# A CONCISE PROOF OF SYLVESTER'S THEOREM

By

F. Csáki

Department of Automation, Technical University, Budapest

(Received August 11, 1969)

## 1. Introduction

In the stability investigations of nonlinear systems the LYAPUNOV technique is one of the fundamental methods. As it is well known the LYAPUNOV method decides upon the stability by choosing a positive definite scalar function $V(\mathbf{x})$ with state vector argument $\mathbf{x}$. First, if the time derivative of this function $V(\mathbf{x})$ is negative definite along the trajectory, then the system described by the differential equation $\dot{\mathbf{x}} = f(\mathbf{x})$ is asymptotically stable, secondly, if $V(\mathbf{x})$ is negative semidefinite and $\dot{V}(\mathbf{x}) = 0$ does not constitute a trajectory then the system is also asymptotically stable, while in the reversed case, when $\dot{V}(\mathbf{x}) = 0$ supplies a trajectory, it is only stable, finally, if $V(\mathbf{x})$ is indefinite, a new $V(\mathbf{x})$ function may be tried.

Here only autonomous systems are mentioned, but the LYAPUNOV method can also be extended to nonautonomous systems with differential equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$. Very often quadratic forms are chosen as LYAPUNOV functions $V(\mathbf{x})$ or as the time rate of changes $\dot{V}(\mathbf{x})$, especially when the system is linear and its differential equation is of the form $\dot{\mathbf{x}} = \mathbf{A}\,\mathbf{x}$. In such cases SYLVESTER's theorem is of particular importance for the decision of the positive definiteness of quadratic forms. By the way, negative definiteness can equally well be determined by introducing a sign change into the quadratic form investigated. If the quadratic form is positive definite the matrix figuring in the form is also said positive definite and vice versa. A manifest advantage of SYLVESTER's theorem is that it yields an indirect method in which the determination of the eigenvalues of the matrix in question can completely be avoided.

It is rather curious that in the textbooks treating the LYAPUNOV method, [e.g. 6, 8, 9, 10, 12], the proof of SYLVESTER's theorem is generally omitted, although this is a salient point of the LYAPUNOV method. If there is exceptionally a proof, a somewhat lengthy and awkward derivation is only published [2, 3, 4, 5, 7], without speaking of the case when a completely incorrect proof is given [11] in an otherwise good book. In what follows, this difficulty will be eliminated by presenting a brief and concise proof of the SYLVESTER theorem.

## 2. Preliminary remarks

Before going into the details of the proof, two definitions and two auxiliary theorems are mentioned and some remarks are also made.

*Definition* 1. A quadratic form is a homogeneous quadratic polynomal of the variables $x_1, x_2, \ldots, x_n$ or the vector $\mathbf{x}$.

*Auxiliary theorem* 1. Every asymmetric real quadratic form can be expressed as a real *symmetric* quadratic form.

*Proof.* Let us have the asymmetric quadratic form

$$Q'(\mathbf{x}) = \mathbf{x}^{\mathrm{T}} \mathbf{P}' \mathbf{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} p'_{ij} x_i x_j \tag{1}$$

with an asymmetric quadratic matrix $P'_{ji} \neq$, that is, with matrix elements $p'_{ji}$, if $i \neq j$. (Here the superscript suffix $\mathbf{T}$ refers to the transposed of a matrix.) Then taking the symmetric quadratic form

$$Q(\mathbf{x}) = \mathbf{x}^{\mathrm{T}} \mathbf{P} \mathbf{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} x_i x_j \tag{2}$$

with symmetric quadratic matrix $\mathbf{P}$, that is, with matrix elements $p_{ij} = p_{ji}$; $i, j = 1,2, \ldots, n$, it can readily be shown that

$$Q'(\mathbf{x}) = Q(\mathbf{x}) \tag{3}$$

if

$$\mathbf{P} = \frac{1}{2} (\mathbf{P}' + \mathbf{P}'^{\mathrm{T}}) \tag{4}$$

or

$$p_{ij} = p_{ji} = \frac{1}{2} (p'_{ij} + p'_{ji}); \quad i, j = 1, 2, \ldots, n. \tag{5}$$

The asymmetric quadratic form $Q'(\mathbf{x})$ can namely be written as

$$Q'(\mathbf{x}) = \frac{1}{2} [\mathbf{x}^{\mathrm{T}} \mathbf{P}' \mathbf{x} + \mathbf{x}^{\mathrm{T}} \mathbf{P}'^{\mathrm{T}} \mathbf{x}] = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{1}{2} (p'_{ij} + p'_{ij}) x_i x_j \tag{6}$$

which yields the desired result, indeed. Thus there is no loss of generality when assuming in the following treatise the quadratic form to be symmetric.

*Definition* 2. Two real quadratic forms $Q_1(\mathbf{x}) = \mathbf{x}^{\mathrm{T}} \mathbf{P}_1 \mathbf{x}$ and $Q_2(\mathbf{x}) = \mathbf{y}^{\mathrm{T}} \mathbf{P}_2 \mathbf{y}$ are said to be *similar* or, from the point of view of definiteness, *equivalent* if the one can be transformed into the other through a nonsingular transformation of co-ordinates. By the way, if the transformation is orthogonal, that is the transformation matrix $\mathbf{T}$ is an orthogonal matrix with the property $\mathbf{T}^{-1} = \mathbf{T}^{\mathrm{T}}$, or in other words $\mathbf{T}^T\mathbf{T} = \mathbf{I}$, then the two forms are said to be *orthogonally similar* and, from the definiteness point of view *orthogonally equivalent.* When the eigenvalues of the matrices in the two quadratic forms are the same then the forms are said to be *congruent* and *orthogonally congruent*, respectively.

*Auxiliary theorem* 2. [e.g. 2, 3, 4, 5]. Every real symmetric quadratic form as given in (2) can be transformed by an appropriate orthogonal transformation $\mathbf{T}$ into the canonical form

$$Q_c(\mathbf{v}) = \mathbf{v}^T \mathbf{\Lambda} \mathbf{v} \tag{7}$$

where

$$\mathbf{\Lambda} = \operatorname{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n] \tag{8}$$

and $\lambda_i$ $(i = 1,2, \ldots, n)$ are the real eigenvalues of the symmetric matrix $\mathbf{P}$. (As it is well known, the eigenvalues of a real symmetric matrix are real quantities, see e.g. [4].) For the simple case of distinct eigenvalues, the proof of the theorem is easy. Suppose that $\mathbf{p}_i$ $(i = 1,2, \ldots, n)$ are the corresponding eigenvectors, that is, the following relations are valid:

$$\mathbf{P}\,\mathbf{p}_i = \lambda_i\,\mathbf{p}_i; \quad i = 1,2, \ldots, n \tag{9}$$

and

$$||\mathbf{p}_i|| = \mathbf{p}_i^T\,\mathbf{p}_i = 1; \quad i = 1,2, \ldots, n \tag{10}$$

As its is well known, the eigenvectors are orthogonal to each other:

$$\mathbf{p}_i^T\,\mathbf{p}_j = 0 \quad i \neq j \tag{11}$$

Thus, the appropriate transformation matrix is

$$\mathbf{T}_0 = [\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n] \tag{12}$$

which, by property (10) and (11), is easily seen to be orthogonal: $\mathbf{T}_0^T\mathbf{T}_0 = \mathbf{I}$. Then, let us introduce the new variable $\mathbf{v}$ by the following relationships:

$$\mathbf{x} = \mathbf{T}_0\,\mathbf{v} \quad \text{or} \quad \mathbf{v} = \mathbf{T}_0^{-1}\mathbf{x} = \mathbf{T}_0^T\,\mathbf{x} \tag{13}$$

Applying this orthogonal transformation the canonical form (7) results:

$$Q_c(\mathbf{v}) = \mathbf{v}_0^T\,\mathbf{T}_0^T\,\mathbf{P}\mathbf{T}\mathbf{v} \tag{14}$$

As, on the one hand

$$\mathbf{T}_0\mathbf{\Lambda} = [\mathbf{p}_1\,\lambda_1, \mathbf{p}_2\,\lambda_2\,, \ldots, \mathbf{p}_n\lambda_n] \tag{15}$$

while on the other hand, according to (12) and (9):

$$\mathbf{P}\,\mathbf{T}_0 = [\mathbf{P}\,\mathbf{p}_1, \mathbf{P}\mathbf{p}_2\,, \ldots, \mathbf{P}\mathbf{p}_n] = \\ = [\lambda_1\mathbf{p}_1, \lambda_2\mathbf{p}_2\,, \ldots, \lambda_n\mathbf{p}_n] \tag{16}$$

consequently,

$$\mathbf{T}_0\mathbf{\Lambda} = \mathbf{P}\,\mathbf{T}_0 \tag{17}$$

or

$$\mathbf{\Lambda} = \mathbf{T}_0^{-1}\,\mathbf{P}\mathbf{T}_0 = \mathbf{T}_0^{\mathrm{T}}\,\mathbf{P}\mathbf{T}_0\,. \tag{18}$$

Thus, every symmetric quadratic form can be transformed into the canonical form, or in other words, every symmetric quadratic form is orthogonally congruent with the canonic form. The definiteness of the quadratic form can be judged and decided from the canonic form. If every eigenvalue is positive (negative) both quadratic forms are positive (negative) definite, if some eigenvalues are zero, while the others are all positive (negative), then both quadratic forms are positive (negative) semidefinite, finally, if there are positive and negative eigenvalues as well, then both quadratic form are indefinite.

Unfortunately, the numerical determination of the eigenvalues of a matrix of large dimension is a very cumbersome matter. This is the reason why a criterion in terms of eigenvalues is not useful in applications. Thus, the previous criteria of definiteness are merely of theoretical value. Generally, for analytic and computional purposes, the more useful SYLVESTER's theorem are applied.

In the following proof of SYLVESTER's theorem also a diagonal form different from the canonical form will be employed. To obtain it, a special nonorthogonal similarity transform will be used, which, according to the author's knowledge, is first proposed here, so it pretends to priority.

### 3. The statement of Sylvester's theorem

According to SYLVESTER's theorem, necessary and sufficient condition of the positive definiteness of the real symmetric quadratic form

$$Q(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{P}\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}\,x_i\,x_j = Q(x_1, x_2, \ldots x_n)$$

with $p_{ij} = p_{ji}$, is that all the principal minors of the matrix $\mathbf{P}$ be positive:

$$P_1 = p_{11} > 0\,;$$

$$P_2 = \begin{vmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{vmatrix} > 0\,;$$

$$\text{-------------------}$$

$$P_n = \begin{vmatrix} p_{11} & p_{22} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{vmatrix} > 0\,.$$

It is recapitulated that in the expression of $Q(\mathbf{x})$, $\mathbf{x}^T = [x_1, x_2, \ldots, x_n]$ means a transposed column vector, that is, a row vector, while

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1, x_2, \ldots, x_n]^T$$

is the original column vector. Finally, $\mathbf{P}$ is a nxn symmetric matrix with elements $p_{ij} = p_{ji}$.

## 4. The proposed verification of Sylvester's theorem

Now, a concise proof of SYLVESTER's theorem will here be given as follows. Let $P_r$ denote the principal minors in the determinant $|\mathbf{P}|$ of the matrix $\mathbf{P}$, that is,

$$P_r = \begin{vmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & & & \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{vmatrix}, \quad (r=1, 2, \ldots, n). \tag{19}$$

Without loss of generality $P_r \neq 0$ $(r = 1,2, \ldots, n)$ can be assumed, as the necessary condition of the positive definiteness of the matrix $\mathbf{P}$ and the $n$ variable quadratic form $Q$ is that matrix $\mathbf{P}$ must be of rank $n$, that is, $P_n \neq 0$. Diminishing the number of variables successively by one, $P_{n-1} \neq 0$, $P_{n-2} \neq 0$ etc. follows. Furthermore, let $P^r_{ij}$ denote the cofactor pertaining to an arbitrary $p_{ij}$ element in the principal minor $P_r$ of the determinant $P_n = |\mathbf{P}|$. Here in $P^r_{ij}$, the superscript $r$ refers to the principal minor $P_r$. The cofactor $P^r_{ij}$ of the element $p_{ij}$ is $(-1)^{i+j}$ times the determinant of the minor $P_r$ formed by deleting the $i$th row and the $i$th column from $P_r$, that is, $P^r_{ij}$ is the minor with a corresponding sign.

Now, let us introduce a new state vector $\mathbf{v}$ by the nonsingular transformation

$$\mathbf{x} = \mathbf{T}\mathbf{v} \qquad \text{or} \qquad \mathbf{x}^T = \mathbf{v}^T\mathbf{T}^T \tag{20}$$

which, in contrast to (13), is not orthogonal anymore.

Assuming a symmetric matrix $\mathbf{P}$, the transformation matrix proposed should be

$$\mathbf{T} = \begin{bmatrix} P^1_{11}/P^1_{11} & P^2_{21}/P^2_{22} & \ldots & P^n_{n1}/P^n_{nn} \\ 0 & P^2_{22}/P^2_{22} & \ldots & P^n_{n2}/P^n_{nn} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & P^n_{nn}/P^n_{nn} \end{bmatrix} \tag{21}$$

whereas the transposed matrix is, with the symmetry conditions taken into consideration,

$$\mathbf{T}^{\mathrm{T}} = \begin{bmatrix} P^1_{11}/P^1_{11} & 0 & \dots & 0 \\ P^2_{12}/P^2_{22} & P^2_{22}/P^2_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ P^n_{1n}/P^n_{nn} & P^n_{2n}/P^n_{nn} & \dots & P^n_{nn}/P^n_{nn} \end{bmatrix} \tag{22}$$

where $P^1_{11} = 1$; $P^2_{12} = P^2_{21} = -p_{21} = -p_{12}$; $P^2_{22} = p_{11}$, etc. After introducing $P_0 = 1$, the symmetric quadratic form as given in (2) can be converted into the following diagonal form:

$$Q(\mathbf{Tv}) = \mathbf{v}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{PTv} = \mathbf{v}^{\mathrm{T}}\mathbf{P}_0\,\mathbf{v} \tag{23}$$

where

$$\mathbf{P}_0 = \mathbf{T}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}} = \mathrm{diag}\,[P_1/P_0, P_2/P_1, \dots P_{n-1}/P_{n-2}, P_n/P_{n-1}] \tag{24}$$

because $P^r_{rr} = P_{r-1}$ for $r = 1,2, \dots, n$.
Thus, the diagonal form can be expressed as

$$Q_d(\mathbf{v}) = Q(\mathbf{Tv}) = \frac{P_1}{P_0}v_1^2 + \frac{P_2}{P_1}v_2^2 + \dots + \frac{P_{n-1}}{P_{n-2}}v_{n-1}^2 + \frac{P_n}{P_{n-1}}v_n^2. \tag{25}$$

The last relation clearly reveals the sufficiency of SYLVESTER's theorem. If each principal minor $P_r$ $(r = 1,2, \dots, n)$ is positive, then the quadratic form (25) is really positive definite.

The necessity of the theorem is similarly thrown into relief: If not all of the principal minors are positive, zero minors being excluded beforehand, then some terms in the diagonal form (25) turn to negative. Selecting their associated variables sufficiently small as compared to the others, makes $Q$ positive, while sufficiently large variables make it negative. Consequently, when not all the principle minors are positive $\mathbf{P}$ cannot be positive definite.

Thus the proof of the theorem is complete.

## 5. Some supplementary remarks

*Remark* 1.: When obtaining the diagonal form (25) by applying the (23) or (24), we have widely utilized the determinant expansion relations:

$$\sum_{k=1}^{r} p_{ik}\,P^r_{jk} = \begin{cases} P_r, & i = j \\ 0, & i \neq j \end{cases} \tag{26 }$$

or

$$\sum_{k=1}^{r} p_{ki}\,P^r_{kj} = \begin{cases} P_r, & i \neq j \\ 0, & i \neq j \end{cases} \qquad (27$$

*Remark* 2. The diagonal form (25) is very similar to the well known Jacobian form

$$Q_J(\mathbf{y}) = \frac{P_0}{P_1} y_1^2 + \frac{P_1}{P_2} y_2^2 + \ldots + \frac{P_{n-1}}{P_n} y_n^2 \tag{28}$$

Of course, the same conclusions can be drawn from both forms, but the relations (25) and (28) are obtained in quite different manners.

As it is remarked in reference [5], the expression (28) can be obtained by the transformation

$$\mathbf{y} = \mathbf{G}\mathbf{x}\,, \quad \mathbf{y}^{\mathrm{T}} = \mathbf{x}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}} \tag{29}$$

where in our notations

$$\mathbf{G} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ 0 & p_{22}^1 & \cdots & p_{2n}^1 \\ \vdots & & & \vdots \\ 0 & 0 & \ldots & p_{nn}^{n-1} \end{bmatrix}. \tag{30}$$

Here

$$p_{ij}^k = p_{ik}^{(k-1)} - \frac{p_{ik}^{(k-1)}}{p_{kk}^{(k-1)}} p_{kj}^{(k-1)} \tag{31}$$

are the coefficients of the GAUSS elimination algorithm which can also be determined as:

$$p_{ij}^k = \frac{P\begin{pmatrix}1, 2, \ldots, k, i\\ 1, 2, \ldots, k, j\end{pmatrix}}{P\begin{pmatrix}1, 2, \ldots, k\\ 1, 2, \ldots, k\end{pmatrix}} \tag{32}$$

where

$$P\begin{pmatrix} i_1, i_2, \ldots, i_q \\ j_1, j_2, \quad , j_q \end{pmatrix} = \begin{vmatrix} p_{i_1 j_1} & p_{i_1 j_2} & \cdots & p_{i_1 j_q} \\ p_{i_2 j_1} & p_{i_2 j_2} & \cdots & p_{i_2 j_q} \\ \vdots & \vdots & & \vdots \\ p_{i_q j_1} & p_{i_q j_2} & \cdots & p_{i_q j_q} \end{vmatrix}. \tag{33}$$

Thus,

$$Q_J(\mathbf{y}) = \mathbf{y}^{\mathrm{T}}[\mathbf{G}^{\mathrm{T}}]^{-1}\mathbf{P}\mathbf{G}^{-1}\mathbf{y} = \mathbf{y}^{\mathrm{T}}[\mathbf{G}^{-1}]^{\mathrm{T}}\mathbf{P}\mathbf{G}^{-1}\mathbf{y}\,. \tag{34}$$

Because here inverse matrices are encountered the latter method does not seem simpler than the derivation method proposed in this paper.

*Remark* 3. The number of positive, negative and possibly zero terms in diagonal form does not depend upon the determination method of the particular form. This is the SYLVESTER—JACOBI law of inertia [4, 5].

The rank of the quadratic form is $r = \pi + \nu$ where $\pi$ is the number of positive terms, while $\nu$ that of the negative terms. Similarly, the signature is $\sigma = \pi - \nu$ which is an other invariant characteristic of the quadratic forms. After a theorem of JACOBI, $\nu$ can be determined as the number of sign changes in the sequence $P_0, P_1, P_2, \ldots, P_r$, (with $P_0 = 1$), while $\pi$ is the number of sign invariance in the same sequence.

*Remark* 5. With the notation introduced in (33) the cofactor $P_{ij}^r$ can be expressed as

$$P_{ij}^r = (-1)^{i+j} P \begin{pmatrix} 1,2, \ldots, (i-1), (i+1), \ldots, r \\ 1,2, \ldots, (j-1), (j+1), \ldots, r \end{pmatrix}$$

## Summary

The paper proposes a concise derivation and, according to the author's knowledge, an original proof of the SYLVESTER theorem playing an important role when determining the positive definiteness of quadratic forms. The proposed method is compared with some other methods.

## References

1. LEFSCHETZ, S.: Differential Equations. Geometric Theory. Interscience Publishers Inc. New York, 1957.
2. Гантмахер, Ф. Р.: Теория матриц. Гостехиздат, Москва, 1953.
3. GANTMACHER, F. R.: The Theory of Matrices. Chelsea Publishing Co. New York, 1959.
4. BELLMAN, R.: Introduction to Matrix Analysis, McGraw Hill, 1960.
5. Гантмахер, Ф. Р.: Теория матриц. Наука, Москва, 1966.
6. GIBSON, J. E.: Nonlinear Automatic Control. McGraw Hill Book Co. New York, 1963.
7. Четаев, Н. Г.: Устойчивость движения. Наука, Москва, 1965.
8. TIMOTHY, L. K.—BONA, B. E.: State Space Analysis: An Introduction. McGraw Hill Book Co. New York, 1968.
9. SCHULTZ, D. G.—MELSA, J. L.: State Functions and Linear Control Systems. McGraw Hill Book Co. New York, 1967.
10. Красовский, Н. Н.: Теория управления движением. Наука, Москва, 1968.
11. DERUSSO, P. M.—ROY, R. J.—CLOSE, CH. M.: State Variables for Engineers. John Wiley et Sons, Inc. New York, 1966.
12. SAATY, TH. L.—BRAM, J.: Nonlinear Mathematics. McGraw Hill Book Co. New York, 1964.

Prof. Dr. Frigyes CSÁKI, Budapest XI., Egry József u. 18. Hungary