) Pati's result: any two $n \times n$ matrices are (unitarily) similar to an upper–Hessenberg matrix and a matrix with zeros in positions $(3,1) - (n,1)$, and $(n,2)$.

Other experiments suggested that the patterns

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} \tag{4}$$

and

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix} \tag{5}$$

are feasible, via unitary similarity. Since these last two are not upper–forms, the distinction between similarity and unitary similarity must be addressed. The appendix deals with patterns like (4) and (5) as well.

The fact that commuting matrices can be simultaneously triangularized suggests a host of questions about zero patterns. For example, given commuting $n \times n$ matrices, can we move the $2\Delta_n$ zeros that seem to be available to other positions? It seems unclear whether significant results of this type are available. For example, if S is the shift

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \tag{6}$$

then S and S^2 cannot even be put simultaneously in tridiagonal form!

Appendix: Upper-Forms for Pairs of Matrices

by Tomaž Košir and B. A. Sethuraman [2]

Holbrook and Schoch asked whether any pair of matrices in $M_4(\mathbb{C})$ is unitarily similar to a pair of matrices in the upper-Hessenberg form. The purpose of this appendix is to show that by simply restating the results and rewording some of the arguments used by Davidson and Djoković to prove Pati's theorem [DD], that question can be answered affirmatively. While in [DD] the results are over the field of complex numbers and for a pair of hermitian matrices we restate them over an algebraically closed field and for a general pair of matrices. Moreover, using a slight generalization of Davidson and Djoković's arguments, we can obtain a more general "upper-form" result for $n \times n$ matrices; we sketch the proof of this result as well.

Let us first introduce some notation. We denote by F an algebraically closed field and by $M_n(F)$ the set of all $n \times n$ matrices over F. A pair of matrices (A, B) is *simultaneously similar* to a pair of matrices (C, D) if there is an invertible matrix S such that $C = S^{-1}AS$ and $D = S^{-1}BS$. If $F = \mathbb{C}$ and $S^{-1} = S^*$ is a unitary matrix then we say that (A, B) is *unitarily similar* to (C, D). A matrix $A = [a_{ij}] \in M_n(F)$ is in the *upper-Hessenberg form* if $a_{ij} = 0$ whenever $i - j > 1$. We denote the projective space of dimension k over F by $\mathbb{P}^k(F)$. We view the space of all nonzero pairs of matrices, modulo simultaneous multiplication by a nonzero scalar, as the projective space $\mathbb{P}^{2n^2-1}(F)$, so that nonzero pairs of matrices act as homogeneous coordinates for $\mathbb{P}^{2n^2-1}(F)$.

First we state the result over a general F. It was essentially proved by Davidson and Djoković in [DD].

[2] Addresses: T. Košir, Department of Mathematics, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia.
B. A. Sethuraman, Department of Mathematics, California State University Northridge, Northridge CA 91330, U.S.A.
e-mail: tomaz.kosir@fmf.uni-lj.si, al.sethuraman@csun.edu.
The authors were supported in part by US–Slovenian bilateral research grants from the National Science Foundation, USA, and the Ministry of Higher Education, Science and Technology, Slovenia.

Theorem A.1. *A pair of matrices (A, B) in $M_4(F)$ is simultaneously similar to a pair of matrices in the upper-Hessenberg form.*

Proof. We outline a proof which is in essence the same as the proof of Davidson and Djoković in [DD, §2].

We need to establish the existence of a complete flag of subspaces

$$\mathcal{F}:\ 0 \subset V_1 \subset V_2 \subset V_3 \subset V_4 = F^4,\ \dim V_j = j,$$

such that

$$AV_j \subset V_{j+1},\ BV_j \subset V_{j+1};\ j = 1, 2. \tag{7}$$

We denote the projective variety of all complete flags $\mathcal{F}$ by $\mathbb{F}$. The conditions such as (7) can be expressed as polynomial conditions in the Plücker coordinates of various Grassmannians (see, for instance, Part III of [Fu] or [H]). Thus the set of triples $(A, B; \mathcal{F})$ in the projective variety $\mathbb{P}^{31}(F) \times \mathbb{F}$ that satisfy the conditions (7) is closed (in the Zariski topology) and therefore also its projection to the first component is closed (see [Sh, p. 58]). It is irreducible since it is the image of the irreducible variety of all triples $(A, B; S)$ under the morphism $(A, B; S) \mapsto (S^{-1}AS, S^{-1}BS)$, where A and B are matrices in the upper-Hessenberg form and S is invertible. It is enough to show that a generic pair of matrices is in the projection to the first component to conclude that every pair of matrices is in the projection, i.e., to conclude that every pair of matrices is simultaneously similar to a pair of matrices in the upper-Hessenberg form. Following [DD] we make the following generic assumptions:

1. A has four distinct eigenvalues $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ and it is diagonal in a basis $\{e_1, e_2, e_3, e_4\}$.

2. Vectors e_j, Be_j, B^2e_j, ABe_j are linearly independent for each j.

We briefly outline an argument that shows that these indeed are generic assumptions.

Consider map $\varphi : \mathrm{Gl}_4(F) \times U \times V \to F^{16} \times F^{16}$, where $\varphi(G, A, B) = (GAG^{-1}, GBG^{-1})$, $U \subset F^4$ is the open set of all tuples $\begin{pmatrix}\alpha_1 & \alpha_2 & \alpha_3 & \alpha_4\end{pmatrix}$, where the α_i are all distinct and nonzero (we want to think of this 4-tuple as the diagonal of the matrix A) and $V \subset F^{16}$ is the set of all matrices B

such that for each j, the j-th column of I, the j-th column of B, the j-th column of B^2, and the j-th column of AB are linearly independent. Since we are using a fixed basis these are open set conditions on diagonal matrix A and matrix B. We want to show that the closure of the image of φ is F^{32}. To compute the dimension of the image of φ, note that the domain space has dimension 36. The fiber over a typical point (GAG^{-1}, GBG^{-1}) consists of all triples $(GM, M^{-1}AM, M^{-1}BM)$, where M is an invertible monomial matrix, i.e., M is a product of a permutation matrix and a diagonal matrix. Thus, the fiber is 4-dimensional. It follows that the dimension of the image of φ is 32, as needed.

Next we consider the system of equations

$$(sA + tB - \lambda I)u = 0 \text{ and } \det\begin{pmatrix} u & Au & A^2u & BAu \end{pmatrix} = 0 \tag{8}$$

for $([s : t : \lambda]; u) \in \mathbb{P}^2(F) \times \mathbb{P}^3(F)$. The key ideas that lead to these equations are the following: Observe that the generic assumptions imply that matrices A and B have no common eigenvector. Then we consider solutions of the system (8). If the second component u of a solution $([s : t : \lambda]; u)$ is different from all e_j then it gives a chain of distinct subspaces $0 \subset V_1 = \mathrm{span}(u) \subset V_2 = \mathrm{span}(u, Au)$ such that $AV_1 \subset V_2$ and $BV_1 \subset V_2$. If the dimension of the subspace $V_3 = \mathrm{span}(u, Au, A^2u)$ is equal to 3 the subspaces V_j induce a required flag of subspaces. Otherwise the dimension of V_3 is equal to 2 and we consider the subspace $V_3' = \mathrm{span}(u, Au, BAu)$. If it is of dimension 3 we have a required flag. If $\dim V_3 = \dim V_3' = 2$ then any 3-dimensional subspace V_3'' such that $V_2 \subset V_3''$ will give us a required flag. Next, we observe that $([1 : 0 : \alpha_j], e_j)$ for each j is also a solution of the system (8) but it does not yield a required flag $\mathcal{F}$. It is called an extraneous solution. We need to establish the existence of a non-extraneous solution. It follows from Bezout's Theorem for a product of projective spaces (see [Sh, p. 237]) that if the system (8) has finitely many solutions then it has generically 24 solutions (counting multiplicities). Exactly the same proof as that of Lemma 2.2 in [DD] shows that each extraneous solution has multiplicity three and so all of them account for 12 of the solutions. Hence, there is at least one flag which gives the upper-Hessenberg form for (A, B). QED

Remark. From [DD, Example 5.1] it follows that if the characteristic of F is 0 there are generically exactly 12 flags that give upper-Hessenberg form for (A, B). Here note that all the entries of both matrices in the example are integers.

The following is a simple consequence of Theorem A.1. It was essentially proved in [DD].

Corollary A.2. *A pair of matrices* (A, B) *in* $M_4(\mathbb{C})$ *is unitarily similar to a pair of matrices in the upper-Hessenberg form.*

Proof. By Theorem A.1 there is an invertible matrix S such that $S^{-1}AS$ and $S^{-1}BS$ are in the upper-Hessenberg form. Let $S = QR$ be a QR-decomposition of S, i.e., a decomposition where Q is a unitary matrix and R is an invertible upper-triangular matrix. Since R is upper-triangular it follows that $(RS^{-1}ASR^{-1}, RS^{-1}BSR^{-1}) = (Q^*AQ, Q^*BQ)$ is a pair of matrices in the upper-Hessenberg form. QED

As another consequence of Theorem A.1 and Corollary A.2 we have a proof of Pati's Theorem [P]:

Corollary A.3. Every matrix $A \in M_4(\mathbb{C})$ is unitarily similar to a tridiagonal matrix.

Proof. Apply Corollary A.2 to a pair (A, A^*). QED

The Davidson-Djoković arguments can be generalized further for $n \geq 5$. It follows from results of Longstaff [L], Sturmfels [St] and Fong and Wu [FW] on tridiagonalization of a matrix that not all pairs of $n \times n$ matrices for $n \geq 5$ are simultaneously similar to a pair of matrices in the upper-Hessenberg form. However a less restrictive upper-form is possible.

Theorem A.4. *A pair of matrices* (A, B) *in* $M_n(F)$ *is simultaneously similar to a pair* (C, D)*, where* C *is in the upper-Hessenberg form and* D *is a matrix with zeros in positions* $(i, 1)$*,* $i = 3, 4, \ldots, n$ *and* $(n, 2)$*.*

Proof. The proof is a slight generalization of the arguments in [DD]. We sketch a proof following the arguments of the proof of Theorem A.1.

For a pair of $n \times n$ matrices (A, B) we have to find a complete flag of subspaces

$$\mathcal{F}: \ 0 \subset V_1 \subset \cdots \subset V_{n-1} \subset V_n = F^n, \ \dim V_j = j,$$

such that

$$AV_j \subset V_{j+1}, \ j = 1, 2, \ldots, n-1; \ BV_1 \subset V_2, \ BV_2 \subset V_{n-1}. \tag{9}$$

As in the proof of Theorem A.1 we show that the set of all pairs of matrices (A, B) that satisfy conditions (9) for a complete flag $\mathcal{F}$ is an irreducible variety in $\mathbb{P}^{2n^2-1}(F)$. Then it is enough to show that the statement of the theorem is true for a generic pair of matrices. We make the following generic assumptions:

1. A has n distinct eigenvalues $\alpha_1, \alpha_2, \ldots, \alpha_n$ and the corresponding eigenvectors are $e_1, e_2, \ldots, e_n$.
2. The set of vectors $\{e_j, Be_j, B^2 e_j, A^k B e_j \ k = 1, 2, \ldots, n-3\}$ is linearly independent for each j.

The system of equations to consider is

$$(sA + tB - \lambda I)u = 0 \text{ and } \det\begin{pmatrix} u & Au & \cdots & A^{n-2}u & BAu \end{pmatrix} = 0 \tag{10}$$

for $([s : t : \lambda]; u) \in \mathbb{P}^2(F) \times \mathbb{P}^{n-1}(F)$. As in [DD] we see that there are either infinitely many solution or there are $(n^3 - n^2)/2$ solutions counting multiplicities; here we apply Bezout's Theorem for a product of projective spaces (see [Sh, p. 237]). We have n extraneous solutions $([1 : 0 : \alpha_j], e_j)$, which do not give a flag satisfying conditions (9). To find the multiplicity of an extraneous solution we use similar arguments to those in the proof of Lemma 2.2 in [DD]. Let us briefly outline them for $([1 : 0 : \alpha_1], e_1)$. In the local ring at this point we set $s = 1$ and $u_1 = 1$, where we write $u = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix}^T$. The unique maximal ideal is generated by $\{\alpha_1 - \lambda, t, u_2, u_3, \ldots, u_n\}$. Then the lowest degree terms of the equations in the system $(A + tB - l\lambda I)u = 0$ are $\alpha_1 - \lambda, u_2, \ldots, u_n$. Modulo the ideal $\mathcal{J}_0$ that these terms generate we have

$$\begin{aligned}
& \det\begin{pmatrix} u & Au & \cdots & A^{n-2}u & BAu \end{pmatrix} \\
\equiv \ & \det\begin{pmatrix} u & (A - \lambda I)u & \cdots & (A^{n-2} - \lambda^{n-2} I)u & BAu \end{pmatrix} \\
\equiv \ & \det\begin{pmatrix} u & -tBu & \cdots & -(A^{n-3} + \lambda A^{n-4} + \cdots + \lambda^{n-3} I)\, tBu & BAu \end{pmatrix} \\
\equiv \ & (-t)^{n-2} \det\begin{pmatrix} u & Bu & \cdots & A^{n-3}Bu & B(A - \lambda I)u \end{pmatrix} \\
\equiv \ & (-t)^{n-1} \det\begin{pmatrix} u & Bu & \cdots & A^{n-3}Bu & B^2 u \end{pmatrix}.
\end{aligned}$$

By our generic assumption the latter determinant is nonzero and we see that the extraneous solution has multiplicity $n-1$. Then it follows that all the extraneous solutions account for n^2-n solutions and that there are $(n^3-3n^2+2n)/2$ 'good' solutions. We still have to show that a 'good' solution gives a required flag. Given a 'good' solution one considers the flag $V_1 = \text{span}(u)$, $V_2 = \text{span}(u, Au)$, $\ldots$, $V_{n-1} = \text{span}(u, Au, ..., A^{n-2}u)$. Observe that if $([s:t:\lambda]; u)$ is a 'good' solution then u is different from all e_j, $t \neq 0$ and therefore V_1 and V_2 are distinct and such that $AV_1 \subset V_2$ and $BV_1 \subset V_2$. If all the subspaces V_i are distinct it follows from the second of equations (10) that conditions (9) hold and we are done. Otherwise $V_i = V_{i+1}$ for some $i \geq 2$. In this case we extend the flag $V_1, \ldots, V_i$ by adding suitable vectors e_k to obtain $V_{i+1}, ..., V_{n-2}$. If BAu is not in V_{n-2} then add it to obtain V_{n-1}, otherwise just add a suitable e_k. This shows the existence of a required flag. QED

Remark. To show that generically the 'good' solutions in the above proof correspond to distinct flags one would have to prove the existence of at least one example with all distinct flags. Holbrook ran a computer experiment which indicates that for $F = \mathbb{C}$ and $n = 5$ there are 30 distinct solutions.

Remark. The extra zero entry in the second column of D in Theorem A.4 can be moved into some other position (p, r) with $p - r \geq 2$. We discuss the two remaining cases when $n = 5$. One could try to generalize the arguments for some other small $n > 5$. However, the case by case analysis is not suitable for general n.

When $n = 5$ there remain two cases to consider: (a) $(p, r) = (4, 2)$ and (b) $(p, r) = (5, 3)$. To show that the extra zero can be moved to these (p, r) positions we have to adjust the proof of Theorem A.3 appropriately. Let us explain all the necessary changes.

We discuss both cases simultaneously. We are looking for a complete flag of subspaces

$$\mathcal{F}: \; 0 \subset V_1 \subset \cdots \subset V_4 \subset V_5 = F^5, \; \dim V_j = j,$$

such that

$$AV_j \subset V_{j+1}, \; j = 1, 2, 3, 4; \; BV_1 \subset V_2, BV_r \cap V_p \subset V_{p-1}. \qquad (11)$$

Since these are all algebraic conditions it follows that the set of all pairs (A, B) satisfying (11) is closed and irreducible. We keep the first generic assumption that A has distinct eigenvalues with corresponding eigenvectors e_j while we replace the second one by the assumption that the set of vectors $\{e_j, Be_j, B^2e_j, ABe_j, A^3Be_j\}$ and $\{e_j, Be_j, BABe_j, ABe_j, A^2Be_j\}$, in cases (a) and (b) respectively, is linearly independent for each j. The system of equations we consider is now changed to

$$(sA + tB - \lambda I)u = 0 \text{ and } \det\begin{pmatrix} u & Au & A^2u & A^4u & BAu \end{pmatrix} = 0 \qquad (12)$$

and

$$(sA + tB - \lambda I)u = 0 \text{ and } \det\begin{pmatrix} u & Au & A^2u & A^3u & BA^2u \end{pmatrix} = 0, \qquad (13)$$

in cases (a) and (b) respectively. Now the same arguments as in the proof of Theorem A.3 show that if the above systems have finitely many solutions then they have 50 solutions counting multiplicities. To find the multiplicity of each extraneous solution $([1 : 0 : \alpha_j], e_j)$ one has to adjust the proof of Lemma 2.2 of [DD]. Following the arguments there we take the extraneous solution $([1 : 0 : \alpha_1], e_1)$ and view the local ring over that point. We may set $s = 1$ and $u_1 = 1$, where $u = \begin{pmatrix} u_1 & u_2 & \cdots & u_5 \end{pmatrix}^T$. The unique maximal ideal m in the local ring is generated by $\lambda - \alpha_1, u_2, u_3, u_4, u_5$ and t. Let $J_a \subset m$ and $J_b \subset m$ be the ideals generated by equations (12) and (13), respectively. In both cases, the minimal terms of the 5 polynomials of bidegree $(1, 1)$ are $-\lambda + \alpha_1, u_2, u_3, u_4$ and u_5. We denote by J_0 the ideal that they generate. Modulo the ideal J_0 we have in the case (a):

$$\begin{aligned}
& \det\begin{pmatrix} u & Au & A^2u & A^4u & BAu \end{pmatrix} \\
= \; & \det\begin{pmatrix} u & (A - \lambda I)u & (A^2 - \lambda^2 I)u & (A^4 - \lambda^4 I)u & BAu \end{pmatrix} \\
\equiv \; & \det\begin{pmatrix} u & -tBu & -(A + \lambda I)tBu & -(A^3 + \lambda A^2 + \lambda^2 A + \lambda^3 I)tBu & BAu \end{pmatrix} \\
= \; & -t^3 \det\begin{pmatrix} u & Bu & ABu & (A^3 + \lambda A^2)Bu & B(A - \lambda I)u \end{pmatrix} \\
= \; & -t^3 \det\begin{pmatrix} u & Bu & ABu & (A^3 + \lambda A^2)Bu & -tB^2u \end{pmatrix} \\
= \; & t^4 \det\begin{pmatrix} u & Bu & ABu & (A^3 + \lambda A^2)Bu & B^2u \end{pmatrix}.
\end{aligned}$$

The degree 4 term of the latter expression is equal to

$$t^4 \det\begin{pmatrix} e_1 & Be_1 & ABe_1 & A^3Be_1 & B^2e_1 \end{pmatrix},$$

which is nonzero because of our generic assumption. Similarly, we compute in the case (b):

$$\begin{aligned}
& \det\begin{pmatrix} u & Au & A^2u & A^3u & BA^2u \end{pmatrix} \\
= \ & \det\begin{pmatrix} u & (A-\lambda I)u & (A^2-\lambda^2 I)u & (A^3-\lambda^3 I)u & BA^2u \end{pmatrix} \\
\equiv \ & \det\begin{pmatrix} u & -tBu & -(A+\lambda I)tBu & -(A^2+\lambda A+\lambda^2 I)tBu & BA^2u \end{pmatrix} \\
= \ & -t^3\det\begin{pmatrix} u & Bu & ABu & A^2Bu & B(A^2-\lambda^2 I)u \end{pmatrix} \\
= \ & -t^3\det\begin{pmatrix} u & Bu & ABu & A^2Bu & -tB(A+\lambda I)Bu \end{pmatrix} \\
= \ & t^4\det\begin{pmatrix} u & Bu & ABu & A^2Bu & (BAB+\lambda B^2)u \end{pmatrix}.
\end{aligned}$$

The degree 4 term of the latter expression is equal to

$$t^4\det\begin{pmatrix} e_1 & Be_1 & ABe_1 & A^2Be_1 & BABe_1 \end{pmatrix}.$$

It is nonzero because of the generic assumption. This shows that the extraneous solutions account for 20 solutions in both cases. Thus there are 30 'good' solutions counting multiplicities. Given a 'good' solution one considers the flag $V_1 = \operatorname{span}(u)$, $V_2 = \operatorname{span}(u, Au)$, $V_3 = \operatorname{span}(u, Au, A^2u)$, $V_4 = \operatorname{span}(u, Au, A^2u, A^3u)$, $V_5 = F^5$. If these subspaces are all distinct we are done. Otherwise $V_i = V_{i+1}$ for $i = 2$ or $i = 3$. In these cases we have to consider some further subcases. We list all the different cases and give a required flag $\mathcal{F}$ in each case. Vectors e_j and e_k below are always chosen so that the corresponding subspace is of appropriate dimension. No other condition is imposed on their choice. Such a choice is always possible since vectors e_j form a basis.

Case (a), $i = 2$, $BAu \notin V_2$ and $ABAu \notin \operatorname{span}(u, Au, BAu)$:

$$\mathcal{F} : V_1 \subset V_2 \subset \operatorname{span}(u, Au, BAu) \subset \operatorname{span}(u, Au, BAu, ABAu) \subset F^5$$

Case (a), $i = 2$, $BAu \notin V_2$ and $ABAu \in \operatorname{span}(u, Au, BAu)$:

$$\mathcal{F} : V_1 \subset V_2 \subset \operatorname{span}(u, Au, BAu) \subset \operatorname{span}(u, Au, BAu, e_k) \subset F^5$$

Case (a), $i = 2$ and $BAu \in V_2$:

$$\mathcal{F} : V_1 \subset V_2 \subset \operatorname{span}(u, Au, e_k) \subset \operatorname{span}(u, Au, e_k, e_l) \subset F^5$$

Case (a), $i = 3$ and $BAu \notin V_3$:

$$\mathcal{F} : V_1 \subset V_2 \subset V_3 \subset \operatorname{span}(u, Au, A^2u, BAu) \subset F^5$$

Case (a), $i = 3$ and $BAu \in V_3$:

$$\mathcal{F} : V_1 \subset V_2 \subset V_3 \subset \text{span}(u, Au, A^2u, e_k) \subset F^5$$

Case (b) and $i = 2$: We can find a vector $w \notin V_2$ such that either w is an eigenvector of B or $\{w, w'\}$ or $\{w, w', w''\}$ is a Jordan chain for B with $w', w'' \in V_2$. In all cases $Bw \in \text{span}(u, Au, w)$. Case $Aw \notin \text{span}(u, Au, w)$:

$$\mathcal{F} : V_1 \subset V_2 \subset \text{span}(u, Au, w) \subset \text{span}(u, Au, w, Aw) \subset F^5$$

Case $Aw \in \text{span}(u, Au, w)$:

$$\mathcal{F} : V_1 \subset V_2 \subset \text{span}(u, Au, w) \subset \text{span}(u, Au, w, e_k) \subset F^5$$

Case (b), $i = 3$ and $BA^2u \notin V_3$:

$$\mathcal{F} : V_1 \subset V_2 \subset V_3 \subset \text{span}(u, Au, A^2u, BA^2u) \subset F^5$$

Case (b), $i = 3$ and $BA^2u \in V_3$:

$$\mathcal{F} : V_1 \subset V_2 \subset V_3 \subset \text{span}(u, Au, A^2u, e_k) \subset F^5.$$

References

[Cam] M. Camarri, Tridiagonalization and triangularization of linear operators, Honours Dissertation, University of Western Australia, 1993.

[DD] K. R. Davidson and D. Z. Djoković, Tridiagonal forms in low dimensions, Linear Alg. Appl. 407 (2005), 169–188.

[FW] C.–K. Fong and P.–Y. Wu, Band diagonal operators, Linear Alg. Appl. 248 (1996), 195–204.

[Fu] W. Fulton, Young tableaux, Cambridge University Press, 1997.

[H] J. Harris, Algebraic Geometry, A First Course, Springer 1992.

[L] W. E. Longstaff, On the tridiagonalization of matrices, Linear Alg. Appl.

109 (1988), 153–163.

[P] V. Pati, Unitary tridiagonalisation in $M(4, \mathbb{C})$, Proc. Indian Acad. Sci. (Math. Sci.) 111, 381–397.

[Sch] J.–P. Schoch, Theory vs Experiment in Matrix Analysis, PhD thesis, University of Guelph 2002.

[Sh] I. R. Shafarevich, Basic Algebraic Geometry 1. Springer, 1994.

[St] B. Sturmfels, Tridiagonalization of complex matrices and a problem of Longstaff, Linear Alg, Appl. 109 (1988), 164–165.