
Canonical forms for congruence of matrices and T -palindromic matrix pencils: a tribute to H. W. Turnbull and A. C. Aitken

Fernando De Terán

the date of receipt and acceptance should be inserted later

Abstract A canonical form for congruence of matrices was introduced by Turnbull and Aitken in 1932. More than 70 years later, in 2006, another canonical form for congruence has been introduced by Horn and Sergeichuk. The main purpose of this paper is to draw the attention on the pioneer work of Turnbull and Aitken and to compare both canonical forms. As an application, we also show how to recover the spectral structure of T -palindromic matrix pencils from the canonical form of matrices by Horn and Sergeichuk.

Keywords matrices, canonical forms, congruence, T -palindromic matrix pencils, equivalence

Mathematics Subject Classification (2000) 15A21, 15A22

1 Introduction

1.1 Congruence and similarity

We consider the following two actions of the general linear group of $n \times n$ invertible matrices with complex entries (denoted by $\text{GL}(n, \mathbb{C})$) on the set of $n \times n$ matrices with complex entries (denoted by $\mathbb{C}^{n \times n}$):

$$\begin{array}{ccc} \text{GL}(n, \mathbb{C}) \times \mathbb{C}^{n \times n} & \longrightarrow & \mathbb{C}^{n \times n} \\ (P, A) & \longmapsto & PAP^{-1}, \end{array} \quad \text{and} \quad \begin{array}{ccc} \text{GL}(n, \mathbb{C}) \times \mathbb{C}^{n \times n} & \longrightarrow & \mathbb{C}^{n \times n} \\ (P, A) & \longmapsto & PAP^T. \end{array}$$

The first one is known as the action of *similarity* and the second one as the action of *congruence*. Accordingly, two matrices $A, B \in \mathbb{C}^{n \times n}$ are said to be *similar* if there is a nonsingular $P \in \mathbb{C}^{n \times n}$ such that $PAP^{-1} = B$ and they are said to be *congruent* if there is a nonsingular P such that $PAP^T = B$. Similar matrices correspond to the same linear transformation in different bases, whereas congruent matrices correspond to equivalent bilinear forms. Both similarity and congruence determine equivalence relations in $\mathbb{C}^{n \times n}$. These relations allow us to classify the set $\mathbb{C}^{n \times n}$ into equivalence classes (also called *orbits*) consisting of equivalent matrices. A usual approach to identify these equivalence classes is by introducing a canonical form for the corresponding relation. Then the problem of determining whether or not two matrices are equivalent reduces to the problem of finding the canonical form of these matrices. There are several interesting connections between canonical forms for similarity and congruence. For instance, the *Sylvester equation* $XA - AX = 0$ can be solved in terms of the canonical form for similarity of A [5, Chapter VIII, §2], whereas the solution of the T -Sylvester matrix equation $XA + AX^T = 0$ depends on the canonical form for congruence [4]. But there is a remarkable difference between them. On the one hand, the canonical form for similarity goes back to Jordan (1838-1922) and it is widely known by the mathematical community (it is the well-known *Jordan Canonical Form*, from now on, the JCF). Moreover, it is taught in undergraduate courses in mathematics and in other scientific disciplines, and many textbooks in basic (and not so basic) linear algebra deal with it.

Partially supported by the Ministerio de Educación y Ciencia of Spain through grant MTM-2012-32542. A previous version of this work has been published in: Actas del II Congreso ALAMA 2010, Valencia (Spain), ISBN: 978-84-8363-544-5

Departamento de Matemáticas, Universidad Carlos III de Madrid · Avda. Universidad 30, 28911 Leganés, Spain
E-mail: fteran@math.uc3m.es

By contrast, the canonical form for congruence was introduced in 1932 and it is not sufficiently known even by linear algebra researchers. Moreover, several canonical forms for congruence have been proposed later by different authors and none of them is still completely accepted by researchers (this would not be surprising, since two of them have been introduced in 2006). The canonical form for congruence is the subject of the present paper.

1.2 Some history about the canonical form for congruence

A canonical form for congruence of complex matrices was derived by Turnbull and Aitken in a classical reference on matrix theory [20, p. 139]. Turnbull and Aitken obtained first a canonical form for congruence of pencils of the type $\mu A + \lambda A^T$ (here μ, λ are complex variables) and, from this canonical form they derive the canonical form for matrices just by taking $\mu = 1, \lambda = 0$. The most immediate difference between the JCF and the congruence canonical form of Turnbull and Aitken (from now on, the TACF for brevity) is that whereas the JCF consists of just one type of blocks (the *Jordan blocks* associated with a particular eigenvalue) the TACF consists of six types of blocks (see Section 3).

In 1947, Hodge and Pedoe, without mentioning the work of Turnbull and Aitken at all, derived in [7] the same canonical form as Turnbull and Aitken (except for a slight difference in the presentation of one type of blocks). Their canonical form for $A \in \mathbb{C}^{n \times n}$ comes from the canonical reduction of matrix pairs (R, S) , with R symmetric and S skew-symmetric, applied to the symmetric and skew-symmetric parts of A . This reduction is obtained by the authors using similar techniques as the ones introduced by Bromwich forty years before in [1] for pairs of symmetric matrices.

A long time after the publication of Turnbull and Aitken's book (we are using a reprinted edition [20] of the original book, which goes back to 1932), Sergeichuk provided in [18] canonical forms for congruence over any field \mathbb{F} of characteristic not 2 up to classification of Hermitian forms over finite extensions of \mathbb{F} . Based on that work Horn and Sergeichuk established in [9] a canonical form for congruence over fields of characteristic not 2 with an involution (later extended in [12] to bilinear forms over arbitrary algebraically closed fields and real closed fields). When particularized to the complex field, they achieve in [11] another canonical form for congruence in $\mathbb{C}^{n \times n}$. From now on, for brevity we will refer to this form as the HSCF. In the meantime, another canonical form was proposed in [3]. Nonetheless, as it was noticed by Horn and Sergeichuk in [11, p. 1013], some of the canonical matrices in this form are "cumbersome and not canonical".

In 1992 Thompson published an expository paper [19], where he showed canonical forms for congruence of pencils $\rho R + \sigma S$, with R, S being symmetric or skew-symmetric (both real and complex). The case where R is symmetric and S skew-symmetric is closely related to the type $\mu A + \lambda A^T$. More precisely, as noticed by Turnbull and Aitken [20, p. 136], an elementary change of variables takes the pencil $\rho R + \sigma S$ into the pencil $\mu A + \lambda A^T$. In 1996, apparently unaware of [7] and [20], Lee and Weinberg [15] followed Thompson's paper to derive another canonical form for congruence of matrices from the canonical form of pencils $\rho R + \sigma S$, with R symmetric and S skew-symmetric. Their approach is the same as the one of Hodge and Pedoe. These authors focus only on the real case, and they just mention that a similar procedure can be followed for complex matrices. This procedure should lead to the TACF. This is precisely the approach followed by Schröder [17], who actually obtains a canonical form for congruence with six types of blocks. Nonetheless, these blocks have a different appearance than the ones of the TACF due to some further modifications addressed by the author. This amounts to, at least, three canonical forms for congruence after Turnbull and Aitken's one (without considering the one of Lee and Weinberg).

The most surprising fact in all this story is that the work of Turnbull and Aitken is not mentioned in the papers [3, 9, 11, 15, 17, 18]. However, most of them cite Thompson's paper [19], where Turnbull and Aitken's book is included among the references (though it is just cited for historical remarks [19, pp. 335, 358]). One of the main purposes of the present paper is to point out the work of Turnbull and Aitken. Nonetheless, the HSCF is the simplest one among all canonical forms mentioned above (including the TACF). Hence, we consider interesting to establish the relationship between this form and the pioneer one of Turnbull and Aitken. Since both forms provide reductions under congruence for matrices in $\mathbb{C}^{n \times n}$, there should be a correspondence between blocks in each of these canonical forms. To explicitly state this correspondence is the second purpose of this paper (Theorem 5). We want to point out that there is a remarkable difference between both canonical forms. In particular, the HSCF consists of just three types of blocks instead of the six types in the TACF. This lower number of blocks makes this form more manageable and it is, for instance, the form that has been used to solve the equation $XA + AX^T = 0$ in [4].

One of the advantages of the TACF of A is that it immediately gives the spectral structure of the associated pencil $\mu A + \lambda A^T$. These pencils are known nowadays as T -palindromic pencils, though this name has been introduced quite recently¹. They are becoming a subject of interest in the linear algebra community mostly motivated by its appearance in several applied problems, like the analysis of rail track vibration produced by high speed trains [6, 13, 14, 16], or the mathematical modeling and numerical simulation of the behavior of periodic surface acoustic wave filters [8, 21]. As noted before, the T -palindromic pencils $\mu A + \lambda A^T$ are in one-to-one correspondence with pencils of the form $\rho R + \sigma S$, with R symmetric and S skew-symmetric. Horn and Sergeichuk derive in [11] canonical forms for congruence of this last kind of pencils from their canonical form for congruence of matrices. Nonetheless, the HSCF does not explicitly show the spectral structure of the pencil. The third contribution of the present paper is to show how to recover explicitly the spectral structure of the associated palindromic pencil $\mu A + \lambda A^T$ from the HSCF of A (Theorem 4).

We close this Introduction by noticing that this is not the first time that the book by Turnbull and Aitken is “rediscovered”. Brualdi published in [2] an “old” proof for the JCF whose basic idea was found in that book. This leads us to ask : What other pearls may be hidden in this old book?

2 Spectral structure of T -palindromic matrix pencils

We recall in this section some basic notions of matrix pencils that will be used along the manuscript. Only in this section we allow for rectangular matrices $A, B \in \mathbb{C}^{m \times n}$. We will say that two matrix pencils $\mu A + \lambda B$ and $\mu C + \lambda D$ are *equivalent* if there exist two nonsingular matrices $P \in \mathbb{C}^{m \times m}$ and $Q \in \mathbb{C}^{n \times n}$ such that

$$P(\mu A + \lambda B)Q = \mu C + \lambda D.$$

Each matrix pencil $\mu A + \lambda B$ is equivalent to a block diagonal pencil of the form

$$\text{diag}(J, N, L_{\varepsilon_1}, \dots, L_{\varepsilon_p}, L_{\eta_1}^T, \dots, L_{\eta_q}^T), \quad (1)$$

where

$$J = \text{diag}(\mu J_{n_1}(\lambda_1) + \lambda I_{n_1}, \dots, \mu J_{n_r}(\lambda_r) + \lambda I_{n_r}), \quad (2)$$

and

$$N = \text{diag}(\mu I_{m_1} + \lambda J_{m_1}(0), \dots, \mu I_{m_s} + \lambda J_{m_s}(0)), \quad (3)$$

with

$$J_d(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{bmatrix}$$

being a $d \times d$ Jordan block associated with $\lambda_i \in \mathbb{C}$. Also

$$L_{\varepsilon_i} = \begin{bmatrix} \mu & \lambda & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \mu & \lambda \end{bmatrix}_{\varepsilon_i \times (\varepsilon_i + 1)}$$

is an $\varepsilon_i \times (\varepsilon_i + 1)$ right singular block, and $L_{\eta_i}^T$ is an $(\eta_i + 1) \times \eta_i$ left singular block. The pencils (2) and (3) comprise the *regular structure* (finite and infinite, respectively) of $\mu A + \lambda B$, and the remaining diagonal blocks of (1) comprise the so-called *singular structure*.

The numbers $-\lambda_1, \dots, -\lambda_r$ are the *finite eigenvalues* of $\mu A + \lambda B$, and it is said that $\mu A + \lambda B$ has an *infinite eigenvalue* if N in (1) is nonempty. Accordingly, we will refer to the diagonal blocks of N as *Jordan blocks associated with the infinite eigenvalue*.

The numbers η_1, \dots, η_q and $\varepsilon_1, \dots, \varepsilon_p$ are known as, respectively, the *left* and *right minimal indices* of $\mu A + \lambda B$. When $\mu A + \lambda B$ is square (and, in particular, when it is T -palindromic), we have $p = q$.

The pencil (1) is known as the *Kronecker Canonical Form* (KCF) of $\mu A + \lambda B$, and it is unique up to permutation of the diagonal blocks. This canonical form comprises the so-called *spectral structure* of $\mu A + \lambda B$, consisting of the eigenvalues of $\mu A + \lambda B$, together with the sizes of the associated Jordan blocks, and the minimal indices. For a proof of the existence and uniqueness of the KCF we refer the reader to the classical

¹ It was introduced privately by D. Steven Mackey in 2003, but the first reference where this name appears in the literature is [6]

Since congruence of matrix pencils is a particular case of equivalence, congruent matrix pencils have the same KCF. Turnbull and Aitken obtained in [20, pp. 138–139] a canonical form for congruence of T -palindromic matrix pencils that exhibits the KCF of the pencil $\mu A + \lambda A^T$. From this canonical form they derive the canonical form for congruence of matrices just by taking $\mu = 1, \lambda = 0$. As a consequence, the TACF of A allows us to know the KCF of $\mu A + \lambda A^T$. On the other hand, Theorem 4 shows also how to recover the KCF of $\mu A + \lambda A^T$ from the HSCF. If we were able to prove that, conversely, the KCF of $\mu A + \lambda A^T$ uniquely determines the congruence class of A then we could establish the relationship between the HSCF and the TACF. This is the aim of the following Lemma.

Lemma 1 *Two square matrices $A, B \in \mathbb{C}^{n \times n}$ are congruent if and only if the pencils $\mu A + \lambda A^T$ and $\mu B + \lambda B^T$ are equivalent.*

Proof. If A, B are congruent then the pencils $\mu A + \lambda A^T$ and $\mu B + \lambda B^T$ are clearly equivalent. Let us prove the converse. Assume that $\mu A + \lambda A^T$ and $\mu B + \lambda B^T$ are equivalent. Then, there exists two nonsingular matrices P, Q such that $PAQ = B$ and $PA^T Q = B^T$, so

$$P \left(\frac{A + A^T}{2} \right) Q = \frac{B + B^T}{2} \quad \text{and} \quad P \left(\frac{A - A^T}{2} \right) Q = \frac{B - B^T}{2}.$$

Now, by the well-know fact that two symmetric matrices are similar if and only if they are congruent (see Lemma III in [20, p. 130]), there is a nonsingular $H \in \mathbb{C}^{n \times n}$ such that

$$H^T \left(\frac{A + A^T}{2} \right) H = \frac{B + B^T}{2} \quad \text{and} \quad H^T \left(\frac{A - A^T}{2} \right) H = \frac{B - B^T}{2},$$

and the result follows. □

Now we are in the position to state the main result of this paper.

Theorem 5 *Each of the blocks in the TA canonical form is congruent to one of the blocks in the HS canonical form as follows.*

- 1) \mathcal{A}_{2k+1} is congruent to $J_{2k+1}(0)$.
- 2) $\mathcal{B}_{2k}(c)$ is congruent to
 - 2a) $J_{2k}(0)$, if $c = 0$, and
 - 2b) $H_{2k}(c)$, if $c \neq 0$.
- 3) \mathcal{C}_{2k+1} is congruent to Γ_{2k+1} .
- 4) \mathcal{D}_{2k} is congruent to $H_{2k}(1)$ (k even).
- 5) \mathcal{E}_{2k} is congruent to Γ_{2k} .
- 6) \mathcal{F}_{2k} is congruent to $H_{2k}(-1)$ (k odd).

Proof. The result is an immediate consequence of Theorem 4, Lemma 1, and the relationship between canonical blocks in the TA canonical form of A and the KCF of $\mu A + \lambda A^T$, given by [20, pp. 135–139]

- $\mu \mathcal{A}_{2k+1} + \lambda \mathcal{A}_{2k+1}^T$ is equivalent to $L_{2k} \oplus L_{2k}^T$.
- $\mu \mathcal{B}_{2k+1}(c) + \lambda \mathcal{B}_{2k+1}(c)^T$ is equivalent to $(\mu J_k(c) + \lambda I_k) \oplus (\mu J_k(1/c) + \lambda I_k)$ ($0 \neq c \neq \pm 1$).
- $\mu \mathcal{B}_{2k+1}(0) + \lambda \mathcal{B}_{2k+1}(0)^T$ is equivalent to $(\mu J_k(0) + \lambda I_k) \oplus (\mu I_k + \lambda J_k(0))$.
- $\mu \mathcal{C}_{2k+1} + \lambda \mathcal{C}_{2k+1}^T$ is equivalent to $\mu J_{2k+1}(1) + \lambda I_{2k+1}$.
- $\mu \mathcal{D}_{2k+1} + \lambda \mathcal{D}_{2k+1}^T$ is equivalent to $(\mu J_k(1) + \lambda I_k) \oplus (\mu J_k(1) + \lambda I_k)$.
- $\mu \mathcal{E}_{2k} + \lambda \mathcal{E}_{2k}^T$ is equivalent to $\mu J_{2k}(-1) + \lambda I_{2k}$.
- $\mu \mathcal{F}_{2k} + \lambda \mathcal{F}_{2k}^T$ is equivalent to $(\mu J_k(-1) + \lambda I_k) \oplus (\mu J_k(-1) + \lambda I_k)$.

□

Theorem 5 also allows us to explain the difference in the number of blocks between the TACF and the HSCF. This difference comes from the size of type 0 and type I blocks and also on the values associated with type II blocks in the HSCF. More precisely, type 0 blocks give rise to two different types of blocks in the TACF depending on whether the size of the type 0 block is even or odd, and the same happens with the type I blocks (cases 1), 2a) and 3), 5) in Theorem 5). On the other hand, type II blocks associated with, respectively, $c = 1$, $c = -1$ and $c \neq \pm 1$, give rise to three different types of blocks in the TACF (cases 2b), 4) and 6)).

The canonical form for congruence of T -palindromic pencils introduced by Turnbull and Aitken in pages 138–139 contains some redundant blocks. First, their type 4 blocks leading to type (d) blocks for the congruence form of matrices include the following $2k \times 2k$ blocks

$$\left[\left[\begin{array}{ccc} 0 & & \mu + \lambda \\ & \mu + \lambda & \mu - \lambda \\ & \ddots & \ddots \\ \mu + \lambda & \mu - \lambda & 0 \end{array} \right] \setminus \left[\begin{array}{ccc} 0 & & \mu + \lambda \\ & \mu + \lambda & -\mu + \lambda \\ & \ddots & \ddots \\ \mu + \lambda & -\mu + \lambda & 0 \end{array} \right] \right]. \quad (5)$$

These blocks correspond to a pair of Jordan blocks with size k associated with the eigenvalue -1 . But, when k is odd, this type of block is contained in the type 3 blocks, which are the blocks with odd size associated with the eigenvalue -1 . More precisely, (5) is congruent to $[P_k \setminus P_k]$, where P_k is the $k \times k$ pencil

$$P_k = \left[\begin{array}{ccc} 0 & & \mu + \lambda \\ & \mu + \lambda & -\mu + \lambda \\ & \ddots & \ddots \\ \mu + \lambda & \mu - \lambda & 0 \end{array} \right].$$

Second, the type 6 blocks of size $2k \times 2k$ (leading to the type (f) blocks in the canonical form for matrices)

$$\left[\left[\begin{array}{ccc} 0 & & -\mu + \lambda \\ & -\mu + \lambda & \mu + \lambda \\ & \ddots & \ddots \\ -\mu + \lambda & \mu + \lambda & 0 \end{array} \right] \setminus \left[\begin{array}{ccc} 0 & & \mu - \lambda \\ & \mu - \lambda & \mu + \lambda \\ & \ddots & \ddots \\ \mu - \lambda & \mu + \lambda & 0 \end{array} \right] \right] \quad (6)$$

with k even, corresponding to a pair of Jordan blocks of even size associated with the eigenvalue 1, are included in the type 5 blocks. More precisely, (6) is congruent to $[Q_k \setminus Q_k]$, where Q_k is the $k \times k$ pencil

$$Q_k = \left[\begin{array}{ccc} 0 & & \mu - \lambda \\ & \mu - \lambda & \mu + \lambda \\ & \ddots & \ddots \\ -\mu + \lambda & \mu + \lambda & 0 \end{array} \right].$$

Despite this redundancy in the canonical form for T -palindromic pencils, when the authors derive the canonical form for matrices by taking $\mu = 1$, $\lambda = 0$, the redundant blocks disappear without any further comment.

Acknowledgements The author wishes to thank Froilán M. Dopico for posing the main question addressed in the paper, namely, to state the relationship between the TACF and the HSCF.

References

1. T. J. BROMWICH. *Quadratic forms and their classification by means of invariant factors*. Cambridge University Press, Cambridge, 1906.
2. R. A. BRUALDI. *The Jordan canonical form: an old proof*. Amer. Math. Monthly, 94 (1987) 257–267.
3. B. CORBAS, G. D. WILLIAMS. *Bilinear forms over an algebraically closed field*. J. Pure Appl. Algebra 165 (3) (2001) 225–266.
4. F. DE TERÁN, F. M. DOPICO. *The solution of the equation $XA + AX^T = 0$ and its application to the theory of orbits*. Linear Algebra Appl. 434 (2011) 44–67.

5. F. R. GANTMACHER. *The Theory of Matrices*. Chelsea, New York, 1959.
6. A. HILLIGES, C. MEHL, V. MEHRMANN. *On the solution of palindromic eigenvalue problems*. Proceedings of the 4th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS), JyvÄaskylÄa, Finland, 2004. CD-ROM.
7. W. V. D HODGE, D. PEDOE. *Methods of Algebraic Geometry*. Cambridge University Press, Cambridge, 1947.
8. M. HOFER, N. FINGER, J. SCHBERL, S. ZAGLMAYR, U. LANGER, R. LERCH. *Finite element simulation of wave propagation in periodic piezoelectric SAW structures*. IEEE Trans. Ultrason. Ferroelectr. Freq. Control 53 (2006) 1192–1201.
9. R. A. HORN, V. V. SERGEICHUK. *Congruence of a square matrix and its transpose*. Linear Algebra Appl., 389 (2004) 347–353.
10. R. A. HORN, V. V. SERGEICHUK. *A regularization algorithm for matrices of bilinear and sesquilinear forms*. Linear Algebra Appl., 412 (2006) 380–395.
11. R. A. HORN, V. V. SERGEICHUK. *Canonical forms for complex matrix congruence and *congruence*. Linear Algebra Appl., 416 (2006) 1010–1032.
12. R. A. HORN, V. V. SERGEICHUK. *Canonical matrices of bilinear and sesquilinear forms*. Linear Algebra Appl., 428 (2008) 193–223.
13. T.-M. HUANG, W.-W. LIN, J. QIANG. *Structure-preserving algorithms for palindromic quadratic eigenvalue problems arising from vibration of fast trains*. SIAM J. Matrix Anal. Appl., 30 (4) (2009) 1566–1592.
14. I. C. F. Ipsen. *Accurate eigenvalues for fast trains*. SIAM News, 37 (9) (2004) 1–2.
15. J. M. LEE, D. A. WEINBERG. *A note on the canonical form for matrix congruence*. Linear Algebra Appl., 249 (1996) 207–215.
16. D. S. MACKEY, N. MACKEY, C. MEHL, V. MEHRMANN. *Structured polynomial eigenvalue problems: Good vibrations from good linearizations*. SIAM J. Matrix Anal. Appl., 28 (2006) 1029–1051.
17. C. SCHRÖDER. *A canonical form for palindromic pencils and palindromic factorizations*. Preprint. Available at http://www.matheon.de/preprints/3822_Sch06.pdf
18. V. V. SERGEICHUK. *Classification problems for system of forms and linear mappings*. Math. USSR, Izvestiya 31 (3) (1998) 481–501.
19. R. C. THOMPSON. *Pencils of Complex and Real Symmetric Matrices*. Linear Algebra Appl., 147 (1991) 323–371.
20. H. W. TURNBULL, A. C. AITKEN. *An introduction to the Theory of Canonical Matrices*. Dover, New York, 1961 (First published by: Blackie & Son Limited, 1932).
21. S. ZAGLMAYR. *Eigenvalue problems in SAW-filter simulations*. Diplomarbeit, Inst. of Comp. Mathematics, J. Kepler Univ. Linz, Austria, 2002.