ISAS - INTERNATIONAL SCHOOL
FOR ADVANCED STUDIES

A contribution to the Sturmian theory for second-order non-symmetric systems

Thesis submitted for the degree of

"Doctor Philosophiæ"

CANDIDATE

Egone Pertotti

SUPERVISOR

G.Vidossich

EXTERNAL REFEREE

S.Ahmad

November 1995

SISSA - SCUOLA
INTERNAZIONALE
SUPERIORE
DI STUDI AVANZATI
TRIESTE
Strada Costiera 11

TRIESTE

Introduction

Sturmian theory deals with the qualitative properties of higher order linear differential equations and systems, related to the asymptotic behavior of solutions. More precisely, it treats oscillation, separation and comparison of solutions.

If $u(x)$ and $v(x)$ are nontrivial solutions of

$$-u'' + pu = 0$$
$$-v'' + qv = 0$$

for constant p and q, then

(i) $q < p \Rightarrow$ successive zeros of $u(x)$ are separated by a zero of $v(x)$.

(ii) $p \geq 0 \Rightarrow u(x)$ has at most one zero on any interval $[x_0, \infty)$.

(iii) $q < p \Rightarrow v(x)$ has a zero on every interval $[x_0, \infty)$.

These observations constitute the simplest examples of separation , nonoscillation, and oscillation theorems, respectively. They are also the motivation for a large and growing body of mathematical literature.

The theory originated in the last century with the famous paper by Sturm [16] dealing with oscillation and comparison theorems for linear second order scalar equations. Closely associated to Sturm's work was that of Liouville [8] studying the asymptotic form of solutions to second-order linear ODEs in a characteristic parameter, ensuing a BVP known now as the Sturm-Liouville problem. A great emphasis on this topic was given by Bocher's thesis [3] written under the direction of F.Klein. In his research on problems in potential theory, Klein was led to the question of when certain linear homogeneous second-order ODEs involving two parameters had for two given non-overlapping intervals in **R** a pair of solutions which possessed on the respective intervals a prescribed number of zeros. The basic work [6] of Hilbert in the first decade of the twentieth century was fundamental for the study of BVPs associated with self-adjoint differential systems, both with regard to the development of the theory of integral equations and in connection with the interrelations between the Calculus of Variations and the characterization of eigenvalues and eigensolutions of these systems. Moreover, in subsequent years the significance of the calculus of variations for such BVPs was emphasized by G.A.Bliss and M.Morse. In particular, Morse showed in his basic 1930 paper that [9] that variational principles provided an appropriate environment for the extension to selfadjoint differential systems of the classical Surmian theory.

Morse [9] was the first to extend the theory to systems of ordinary differential equations, treating exclusively selfadjoint systems. For a panorama of the results and literature on scalar equations or selfadjoint systems, confront the books by Coppel [4], Morse [10] and Reid [14], [15]. Recently, Ahmad-Lazer [2] initiated the extension of the theory to nonselfadjoint systems. Again, various papers appeared on the nonselfadjoint case, cf. Ahmad [1] for a comprehensive bibliography.

It should also be mentioned that Sturmian theory generated the development of some new mathematical fields. Of these, the most popular is the study of the Riccati equation, that is useful also in numerical analysis. Less popular is transformation theory, one of the key tools in oscillation theory, e.g. Goff-St.Mary [5] for a recent account of the literature and the introduction of the Bohl transformation of a pair of Hermitian matrices by the use of the matrix analogues of the sine and cosine functions.

The purpose of this thesis is to present some results about the Sturmian theory of second-order systems

$$x'' + A(t)x = 0 \tag{1}$$

where $A(t)$ is a given continuous $n \times n$ matrix function. We are mainly interested in the case that $A(t)$ is not symmetric, so that neither the spectral theory of self-adjoint operators nor the methods of the Calculus of Variations are applicable.

In connection with oscillation theory, whose goal is to detect the distribution and the multiplicity of zeros of the solutions, the notion of conjugate and focal points are of fundamental importance because they provide a link between oscillation and BVPs. We recall their definition. Let a, b be real numbers. We say that b is a **conjugate point** of a if there exists a nontrivial solution $x(t)$ of (1) such that $x(a) = 0 = x(b)$. We say that b is a **focal point** of a if $b > a$ and there exists a nontrivial solution $x(t)$ of (1) such that $x'(a) = 0 = x(b)$. In other words, conjugate points correspond to the Picard BVP, while focal points correspond to the Nicoletti BVP.

Another concept of great importance is the matrix analogue of (1), i.e. the equation

$$X'' + A(t)X = 0 \tag{2}$$

where the solution $X(t)$ is an $n \times n$ matrix for each t. In this connection, we denote by

$$X_A(t,\tau) \quad \text{and} \quad Y_A(t,\tau)$$

the solutions to (2) that satisfy the following initial conditions:

$$\begin{cases} X_A(\tau,\tau) = 0_n \\ X'_A(\tau,\tau) = I_n \end{cases}$$

and

$$\begin{cases} Y_A(\tau,\tau) = 0_n \\ Y'_A(\tau,\tau) = 0_n \end{cases}$$

respectively. The interest in these special solutions lies on the following well-known result:

Theorem (Reid [15]): *b is a conjugate point to a iff* $\det X_A(a,b) = 0$ *and b, $a < b$, is a focal point of a iff* $\det Y_A(a,b) = 0$.

This thesis is divided into three chapters, each corresponding to one of the three papers Pertotti [11], [12] and Pertotti-Geaman [13].

The starting point of our research is a deep study of the "elementary" algebraic properties of X_A and Y_A and their interelationships. Perhaps it is worth while remarking that the dimension of the vector space of $n \times n$ matrices is n^2 ,so that the dimension of the space of solutions to (2) is $2n^2$. This means that, in the matrix case, X_A and Y_A do not constitute a basis for the space of solution, contrarily to the scalar case.

These properties of X_A and Y_A are used in connection with ordered Banach space techniques (not only the Krein-Rutman theory!) in the proofs of the main results. These are the following:

- in ch.I, we provide a partial answer to a conjecture of Ahmad [1] as well as conditions for the existence of solutions whose components have a prescribed sign;

- in ch.2, we characterize the existence of symmetric matrix solutions to (2). This is of interest in the oscillation theory and in the construction of the Bohl transformation, cf. Goff-St.Mary [5];

- in ch.3, we reduce the question of the existence of conjugate and focal points to the analysis of a scalar equation, in analogy with the Liouville theorem about the Wronskian of a matrix solution.

References.

1. S.AHMAD : On Sturmian theory for second order systems, *Proc. Amer. Math. Soc.* 87 (1983), 661-665

2. ___and A.C.LAZER: An N-dimensional extension of the Sturm separation and comparison theory to a class of nonselfadjoint systems , *SIAM J. Math. Anal.* 9 (1978), 1137-1150.

3. M. BOCHER *Über die Reihenentwickelungen der Potentialtheorie,* (Dissertation, Gottingen, 1891).

4. W.COPPEL: *Disconjugacy* , Springer L.N.220, New York, 1971.

5. S.GOFF and D.F.ST.MARY: The Bohl transformation and oscillation of a linear differential system, *SIAM J.Math.Anal.*20 (1989), 215-221.

6. D.HILBERT: *Grundzüge einer allgemeiner Theorie der linearen Integralgleichungen*, Teubner, Leipzig, 1912; reprinted by Chelsea, New York, 1952.

7. K.KREITH: *Oscillation Theory*, Springer L.N.324, 1973, New York

8. J.LIOUVILLE: *Sur le developpement des fonctions ou parties de fonctions en series dont les divers termes sont assujettis a satisfaire a una meme equation differentielles du second ordre contenant un parametre variable*, JMPA, 1(1836), 253-265; 2(1837) ,16-35; 418-436.

9. M.MORSE: A generalization of the Sturm separation and comparison theorems in n-space, *Math.Ann.*103 (1930), 52-69.

10. ___: *Variational analysis: critical extremals and Sturmian extension*, John Wiley, New York, 1973.

11. E.PERTOTTI: On Sturmian theory for nonselfadjoint second-order systems, *Dynamic System Appl.*, 4 (1995), 189-198.

12. ___: Characterizing the existence of symmetric solutions for matrix equations $X'' + AX = 0$, *Dynamic System Appl.*

13. ___ and V.GEAMAN: Contribution to the theory of conjugate and focal points of second-order systems, Colloquium on Oscillation Theory, Plovdiv, Bulgaria 13-17 August 1995.

14. W.T.REID: *Ordinary differential equations*, John Wiley, New York, 1971.

15. ___: *Sturmian theory for ordinary differential equations*, Springer, New York, 1980.

16. J.C.F.STURM: Memoire sur les equations differentielles lineaires de second ordre, *J.math.Pures Appl.* 1 (1836), 106-186.

Note on the references

Since each chapter is a reproduction of a distinct paper, references to the equations as well as to the literature listed in the bibliography apply to each chapter separately.

Chapter 1

On the existence of conjugate points

This chapter provides two contributions to the Sturmian theory for nonselfadjoint systems. In §1 we prove some technical properties of solutions to linear ordinary differential equations in space of $n \times n$ matrices and we apply them to get a partial answer to the open problem raised by Ahmad [1],[2]. In §2 we use some properties of disconjugacy to obtain solutions, whose components have a prescribed sign, to nonhomogeneous systems satisfying suitable sign assumptions.

We recall a well-known concept that is used in this chapter. A second-order system

$$x'' + A(t)x = 0$$

is said to be **disconjugate on an interval** I if every non-trivial solution vanishes at most once on I. For example, $x'' = 0$ is disconjugate on any interval.

§1: A Theorem about the set of conjugate points

In this section we study some general properties of conjugate points in the nonselfadjoint case. To every continuous $n \times n$ matrix function A on R and to every $a \in R$ we associate the following Cauchy's matrix problems:

(1) $$X'' + A(t)X = 0_n, \quad X(a) = 0_n, \quad X'(a) = I_n,$$

(2) $$Y'' + A(t)Y = 0_n, \quad Y(a) = I_n, \quad Y'(a) = 0_n$$

and we shall denote by $X_A(t)$ and $Y_A(t)$ their unique solutions. It is well known (cf. Reid [9],[10]) that b is a conjugate point of a with respect to the system $x'' + A(t)x = 0$ if and only if $det\ X_A(b) = 0$. If this is the case then the multiplicity of b is, by definition, equal to dim $Ker\ X_A(b)$.

Theorem 1 *For the matrices just introduced and for every $t \in R$ the following equations hold:*

$$\begin{aligned}
&(a) \quad X'_{A^T}(t)^T X_A(t) - X_{A^T}(t)^T X'_A(t) = 0_n \\
&(b) \quad Y_{A^T}(t)^T X'_A(t) - Y'_{A^T}(t)^T X_A(t) = I_n \\
&(c) \quad X_A(t) Y_{A^T}(t)^T - Y_A(t) X_{A^T}(t)^T = 0_n \\
&(d) \quad X'_A(t) Y_{A^T}(t)^T - Y'_A(t) X_{A^T}(t)^T = I_n
\end{aligned}$$

Proof. Consider the following Cauchy's matrix problem:

$$Z' = \begin{pmatrix} 0_n & I_n \\ -A(t) & 0_n \end{pmatrix} Z, \quad Z(a) = I_{2n}$$

and let $Z(t)$ be the solution. An easy calculation shows that

$$Z(t) = \begin{pmatrix} Y_A(t) & X_A(t) \\ Y'_A(t) & X'_A(t) \end{pmatrix}.$$

Consider now the following Cauchy's matrix problem:

$$W' = -W \begin{pmatrix} 0_n & I_n \\ -A(t) & 0_n \end{pmatrix}, \quad W(a) = I_{2n}$$

and let $W(t)$ be its unique solution. An easy calculation shows that

$$W(t) = \begin{pmatrix} X'_{A^T}(t)^T & -X_{A^T}(t)^T \\ -Y'_{A^T}(t)^T & Y_{A^T}(t)^T \end{pmatrix}.$$

It is easy to prove that the derivative of $W(t)Z(t)$ is zero everywhere. In fact,

$$(W(t)Z(t))' = W'(t)Z(t) + W(t)Z'(t) = -W(t)\begin{pmatrix} 0_n & I_n \\ -A(t) & 0_n \end{pmatrix} Z(t)+$$

$$W(t)\begin{pmatrix} 0_n & I_n \\ -A(t) & 0_n \end{pmatrix} Z(t) = 0_{2n}.$$

From this it follows that $W(t)Z(t) \equiv W(a)Z(a) = I_{2n}$ and so $Z(t)W(t) \equiv I_{2n}$. Therefore

$$\begin{pmatrix} X'_{A^T}(t)^T & -X_{A^T}(t)^T \\ -Y'_{A^T}(t)^T & Y_{A^T}(t)^T \end{pmatrix} \begin{pmatrix} Y_A(t) & X_A(t) \\ Y'_A(t) & X'_A(t) \end{pmatrix} = I_{2n}$$

leads us to the first two equations and

$$\begin{pmatrix} Y_A(t) & X_A(t) \\ Y'_A(t) & X'_A(t) \end{pmatrix} \begin{pmatrix} X'_{A^T}(t)^T & -X_{A^T}(t)^T \\ -Y'_{A^T}(t)^T & Y_{A^T}(t)^T \end{pmatrix} = I_{2n}$$

leads us to the last two equations. Q.E.D.

Theorem 2 *The following relations hold for every $t \in R$:*

(a) $X'_A(t)\ Ker\ X_A(t) = Ker\ X_{A^T}(t)^T$;

(b) *the restriction Q_t of the linear operator $X'_A(t)$ on $Ker\ X_A(t)$ is injective;*

(c) $dim\ Ker\ X_A(t) = dim\ Ker\ X_{A^T}(t)$.

Proof. (a) Let $h \in X'_A(t)\ Ker\ X_A(t)$. Then there exists $\eta \in Ker\ X_A(t)$ such that $h = X'_A(t)\eta$. If we multiply equation (a) of Theorem 1 from the right by η then we obtain that $X_{A^T}(t)^T h = 0$ because $X_A(t)\eta = 0$ by hypothesis. From this it follows that $h \in Ker\ X_{A^T}(t)^T$. Let $k \in Ker\ X_{A^T}(t)^T$. If we multiply equation (d) in Theorem 1 from the right by k then we obtain that $k = X'_A(t)Y_{A^T}(t)^T k$ because $X_{A^T}(t)^T k = 0$ by hypothesis. We must check that $Y_{A^T}(t)^T k \in Ker\ X_A(t)$. By hypothesis $k \in Ker\ X_{A^T}(t)^T$ and so, if we multiply equation (c) in Theorem 1 from the right by k we obtain that $X_A(t)Y_{A^T}(t)^T k = 0$. The relation (a) is therefore proved.

(b) Let $\eta_1, \eta_2 \in Ker\ X_A(t)$ and suppose $Q_t(\eta_1) = Q_t(\eta_2)$. Then $X'_A(t)\eta_1 = X'_A(t)\eta_2$. From this it follows that

$$\eta_1 - \eta_2 \in Ker\ X'_A(t)$$

and, obviously, $\eta_1 - \eta_2 \in Ker\ X_A(t)$. Now, if we multiply equation (b) in Theorem 1 from the right by $\eta_1 - \eta_2$ then we obtain that $0 = \eta_1 - \eta_2$ and so $\eta_1 = \eta_2$.

(c) It follows from (b) that

$$\dim X'_A(t)\ Ker\ X_A(t) = \dim Ker\ X_A(t)$$

and so, from (a) it follows that

$$\dim\, Ker\ X_A(t) = \dim\, Ker\ X_{A^T}(t)^T = \dim\, Ker\ X_{A^T}(t).$$

Q.E.D.

Remark Condition (c) in Theorem 2 implies a simple, direct proof of Proposition 1 in Ahmad-Lazer [4].

Now we apply the above technical results to prove a theorem that answers partially a question raised by Ahmad [1], [2].

Theorem 3 *Let A be a continuous $n \times n$ matrix function on R and consider the system:*

$$x'' + A(t)x = 0. \tag{3}$$

Let a be a real number and define C_a as the set of conjugate points of a with respect to (3) *that have multiplicity strictly greater than $n/2$. Then C_a is discrete.*

Proof. Suppose, by contradiction, that b is an accumulation point of C_a: there exists a sequence (s_i) in C_a such that $\lim_{i\to\infty} s_i = b$. Without loss of generality we may assume that $\dim\, Ker\ X_A(s_i) := m > n/2$ for every $i = 1, 2, 3, \ldots$. Define $K_i = Ker\ X_A(s_i)$ and $L_i = Ker\ X_{A^T}(s_i)$. From the Theorem 2 (a) it follows that

$$X'_A(s_i)K_i = Ker\ X_{A^T}(s_i)^T.$$

If we replace A with A^T in the previous formula then we obtain that

$$X'_{A^T}(s_i)L_i = Ker\ X_A(s_i)^T.$$

From (c) of Theorem 2 or, alternatively, from Proposition 1 of Ahmad-Lazer [4] we have $\dim L_i = \dim K_i = m$ and so

$$\dim X'_{A^T}(s_i)L_i = \dim X'_A(s_i)K_i = m.$$

By hypothesis $m > n/2$. Therefore

$$(X'_{A^T}(s_i)L_i) \cap (X'_A(s_i)K_i) \neq \{0\}$$

for $i = 1, 2, 3, \ldots$. For every i we can therefore select from this intersection an element ρ_i with norm equal to 1. Now, (ρ_i) is a sequence in the unit sphere of R^n, we can therefore find a convergent subsequence (ρ_{i_r}). Put $\rho = \lim_{r\to\infty} \rho_{i_r}$. Obviously $\|\rho\| = 1$. From the definition of ρ_i there exists $\eta_i \in K_i$ and $\eta_i^* \in L_i$ such that $\rho_i = X'_A(s_i)\eta_i$ and $\rho_i = X'_{A^T}(s_i)\eta_i^*$. We claim that the sequences (η_{i_r}) and $(\eta_{i_r}^*)$ are convergent. To see, for example, that the sequence (η_{i_r}) is convergent we proceed as follows. We let $t = s_{i_r}$ in the equation (b) of the Theorem 1 and multiply it from the right by η_{i_r}. We obtain that $\eta_{i_r} = Y_{A^T}(s_{i_r})^T \rho_{i_r}$ and this tends to $Y_{A^T}(b)^T\rho$ for $r \to \infty$. It follows that the sequence (η_{i_r}) is convergent. In the same way we prove that also the sequence $(\eta_{i_r}^*)$ is convergent. Set $\eta = \lim_{r\to\infty} \eta_{i_r}$, $\eta^* = \lim_{r\to\infty} \eta_{i_r}^*$. From

$$X_A(s_{i_r})\eta_{i_r} = 0 \quad \forall r \in N$$

and

$$X_{A^T}(s_{i_r})\eta_{i_r}^* = 0 \quad \forall r \in N$$

it follows that

$$X_A(b)\eta = 0 \text{ and } X_{A^T}(b)\eta^* = 0.$$

So $\eta \in Ker\ X_A(b)$ and $\eta^* \in Ker\ X_{A^T}(b)$. From

$$X'_A(s_{i_r})\eta_{i_r} = \rho_{i_r} \quad \forall r \in N$$

and

$$X'_{A^T}(s_{i_r})\eta_{i_r}^* = \rho_{i_r} \quad \forall r \in N$$

it follows that

(4) $$\rho = X'_A(b)\eta$$

and

(5) $$\rho = X'_{A^T}(b)\eta^*.$$

Since $\eta^* \in Ker\ X_{A^T}(b)$, from (5) and

$$X'_{A^T}(b)\ Ker\ X_{A^T}(b) = Ker\ X_A(b)^T$$

it follows that $\rho \in Ker\ X_A(b)^T$ and so $\rho \in (Range\ X_A(b))^\perp$.
Now, $\langle \rho, X_A(s_{i_r})\eta_{i_r}\rangle = 0$ and $\langle \rho, X_A(b)\eta_{i_r}\rangle = 0$ because $\rho \in (Range\ X_A(b))^\perp$. Set

$$c := \min\{s_{i_r}, b\}, \qquad d := \max\{s_{i_r}, b\}$$

and for every $t \in [c,d]$ define

$$F_r(t) := \langle \rho, X_A(t)\eta_{i_r}\rangle.$$

Then $F_r(t)$ is a real valued, differentiable function, defined on the interval $[c,d]$ and such that $F_r(c) = F_r(d) = 0$. From Rolle's Theorem there exists $c \le t_r \le d$ such that $F'_r(t_r) = 0$ i.e. $\langle \rho, X'_A(t_r)\eta_{i_r}\rangle = 0$. Obviously, $\lim_{i\to\infty} t_i = b$ and so, passing to the limit in the last equation, we obtain that $\langle \rho, X'_A(b)\eta\rangle = 0$. From (4) we have that $\langle \rho,\rho\rangle = 0$ and so $\rho = 0$. But this contradicts the fact that $\|\rho\| = 1$. Q.E.D.

§2: Existence of solutions with prescribed sign

By appealing on the properties of disconjugacy, we study in this section the existence of solutions to nonhomogenous equations with components having a prescribed sign.

Theorem 4 *Let $A(t) = \{a_{i,j}(t)\}$ be a continuous $n \times n$ matrix on R and let $f : R \to R^n$ be continuous. Consider the system:*

$$x'' + A(t)x = f(t). \tag{6}$$

Let a, b be two real numbers with $a < b$. Suppose that the system

$$x'' + A(t)x = 0 \tag{3}$$

is disconjugate on $[a, b]$; suppose further that $a_{i,j}(t) \geq 0$ for $t \in [a, b]$, $1 \leq i, j \leq n$ and $f_i(t) \leq 0$ for $t \in [a, b]$, $1 \leq i \leq n$. Then there exists a unique solution $z(t)$ of (6) such that $z(a) = z(b) = 0$ and $z_i(t) \geq 0$ for $t \in [a, b]$, $1 \leq i \leq n$.

Proof. Let C^0 be the Banach space of continuous functions $\phi : [a, b] \to R^n$ and let $P = \{\phi \in C^0 | \phi_i \geq 0 \text{ on } [a, b] \text{ for all } i\}$. It is well-known that P is a total (i.e. generating according to Krasnoselski [6]) cone in C^0 that makes C^0 an ordered Banach space. Define the following operator: for $\phi \in C^0$ put

$$L(\phi)(t) = \int_a^b G(t, s)A(s)\phi(s)ds.$$

where G is the Green function for the two-point boundary value problem. We know that L is a linear, continuous and compact operator. From the hypoteses it follows that L is a positive operator. Define

$$F(t) = -\int_a^b G(t, s)f(s)ds.$$

From the assumptions it follows that $F_i(t) \geq 0$ for $t \in [a, b]$ and $i = 1, 2, \ldots, n$, hence $F \in P$. Consider now the equation

$$\phi = L(\phi) + F. \tag{7}$$

Notice that ϕ solves (7) if and only if ϕ solves (6) on $[a, b]$. Let r be the spectral radius of L. Recall that r is larger than the absolute value of every eigenvalue of L. If $r = 0$, from Theorem 2.16 of Krasnoselski [6], it follows that there exists a unique solution $\phi \in$ P of (7). In this case the Theorem 4 is proved.

Suppose now that $r > 0$. By the famous Krein-Rutman theorem there exists $\phi \in \mathrm{P} - \{0\}$ such that $r\phi = L(\phi)$. We claim that $r < 1$. Suppose, by contradiction, $r \geq 1$ and set $b_{i,j} = r^{-1}a_{i,j}$, $B(t) = \{b_{i,j}(t)\}$. From

$$r\phi = L(\phi) = \int_a^b G(t,s)A(s)\phi(s)ds$$

it follows that

$$\phi = \int_a^b G(t,s)B(s)\phi(s)ds.$$

This shows that ϕ is a non trivial solution of the following problem:

$$y'' + B(t)y = 0, \quad y(a) = y(b) = 0.$$

From $r \geq 1$ it follows that $b_{i,j}(t) \leq a_{i,j}(t)$ for $t \in [a,b]$ and $1 \leq i,j \leq n$. By Theorem 1 of Ahmad-Lazer [3] there exists a conjugate point c of a with respect to (3) such that $a < c \leq b$, but this contradicts the disconjugacy of (3) on $[a,b]$. Therefore $r < 1$ and from the Theorem 2.16 of Krasnoselski [6] there exists a unique solution $\phi \in \mathrm{P}$ of (6). Q.E.D.

Theorem 5 *Let $A(t) = \{a_{i,j}(t)\}$ be a continuous $n \times n$ matrix on R and let $f : R \to R^n$ be continuous. Consider the system:*

$$x'' + A(t)x = f(t). \tag{6}$$

Let a, b be real numbers, $a < b$. Suppose that the system

$$x'' + A(t)x = 0 \tag{3}$$

is disconjugate on $[a,b]$. Suppose further that there exist two sets I and J such that $I \cup J = \{1,\dots,n\}$, $I \cap J = \emptyset$ and for every $t \in [a,b]$ $a_{i,j}(t) \geq 0$ if $(i,j) \in I \times I \cup J \times J$, $a_{i,j}(t) \leq 0$ if $(i,j) \in I \times J \cup J \times I$, $f_i(t) \leq 0$ if $i \in I$ and $f_i(t) \geq 0$ if $i \in J$. Then there exists a unique solution $z(t)$ of (6) such that $z(a) = z(b) = 0$, for every $t \in [a,b]$ $z_i(t) \leq 0$ if $i \in J$ and $z_i(t) \geq 0$ if $i \in I$.

Proof. Put, by definition, for $i = 1, 2, \dots, n$

$$c_i = \begin{cases} 1 & if \quad i \in I \\ -1 & if \quad i \in J, \end{cases}$$

$$C = diag(c_1, \dots, c_n).$$

Then $C^T = C = C^{-1}$. Let $B(t) = CA(t)C$. If $B(t) = \{b_{i,j}(t)\}$ we have

$$b_{i,j}(t) = c_i c_j a_{i,j}(t) = \begin{cases} a_{i,j}(t) & if \quad (i,j) \in I \times I \cup J \times J \\ -a_{i,j}(t) & if \quad (i,j) \in I \times J \cup J \times I \end{cases} = |a_{i,j}(t)|$$

and then all the entries of $B(t)$ are non negative. An easy calculation shows that

$$X_B(t) = X_{CAC}(t) = CX_A(t)C.$$

So $det\ X_B(t) = det\ X_A(t) \neq 0$ for $a < t \leq b$. From this it follows that the system

$$z'' + B(t)z = 0$$

is disconjugate on $[a, b]$.

Put now $f^*(t) = Cf(t)$. We have

$$f_i^*(t) = c_i f_i(t) = \begin{cases} -f_i(t) & if \quad i \in J \\ f_i(t) & if \quad i \in I \end{cases} = -|f_i(t)|$$

for every $t \in [a, b]$. Then the system

$$z'' + B(t)z = f^*(t) \tag{8}$$

satisfies the condition of Theorem 4 and so there exists a solution $z(t)$ of (8) such that $z(a) = z(b) = 0$, $z_i(t) \geq 0$ for $t \in [a, b]$ and $i = 1, 2, \ldots, n$. If we multiply the equation (8) from the left by C then we obtain:

$$(Cz)'' + A(t)(Cz) = f(t).$$

So $g = Cz$ satisfies (6), $g(a) = g(b) = 0$ and an easy calculation shows that for every $t \in [a, b]$ $g_i(t) \leq 0$ if $i \in J$ and $g_i(t) \geq 0$ if $i \in I$. Q.E.D.

Mutatis mutandi in the proofs of Theorems 4 and 5 respectively, we can prove the following results:

Theorem 6 *Let $A(t) = \{a_{i,j}(t)\}$ be a continuous $n \times n$ matrix on R and let $f : R \to R^n$ be continuous. Consider the system:*

$$x'' + A(t)x = f(t). \tag{6}$$

Let a, b be two real numbers with $a < b$. Suppose that the system

$$x'' + A(t)x = 0 \tag{3}$$

is disfocal on $[a, b]$; suppose further that $a_{i,j}(t) \geq 0$ for $t \in [a, b]$, $1 \leq i, j \leq n$ and $f_i(t) \leq 0$ for $t \in [a, b]$, $1 \leq i \leq n$. Then there exists a unique solution $z(t)$ of (6) such that $z'(a) = z(b) = 0$ and $z_i(t) \geq 0$ for $t \in [a, b]$ and $1 \leq i \leq n$.

Theorem 7 *Let $A(t) = \{a_{i,j}(t)\}$ be a continuous $n \times n$ matrix on R and let $f : R \to R^n$ be continuous. Consider the system*

$$x'' + A(t)x = f. \tag{6}$$

Let a, b be real number, $a < b$. Suppose that the system

$$x'' + A(t)x = 0 \tag{3}$$

is disfocal on $[a,b]$. Suppose further that there exist two sets I and J such that $I \cup J = \{1, \cdots, n\}$, $I \cap J = \emptyset$ and for every $t \in [a,b]$ $a_{i,j}(t) \geq 0$ if $(i,j) \in I \times I \cup J \times J$, $a_{i,j}(t) \leq 0$ if $(i,j) \in I \times J \cup J \times I$, $f_i(t) \leq 0$ if $i \in I$ and $f_i(t) \geq 0$ if $i \in J$. Then there exists a unique solution $z(t)$ of (6) such that $z'(a) = z(b) = 0$, for every $t \in [a,b]$ $z_i(t) \leq 0$ if $i \in J$ and $z_i(t) \geq 0$ if $i \in I$.

References.

1. S. AHMAD: On positivity of solutions and conjugate points of nonselfadjoint systems, Bull. Acad. Polon. Sci. Math. Astronom. Phys. 27 (1979), 71-75.

2. ______: On Sturmian theory for second order systems, Proc. Amer. Math. Soc. 87 (1983), 661-665.

3. ______ and A.C.LAZER: An N-dimensional extension of the Sturm separation and comparison theory to a class of nonselfadjoint systems, SIAM J. Math. Anal.9 (1978), 1137-1150.

4. ______and ______: On an extension of Sturm's comparison theorems to a class of nonselfadjoint second-order systems, Nonlinear Anal.TMA 4 (1980), 497-501.

5. W. COPPEL: Disconjugacy, Springer, New York, 1971.

6. M.A.KRASNOSELSKI: Positive solutions of operator equations, Nordhoff, Groningen, 1964.

7. M.MORSE: A generalization of the Sturm separation and comparison theorems in n-space, Math.Ann.103 (1930), 52-69.

8. ______: Variational analysis: critical extremals and Sturmian extension, John Wiley, New York, 1973.

9. W.T.Reid: Ordinary differential equations, John Wiley, New York, 1971.

10. ______: Sturmian theory for ordinary differential equations, Springer, New York, 1980.

11. J.C.F. STURM: Mémoire sur les équations différentielles linéaires de second ordre, J.Math.Pures Appl. 1 (1836), 106-186.

Chapter 2

Characterization of the existence of symmetric solutions

§1. Introduction

In recent years there has been considerable interest in the existence of symmetric solutions to second-order matrix equations, partly due to their link to oscillation theory (e.g., Goff-St.Mary [1], Reid [2], Shreve [3] and literature cited there in).

The aim of this chapter is to characterize the existence of symmetric solutions to the second- order linear matrix equation

$$X'' + AX = 0$$

when A is a given constant matrix. The characterization is obtained in the case of Cauchy problems as well as in the case of two-point boundary value problems. It seems a remarkable fact that the matrix A is not assumed to be symmetric in the Cauchy problem.

The proofs appeal to a direct argument based on an explicit description of two fundamental solutions as well as to ordered Banach space techniques.

§2. The Cauchy problem

In this section we are interested in the existence of symmetric solutions to Cauchy matrix problems

$$X'' + AX = 0, \tag{1}$$

$$X(\tau) = R, \quad X'(\tau) = S \tag{2}$$

where A, S and R are given $n \times n$ matrices.

A solution X to (1) is called *symmetric* if $X(t)$ is a symmetric matrix, i.e. $X(t) = X(t)^T$, for all $t \in \mathbf{R}$.

We shall denote by

$$X_A(t,\tau) \quad \text{and} \quad Y_A(t,\tau)$$

the unique solution to (1) satisfying the initial conditions

$$X_A(\tau,\tau) = 0, \quad X'(\tau,\tau) = I \quad \text{and}$$

$$Y_A(\tau,\tau) = I, \quad Y'_A(\tau,\tau) = 0$$

respectively, I being the identity matrix.

We start with a property of these solutions:

Lemma 1. *For every a, b, $t \in \mathbf{R}$ we have that*

$$X'_{A^T}(t,b)^T X_A(t,a) - X_{A^T}(t,b)^T X'_A(t,a) = X_A(b,a), \tag{a}$$

$$Y_A(t,b)^T X'_{A^T}(t,a) - Y'_A(t,b)^T X_{A^T}(t,a) = X'_{A^T}(b,a). \tag{b}$$

In particular, we have $Y_A(a,b) = X'_{A^T}(b,a)^T$ *and* $X_{A^T}(a,b) = -X_A(b,a)^T$.

Proof. Call $f(t)$ the left hand side of (a). Then for every $t \in \mathbf{R}$ we have that

$$\begin{gathered} f'(t) = X''_{A^T}(t,b)^T X_A(t,a) + X'_{A^T}(t,b)^T X'_A(t,a) - \\ -X'_{A^T}(t,b)^T X'_A(t,a) - X_{A^T}(t,b)^T X''_A(t,a) = \\ -X_{A^T}(t,b)^T A X_A(t,a) + X_{A^T}(t,b)^T A X_A(t,a) = 0. \end{gathered}$$

Therefore $f(t) \equiv f(b) = X_A(b,a)$. This proves (a). The proof of (b) is analogous. The last assertion follows by taking $t = a$. **QED**

Now we analize two special cases of (1), whose solutions play a fundamental role in the proof of the wanted characterization.

Lemma 2. *Let A be an arbitrary $n \times n$ matrix. The solution of the Cauchy matrix problem*

$$\begin{cases} X'' + AX = 0 \\ X(\tau) = 0 \\ X'(\tau) = S \end{cases}$$

is a symmetric solution if and only if

(i) S *is symmetric, and*

(ii) $SA^T = AS$.

Proof. We have

$$X_A(t,\tau)^T = X_A(t-\tau,0)^T = -X_A(\tau-t,0)^T = -X_A(\tau,t)^T.$$

From Lemma 1 it follows that $-X_A(\tau,t)^T = X_{A^T}(t,\tau)$. Therefore,

$$X_A(t,\tau)^T = X_{A^T}(t,\tau). \tag{3}$$

Moreover, $X_A(t,\tau)S$ is the unique solution of the given Cauchy problem.

Assume first that conditions (i) and (ii) hold. If we multiply the equation $X''_{A^T}(t,\tau) + A^T X_{A^T}(t,\tau) = 0$ from the left by S, then we obtain

$$(SX_{A^T}(t,\tau))'' + (SA^T)X_{A^T}(t,\tau) = 0.$$

By virtue of (ii), $SA^T = AS$ and hence

$$(SX_{A^T}(t,\tau))'' + A(SX_{A^T}(t,\tau)) = 0.$$

Moreover, $SX_{A^T}(\tau,\tau) = 0$ and $SX'_{A^T}(\tau,\tau) = S$. This means that $SX_{A^T}(t,\tau)$ is a solution of the given Cauchy problem. By uniqueness we have

$$SX_{A^T}(t,\tau) = X_A(t,\tau)S.$$

From this, (i) and (3) we obtain that

$$(X_A(t,\tau)S)^T = S^T X_A(t,\tau)^T = SX_{A^T}(t,\tau) = X_A(t,\tau)S.$$

This means that the solution $X_A(t,\tau)S$ is symmetric.

Conversely, suppose that the solution $X_A(t,\tau)S$ is symmetric. Then $X'_A(t,\tau)S$ is also a symmetric matrix for every t. From $S = X'_A(\tau,\tau)S$ it follows that $S = S^T$. This proves (i). From

$$SX_A(t,\tau)^T = (X_A(t,\tau)S)^T = X_A(t,\tau)S \quad \text{for every } t$$

it follows that

$$SX'''_A(t,\tau)^T = X'''_A(t,\tau)S.$$

Thus by differentiating the given equation (1) we get

$$-SX'_A(t,\tau)^T A^T = -AX'_A(t,\tau)S.$$

Taking $t = \tau$ we obtain $SA^T = AS$. This proves (ii). **QED**

Lemma 3. *Let A be an arbitrary $n \times n$ matrix. The solution of the Cauchy matrix problem*

$$\begin{cases} X'' + AY = 0 \\ X(\tau) = R \\ X'(\tau) = 0 \end{cases}$$

is a symmetric solution if and only if
(i) R *is symmetric, and*
(ii) $RA^T = AR$.

Proof. From Lemma 1 we derive that $Y_A(t,\tau) = X'_{A^T}(\tau,t)^T$. For each t we have that

$$X'_{A^T}(t,\tau)^T = X'_{A^T}(t-\tau,0)^T = X'_{A^T}(\tau,t)^T,$$

hence

$$Y_A(t,\tau)R = X'_{A^T}(t,\tau)^T R.$$

Moreover, $Y_A(t,\tau)R$ is the unique solution of the given Cauchy problem.

Therefore $Y_A(t,\tau)R$ is symmetric for each t if and only if $X'_{A^T}(t,\tau)^T R$ is symmetric for each t. From

$$X_{A^T}(t,\tau)^T R = \left(\int_\tau^t X'_{A^T}(\sigma,\tau)^T d\sigma\right) R$$

it follows that $X'_{A^T}(t,\tau)^T R$ is symmetric for each t if and only if $X_{A^T}(t,\tau)^T R$ is symmetric for each t. But $X_{A^T}(t,\tau)^T R = X_A(t,\tau)R$ by (3). According to Lemma 2 this can happen if and only if R is symmetric and $RA^T = AR$. **QED**

Now we are ready to prove the main result of this section:

Theorem 1. *Let A be an arbitrary $n \times n$ matrix. The solution of the Cauchy matrix problem (1), (2) is a symmetric solution if and only if*
(i) R *and* S *are symmetric matrices, and*
(ii) $RA^T = AR$ *and* $SA^T = AS$.

Proof. The solution of (1), (2) is given by the formula

$$X(t) = X_A(t,\tau)S + Y_A(t,\tau)R.$$

So, if conditions (i) and (ii) hold, then Lemma 2 and Lemma 3 imply that $X(t)$ is symmetric.

Suppose now that $X(t)$ is symmetric. Then the matrix $R = X(\tau)$ is symmetric. From $X(t) = X(t)^T$ we get $X''(t) = (X(t)'')^T$. Hence $-AX(t) = -X(t)^T A^T$ for each t. If we put $t = \tau$ then we obtain $AR = R^T A^T = RA^T$. By Lemma 3, $Y_A(t,\tau)R$ is a symmetric solution of (1). From $X_A(t,\tau)S = X(t) - Y_A(t,\tau)R$ we obtain that $X_A(t,\tau)S$ is a symmetric solution of (1). Lemma 2 now implies that S is a symmetric matrix and that $AS = SA^T$. **QED**

Now we derive two interesting consequences:

Corollary 1. *X is a symmetric solution to (1) if and only if X is locally symmetric, i.e. $X(t)$ is symmetric for all t in an open interval.*

Proof. We need only to show that if $X(t)$ is symmetric for $a < t < b$, then $X(t)$ is symmetric for every t. Fix $a < \tau < b$ and set $R = X(\tau)$, $S = X'(\tau)$. The matrices R and S are symmetric because $X(t)$ is. From $X(t) = X(t)^T$ we get $X''(t) = (X(t)^T)''$. Hence

$$-AX(t) = -X(t)^T A^T \qquad (a < t < b).$$

Differentiating both sides, we obtain that

$$AX'(t) = X'(t)^T A^T \qquad (a < t < b).$$

Choosing $t = \tau$ in these two relations, we get

$$AR = RA^T \quad \text{and} \quad AS = SA^T$$

respectively (R and S being symmetric). At this point Theorem 1 can be applied to get the desired conclusion. **QED**

Corollary 2. *Let A be an arbitrary $n \times n$ matrix. The matrix problem (1) has at least $2n$ linearly indipendent symmetric solutions.*

Proof. Let V be the vector space of all symmetric $n \times n$ matrices and let W be the vector space of all skew-symmetric $n \times n$ matrices. Let L be the linear map from $V \times V$ to $W \times W$ defined in the following way: $L(x, y) = (Ax - xA^T, Ay - yA^T)$. Let $X(t)$ be a solution of (1). By Theorem 1, $X(t)$ is a symmetric solution of (1) if and only if $L(X(0), X'(0)) = 0$. We know from Linear Algebra that

$$\dim\,(V \times V) = \dim\,\text{Ker}\,(L) + \dim\,L(V \times V),$$

$$\dim\,V = n(n+1)/2, \quad \dim\,W = n(n-1)/2.$$

Combining these relations, we obtain the conclusion of the Corollary since $L(V \times V) \subseteq W \times W$.
QED

Now we assume that A is symmetric and we deduce from Theorem 1 a result of interest in Sturmian theory (cf. Reid [2]):

Theorem 2. *Let A be a symmetric matrix and $a < b$ be real numbers such that the interval $]a, b]$ contains no focal points of a with respect to the system*

$$x'' + Ax = 0. \tag{4}$$

Then there exists a symmetric solution $X(t)$ of (1) such that:
(a) $X(a) \neq 0$ *and* $X'(a) = 0$;
(b) $X(a)$ *is positive semidefinite;*
(c) $X(b) = \lambda X(a)$ *for a suitable* $\lambda \geq 0$.

Proof. Let V be the vector space of all symmetric matrices and let P be the cone of all positive semidefinite matrices. We denote by $\leq$ the order induced on V by P. Let V_A be the vector subspace of V consisting of all symmetric matrices S commuting with A: $AS = SA$. Let $P_A = V_A \cap P$.

By Lemma 3, $Y_A(t,a)$ is symmetric for every t. We prove that

$$Y_A(b,a) \geq 0. \tag{5}$$

Suppose, by contradiction, that $Y_A(b,a) \notin P$. From $Y_A(a,a) = I \in P^\circ$, from the fact that $Y_A(b,a)$ belongs to the open set $V \backslash P$, and from the continuity of the function $Y_A(t,a)$ on the connected set $]a,b[$, it follows that there exists ν, $a < \nu < b$, such that $Y_A(\nu,a)$ is an element of the boundary of P. Therefore $\det Y_A(\nu,a) = 0$. This implies that ν is a focal point of a with respect to (4), contradicing the assumption.

We prove next that if $S \in V_A$ then $Y_A(b,a)S \in V_A$. If $S \in V_A$ then $S^T = S$ and $AS = SA = SA^T$. By Lemma 3, $Y_A(t,a)S$ is a symmetric solution of (1). By uniqueness, this implies that the solution of the following matrix Cauchy problem

$$\begin{cases} X'' + AX = 0 \\ X(b) = Y_A(b,a)S \\ X'(b) = Y'_A(b,a)S \end{cases}$$

is symmetric. Then (ii) of Theorem 1 implies that $Y_A(b,a)S \in V_A$.

Define a linear operator L as follows: for $S \in V_A$ put $L(S) = Y_A(b,a)S$. We have just shown that L maps V_A in V_A. From (5) it follows that if $R \in V$ and $R \geq 0$ then $Y_A(b,a)R \geq 0$, which shows that L is a positive operator.

If $L \equiv 0$ on P_A, then $L \equiv 0$ on V_A since P_A has non-empty interior because $I \in P_A$ and I is an interior point of P. In this case the conclusion of the theorem follows trivially.

If $L \not\equiv 0$ on P_A, then there is $Z \in P_A$ such that $L(Z) > 0$. Define then a sequence of non-linear operators

$$L_n(X) = \frac{L(X + \frac{1}{n}Z)}{\|L(X + \frac{1}{n}Z)\|}.$$

This is a good definition since $L(X + \frac{1}{n}Z) \geq \frac{1}{n}L(Z) > 0$. L_n maps continuously the closed and convex set

$$K = \{X \in P_A \mid \|X\| \leq 1\}$$

into itself. By Brouwer fixed point theorem, for every n there is $X_n \in K$ such that $L(X_n) = X_n$. This means

$$\|L(X_n + \frac{1}{n}Z)\|X_n = L(X_n + \frac{1}{n}Z). \tag{6}$$

By the compactness of K, there is a convergent subsequence $X_{n_k} \to X_\infty$ with $\|X_\infty\| = 1$. Passing to the limit in (6) with $n = n_k$, we obtain

$$\lambda X_\infty = L(X_\infty)$$

for a suitable $\lambda \geq 0$. Define $X(t) := Y_A(t,a)X_\infty$. From $X_\infty \in V_A$ and from Lemma 3 we have that $X(t)$ is a symmetric solution such that $X(a) = Y_A(a,a)X_\infty = X_\infty \neq 0$, $X'(a) = 0$ and $X(b) = Y_A(b,a)X_\infty = L(X_\infty) = \lambda X_\infty = \lambda X(a)$. **QED**

§3. The two-point boundary value problem

In this section we are interested in knowing when the two-point matrix boundary value problem

$$X'' + AX = 0, \quad X(0) = 0 = X(1)$$

has a symmetric, non-trivial solution X: $X(t)$ must be symmetric for every $0 \leq t \leq 1$. By Corollary 1 of Theorem 1, this implies that $X(t)$ is symmetric for every real t.

It is a consequence of Theorem 3 below that this problem is solvable if and only if $\lambda = 1$ is an eigenvalue of the two-point boundary value for the associated system in $\mathbf{R}^n$:

$$y'' + Ay = 0, \quad y(0) = 0 = y(1)$$

and $y(t) \in \mathbf{R}^n$.

In order to achieve this result we shall compare the following eigenvalue problems

$$(7) \qquad \begin{cases} X'' + \lambda AX = 0, \quad X(0) = 0 = X(1), \\ X(t) = X(t)^T \quad 0 \leq t \leq 1; \end{cases}$$

$$(8) \qquad \begin{cases} y'' + \lambda Ay = 0, \quad y(0) = 0 = y(1), \\ y(t) \in \mathbf{R}^n \quad 0 \leq t \leq 1. \end{cases}$$

In this context, we need to assume the symmetry of A. The main result is the following:

Theorem 3. *Let A be a symmetric $n \times n$ matrix. The real number λ is an eigenvalue of (7) if and only if λ is an eigenvalue of (8).*

In order to prove Theorem 3 we first prove two lemmas that are consequences of the results in §2.

Lemma 4. *Let A be a symmetric $n \times n$ matrix. A positive real number $\lambda > 0$ is an eigenvalue of (8) if and only if $detX_A(\lambda^{1/2}, 0) = 0$.*

Proof. *Necessity*: A direct calculation shows that $X_A(\lambda^{1/2}t, 0)$ is the solution of

$$\begin{cases} X'' + \lambda AX = 0 \\ X(0) = 0 \\ X'(0) = \lambda^{1/2} I. \end{cases}$$

If λ is an eigenvalue for (8), then there exists a non-trivial solution $x(t)$ of (8). From uniqueness, $x(t) = \lambda^{-1/2} X_A(\lambda^{1/2}t, 0)x'(0)$. The condition $x(1) = 0$ then imply that $\lambda^{-1/2} X_A(\lambda^{1/2}, 0)x'(0) = 0$ and so $\det X_A(\lambda^{1/2}, 0) = 0$.

Sufficiency: If $\det X_A(\lambda^{1/2}, 0) = 0$ then there exists a non-null $x_0 \in \mathbf{R}^n$ such that

$$X_A(\lambda^{1/2}, 0)x_0 = 0.$$

From this it follows that $X_A(\lambda^{1/2}t, 0)x_0$ is a non-trivial solution of (8). QED

In the same way we can prove the following:

Lemma 5. *Let A be a $n \times n$ symmetric matrix. A negative real number $\lambda < 0$ is an eigenvalue of (8) if and only if $\det X_{-A}((-\lambda)^{1/2}, 0) = 0$.*

Proof of Theorem 3. *Necessity*: Let λ be an eigenvalue of (6) and X a corresponding symmetric eigenfunction. Since X is non-trivial, there exist $z \in \mathbf{R}^n$ and $0 < \tau < 1$, such that $X(\tau)z \neq 0$. Then $y(t) = X(t)z$ is a non-trivial solution to (7) for the same λ.

Sufficiency: We confine the proof to the case $\lambda > 0$, since the case $\lambda < 0$ is similar.

Suppose that $\lambda > 0$ is an eigenvalue for (8). From Lemma 4, $\det X_A(\lambda^{1/2}, 0) = 0$. Applying Lemma 2 with $S = I$ we see that the solution $X(t)$ of the following matrix Cauchy problem

$$\begin{cases} X'' + AX = 0 \\ X(0) = 0 \\ X'(0) = I \end{cases}$$

is symmetric. By definition, $X(t) = X_A(t, 0)$. Put $X_0 := X(\lambda^{1/2})$. Then $\det X_0 = \det X_A(\lambda^{1/2}, 0) = 0$. Obviously, $X(t)$ is the unique solution of the following problem:

$$\begin{cases} Y'' + AY = 0 \\ Y(\lambda^{1/2}) = X(\lambda^{1/2}) \\ Y'(\lambda^{1/2}) = X'(\lambda^{1/2}). \end{cases}$$

But $X(t)$ is also a symmetric solution of this problem. We can therefore apply Theorem 1 and obtain that $X_0^T = X_0$ and $AX_0 = X_0 A^T$. By induction it is easy to see that for every integer k, $AX_0^k = X_0^k A^T$ (remember that, by assumption, $A^T = A$). Obviously, $(X_0^k)^T = X_0^k$. If Q is a linear combination of the type

$$Q = \sum_{\nu=0}^{k} a_\nu X_0^\nu, \qquad a_1, \ldots, a_k \in \mathbf{R}$$

then it is easily seen that $Q^T = Q$ and $AQ = QA^T$.

Let $p(x)$ be the minimal polynomial of X_0. Then $p(x)$ has the form $p(x) = \sum_{\nu=0}^{k} a_\nu x^\nu$. From the fact that $\det X_0 = 0$ it follows that $a_0 = 0$. So $p(x) = xp_1(x)$, where $p_1(x) = \sum_{\nu=1}^{k} a_\nu x^{\nu-1}$. Put $Q_1 := p_1(X_0)$. Then, from the properties of the minimal polynomial it follows that $Q_1 \neq 0$ and $X_0 Q_1 = 0$. From the above argument we have $Q_1^T = Q_1$ and $AQ_1 = Q_1 A^T$. Put $Y(t) := X_A(\lambda^{1/2}t, 0)Q_1$. From Lemma 2 we have $Y(t)^T = Y(t)$ for every t. From $Y'(0) = \lambda^{1/2} Q_1 \neq 0$ it follows that $Y(t)$ is non-trivial. Finally, $Y(1) = X_0 Q_1 = 0$ and $Y(0) = 0$. Thus $Y(t)$ is a symmetric, non-trivial solution of (7). **QED**

References

[1]. S.GOFF and D.F.St.MARY: The Bohl transformation and oscillation of linear differential systems, *SIAM J.Math.Anal.* 20(1989), 215-221.

[2]. W.T.REID: *Sturmian theory for Ordinary Differential Equations*, Springer, New York 1980.

[3].W.E.SHREVE: Boundary value problems for second order nonlinear matrix differential equations, *Proc.Amer.Math.Soc.* 83(1981), 63-68.

Chapter 3

On the existence of conjugate and focal points

This chapter provides a contribution to the Sturmian theory for systems

$$x'' + A(t)x = 0 \tag{1}$$

where $A(t)$ is a 2×2 matrix depending continuously on t. Purposely, $A(t)$ is not assumed symmetric, so that neither the spectral theory of self-adjoint operators nor the methods of the Calculus of Variation are applicable.

Sturmian theory deals with oscillation, separation and comparison of solution. Its goal is to detect the distribution and the multiplicity of zeros of solutions. In this connections, the concepts of conjugate and of focal points play a fundamental role. The major contribution of this chapter is the reduction of the question of the existence of conjugate and of focal points of

$$x'' + A(t)x = 0 \tag{1}$$

to the analysis of a scalar equation- in a deep analogy with Liouville theorem about the Wronskian of matrix solution. Our argument works only for 2×2 matrices $A(t)$, hence the origin of our restriction in the dimension [the extension to $n \times n$ matrices $A(t)$ is an open problem]. From the scalar equation we derive a variety of information about conjugate and focal points of (1).

§2 Terminology, notations and preliminary results

In what follows, we shall denote by

$$X_A(t,\tau) \quad \text{and} \quad Y_A(t,\tau)$$

the unique matrix solution to (1) such that

$$X_A(\tau,\tau)=0, \quad X'_A(\tau,\tau)=I \quad \text{and} \quad Y_A(\tau,\tau)=I, \quad Y_A(\tau,\tau)=0$$

respectively.

Let a and b be real numbers. We say that b is a *conjugate point* of a if there exists a nontrivial solution $x(t)$ of (1) such that $x(a)=x(b)=0$. We say that b is a *focal point* of a if $b>a$ and there exists a nontrivial solution $y(t)$ of (1) such that $y'(a)=y(b)=0$. It is well-known that b is a conjugate point of a if and only if $\det X_A(b,a)=0$ and $b>a$ is a focal point of a if and only if $\det Y_A(b,a)=0$ (see [3]).

We prove now some elementary properties concerning the matrices X_A and Y_A.

Consider the following Cauchy matrix problem:

$$Z' = \begin{pmatrix} 0_n & I_n \\ -A(t) & 0_n \end{pmatrix} Z, \quad Z(a)=I_{2n} \tag{2}$$

and let $Z(t)$ be the solution. An easy calculation shows that

$$Z(t) = \begin{pmatrix} Y_A(t,a) & X_A(t,a) \\ Y'_A(t,a) & X'_A(t,a) \end{pmatrix}.$$

Consider now the following Cauchy matrix problem:

$$W' = -W \begin{pmatrix} 0_n & I_n \\ -A(t) & 0_n \end{pmatrix} \quad W(a)=I_{2n}$$

and let $W(t)$ be its unique solution. An easy calculation shows that

$$W(t) = \begin{pmatrix} X'_{A^T}(t,a)^T & -X_{A^T}(t,a)^T \\ -Y'_{A^T}(t,a)^T & Y_{A^T}(t,a)^T \end{pmatrix}.$$

We prove that the derivative of $W(t)Z(t)$ is zero everywhere. In fact,

$$(W(t)Z(t))' = W'(t)Z(t) + W(t)Z'(t) =$$

$$-W(t)\begin{pmatrix} 0_n & I_n \\ -A(t) & 0_n \end{pmatrix} Z(t) + W(t)\begin{pmatrix} 0_n & I_n \\ -A(t) & 0_n \end{pmatrix} Z(t) = 0_n.$$

From this it follows that $W(t)Z(t) \equiv W(a)Z(a) = I_n$ and so $Z(t)W(t) = I_{2n}$. If we rewrite the last equation in explicit form then we obtain:

$$\begin{pmatrix} X'_{A^T}(t,a)^T & -X_{A^T}(t,a)^T \\ -Y'_{A^T}(t,a)^T & Y_{A^T}(t,a)^T \end{pmatrix}\begin{pmatrix} Y_A(t,a) & X_A(t,a) \\ Y'_A(t,a) & X'_A(t,a) \end{pmatrix} = I_{2n}. \tag{3}$$

We require now the following result from Linear Algebra, (cf. [2], section 0.8.4.)

Theorem A Let

$$M = \begin{pmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \end{pmatrix}$$

be a non singular $2n \times 2n$ matrix and suppose that all the blocks $M_{i,j}$ are $n \times n$ matrices. Put

$$M^{-1} = \begin{pmatrix} N_{1,1} & N_{1,2} \\ N_{2,1} & N_{2,2} \end{pmatrix}$$

where all the blocks $N_{i,j}$ are $n \times n$ matrices. Then

$$|\det N_{2,2}| = |(\det M)^{-1} \det M_{1,1}|, \qquad |\det N_{2,1}| = |(\det M)^{-1} \det M_{2,1}|,$$

$$|\det N_{1,2}| = |(\det M)^{-1} \det M_{1,2}|, \qquad |\det N_{1,1}| = |(\det M)^{-1} \det M_{2,2}|.$$

If we apply Theorem A to the relation (3), then we obtain the following result:

Theorem 1. *For every $a,\ t \in \mathbf{R}$ we have that*

$$|\det Y_{A^T}(t,a)| = |\det Y_A(t,a)|,$$

$$|\det X_{A^T}(t,a)| = |\det X_A(t,a)|,$$

$$|\det X'_{A^T}(t,a)| = |\det X'_A(t,a)|.$$

Theorem 2. *For every $a, b, t \in \mathbf{R}$ we have that*

(a) $$X'_{A^T}(t,b)^T X_A(t,a) - X_{A^T}(t,b)^T X'_A(t,a) = X_A(b,a),$$

(b) $$Y_{A^T}(t,b)^T X'_A(t,a) - Y'_{A^T}(t,b)^T X_A(t,a) = X'_A(b,a).$$

In particular, we have $Y_{A^T}(a,b)^T = X'_A(b,a),\ \ X_{A^T}(a,b) = -X_A(b,a)^T$.

Proof. Call $f(t)$ the left hand side of (a). Then for every $t \in R$ we have that

$$f'(t) = X''_{A^T}(t,b)^T X_A(t,a) + X'_{A^T}(t,b)^T X'_A(t,a) -$$

$$-X'_{A^T}(t,b)^T X'_A(t,a) - X_{A^T}(t,b)^T X''_A(t,a) =$$

$$-X_{A^T}(t,b)^T A(t) X_A(t,a) + X_{A^T}(t,b)^T A(t) X_A(t,a) = 0.$$

Therefore $f(t) \equiv f(b) = X_A(b,a)$. This proves (a). The proof of (b) is analogous. The last assertion follows by taking $t = a$. **QED**

§3 Existence of focal points for 2×2 systems

In the scalar case ($A(t)$ is a real valued function), the existence of a $c \in]a, b[$ such that b is a focal point of c follows from Rolle's theorem. We prove the same result for two-dimensional systems (1) under suitable conditions on the matrix $A(t)$, cf. Theorems 4 and 5 below. In order to reach this goal we need to state some technical results. Of these, the following one plays, in our context, a similar role as Liouville theorem on the Wronskian since it helps to reduce the analysis of a system to the analysis of a scalar equation.

Theorem 3. *For every continuous 2×2 matrix $A(t)$ and every matrix solution $X(t)$ of (1), the following scalar equation*

$$(\det X(t))'' = -\mathrm{Tr}(A(t))\det X(t) + 2\det X'(t)$$

is satisfied.

Proof. Let

$$X(t) = \begin{pmatrix} x_{11}(t) & x_{12}(t) \\ x_{21}(t) & x_{22}(t) \end{pmatrix}, \qquad A(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{pmatrix}.$$

Then

$$\det X(t) = x_{11}(t)x_{22}(t) - x_{12}(t)x_{21}(t),$$

$$(\det X(t))' = x'_{11}(t)x_{22}(t) + x_{11}(t)x'_{22}(t) - x'_{12}(t)x_{21}(t) - x_{12}(t)x'_{21}(t),$$

$$(\det X(t))'' = x''_{11}(t)x_{22}(t) - x''_{12}(t)x_{21}(t) + x_{11}(t)x''_{22}(t) - x_{12}(t)x''_{21}(t)+$$

$$+x'_{11}(t)x'_{22}(t) - x'_{12}(t)x'_{21}(t) + x'_{11}(t)x'_{22}(t) - x'_{12}(t)x'_{21}(t).$$

From the relation

$$\begin{pmatrix} x''_{11}(t) & x''_{12}(t) \\ x''_{21}(t) & x''_{22}(t) \end{pmatrix} = -\begin{pmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{pmatrix} \begin{pmatrix} x_{11}(t) & x_{12}(t) \\ x_{21}(t) & x_{22}(t) \end{pmatrix}$$

it follows that

$$x''_{11}(t) = -[a_{11}(t)x_{11}(t) + a_{12}(t)x_{21}(t)],$$

$$x''_{12}(t) = -[a_{11}(t)x_{12}(t) + a_{12}(t)x_{22}(t)],$$

$$x''_{21}(t) = -[a_{21}(t)x_{11}(t) + a_{22}(t)x_{21}(t)],$$

$$x''_{22}(t) = -[a_{21}(t)x_{12}(t) + a_{22}(t)x_{22}(t)].$$

If we replace the corresponding quantity in the relation written before, then we obtain:

$$(\det X(t))'' = -a_{11}(t)x_{11}(t)x_{22}(t) - a_{12}(t)x_{21}(t)x_{22}(t) + a_{11}(t)x_{12}(t)x_{21}(t)+$$

$$+a_{12}(t)x_{22}(t)x_{21}(t) - a_{21}(t)x_{12}(t)x_{11}(t) - a_{22}(t)x_{22}(t)x_{11}(t)+$$

$$+a_{21}(t)x_{11}(t)x_{12}(t) + a_{22}(t)x_{21}(t)x_{12}(t) + 2[x'_{11}(t)x'_{22}(t) - x'_{12}(t)x'_{21}(t)] =$$

$$-a_{11}(t)[x_{11}(t)x_{22}(t) - x_{12}(t)x_{21}(t)] - a_{22}(t)[x_{11}(t)x_{22}(t) - x_{12}(t)x_{21}(t)]+$$

$$+2\det X'(t) = -\mathrm{Trace}(A(t))\det X(t) + 2\det X'(t).$$ **QED**

Corollary. *Let $A(t)$ be a continuous 2×2 matrix. Then a is a point of strict relative minimum for* $\det X_A(t,a)$ *as a function of t. In particular, for every a there exists a neighbourhood I of a such that $\det X_A(t,a) \neq 0$ for every $t \in I$ and $t \neq a$.*

Proof. Put

$$X_A(t,a) = \begin{pmatrix} x_{11}(t) & x_{12}(t) \\ x_{21}(t) & x_{22}(t) \end{pmatrix}.$$

By definition $X_A(a,a) = 0$, and so the first derivative of the function $\det X_A(t,a)$ evaluated at the point $t = a$ is equal to

$$x'(a)_{11}x_{22}(a) + x_{11}(a)x'_{22}(a) - x'_{12}(a)x_{21}(a) - x_{12}(a)x'_{21}(a) = 0$$

.

From Theorem 3 and from $X'_A(a,a) = I$ it follows that the second derivative of the function $\det X_A(t,a)$ evaluated at the point $t = a$ is equal to 2. Thus a is a point of minimum for $\det X_A(.,a)$. **QED**

Theorem 4. *Between any pair $a < b$ of conjugate points for (1) there is a focal point for (1) provided $A(t)$ is symmetric and positive semi-definite for every $a \leq t \leq b$.*

Proof. Assume, for contradiction, that there exists no number c, $a \leq c \leq b$ such that b is a focal point of c for (1). This implies that $\det Y_A(b,t) \neq 0$ for $a \leq t \leq b$. From Theorem 2 we have that $\det X'_A(t,b) \neq 0$ for $a \leq t \leq b$. Without loss of generality we may assume that a is the first conjugate point of b. We therefore assume that $\det X_A(t,b) \neq 0$ for $a < t < b$. Define $Z(t) = X'_A(t,b)X_A^{-1}(t,b)$ for $a < t < b$. We have

$$Z'(t) = X''_A(t,b)X_A^{-1}(t,b) - X'_A(t,b)X_A^{-1}(t,b)X'_A(t,b)X_A^{-1}(t,b) = -A(t) - Z^2(t).$$

This implies that $Z(t)$ is a solution of the following differential equation:

$$Z'(t) = -A(t) - Z^2(t). \tag{4}$$

Note that $Z(t)$ is a non-extendable solution of (4). In fact: suppose that $\lim_{t\to a^+} Z(t)$ exists and call this limit Z_a. Then we have that $Z_a X_A(a,b) = X'_A(a,b)$. By assumption $X'_A(a,b)$ is non-singular and so $X_A(a,b)$ must be non-singular. But this contradicts the fact that a is a conjugate point of b. In the same way we see that $Z(t)$ is non- extendable at the point b.

From the global existence theorem it follows that

$$\lim_{t\to a^+} \|Z(t)\| = \infty, \tag{5}$$

$$\lim_{t\to b^-} \|Z(t)\| = \infty. \tag{6}$$

It is well-known, [see 3], that

$$X_A^{-1}(t,b)^T X_A'(t,b)^T = X_A'(t,b)X_A^{-1}(t,b)$$

for $a < t < b$. This implies that $Z^T(t) = Z(t)$ for $a < t < b$.

In what follows, we denote by S the space of 2×2 symmetric matrices endowed with the order defined by the cone P of positive semi-definite matrices:

$$U \leq V \Leftrightarrow V - U \text{ positive semi} - \text{definite } \Leftrightarrow V - U \in P.$$

From the fact that $0 \leq A(t)$ and $0 \leq Z^2(t)$ for $a < t < b$ it follows that for $a < t_1 \leq t_2 < b$

$$Z(t_2) - Z(t_1) = \int_{t_1}^{t_2} Z'(\sigma)d(\sigma) = -\int_{t_1}^{t_2} (A(\sigma) + Z^2(\sigma))d(\sigma) \leq 0,$$

so that

$$Z(t_2) \leq Z(t_1). \tag{7}$$

From $\det X_A(b,b) = 0$ and from the Corollary of Theorem 3, we get the existence of a neighbourhood I of b such that $\det X_A(t,b) > 0$ for $t \in I$, $t \neq b$. From $\det X_A'(b,b) = 1$ and from the assumption that $\det X_A(t,b) \neq 0$ and $\det X_A'(t,b) \neq 0$ for $a < t < b$ it follows that $\det Z(t) > 0$ for $a < t < b$. This implies that the eigenvalues of $Z(t)$ have the same sign for $a < t < b$ and from this it follows that $Z(t) \leq 0$ for every $a < t < b$ or $Z(t) \geq 0$ for every $a < t < b$. Suppose first that $Z(t) \geq 0$ for every $a < t < b$. Then, for $a < t_1 \leq t_2 < b$ we deduce from (7) that

$$0 \leq Z(t_2) \leq Z(t_1). \tag{8}$$

On the ordered Banach space (S, P) we take the following norm:

$$\|U\| = \sup|u_{i,j}|$$

for every matrix $U = (u_{i,j})$. This norm turns out to be monotone, i.e.

$$U \leq V \Rightarrow \|U\| \leq \|V\|.$$

Applying this result to (8) we get

$$\|Z(t_2)\| \leq \|Z(t_1)\|$$

which contradicts (6) when $t_2 \uparrow b$. Suppose now that $Z(t) \leq 0$ for $a < t < b$. If $a < t_1 \leq t_2 < b$ then we obtain from (7) that $Z(t_2) \leq Z(t_1) \leq 0$ and so

$$0 \leq -Z(t_1) \leq -Z(t_2). \tag{9}$$

The monotonicity of the norm of (S,P) implies that $\|Z(t_1)\| \le \|Z(t_2\|$ and from this we obtain a contradiction of (5) when $t_1 \downarrow a$. **QED**

In case $A(t)$ is not symmetric, we have the following

Theorem 5. *Between any pair $a < b$ of conjugate points for (1) there is a focal point for (1) provided* Tr $A(t)$ *is negative semi-definite.*

Proof. Without loss of generality we may assume that a is the first conjugate point of b. Assume, for contradiction, that $\det Y_A(b,t) \neq 0$ for $a \le t \le b$. Theorem 2 implies that $\det X'_{A^T}(t,b) \neq 0$ for $a \le t \le b$. From Theorem 1 and $\det X'_A(b,b) = 1$ we obtain $\det X'_A(t,b) > 0$ for $a \le t \le b$. The Corollary of Theorem 3 implies that $\det X_A(t,b) > 0$ for $a < t < b$. From this and from the fact that $\det X_A(b,b) = 0 = \det X_A(a,b)$ it follows that there exists c, $a < c < b$, such that c is a point of relative maximum for the function $\det X_A(t,b)$. Since $\mathrm{Tr}A(t)$ is assumed negative semi-definite, Theorem 3 implies therefore that c is a point of strict relative minimum for the function $\det X_A(t,b)$. A contradiction. **QED**

§4 Convergent sequences of conjugate points for 2×2 systems.

In this section we assume that $A(t)$ is a continuous 2×2 matrix and that there exists a convergent sequences $(b_i)_i$ of conjugate points of a with respect to (1). We list some consequences of this assumption. Obviously, if we put $b := \lim_{i\to\infty} b_i$, then b is a conjugate point of a.

(i) The function $\det X_A(t,a)$ vanishes in every neighbourhood of b at a point distinct from b. So b cannot be a point of strict minimum or maximum for the function $\det X_A(t,a)$. This implies, by Theorem 3, that $\det X'_A(b,a) = 0$.

In particular, if $b < a$, then a is a focal point of b.

(ii) $\mathrm{Rank}X_A(b,a) = \mathrm{Rank}X'_A(b,a) = 1$. In fact from the assumption and from (i) we have that $\mathrm{Rank}X_A(b,a) < 2$ and $\mathrm{Rank}X'_A(b,a) < 2$. If $\mathrm{Rank}X_A(b,a) = 0$ or $\mathrm{Rank}X'_A(b,a) = 0$, then $\det Z(b) = 0$ where $Z(t)$ is the the solution of (2). But this cannot hold because $Z(a) = I$.

(iii) $\mathrm{Range}X_A(b,a) = \mathrm{Range}X'_A(b,a)$.

Proof. Let σ be a non-null vector orthogonal to $\mathrm{Range}X_A(b,a)$. Then (ii) implies that for every $\alpha \in \mathbf{R}^2$, $\alpha \in \mathrm{Range}X_A(b,a)$ if and only if $\alpha \perp \sigma$.

For every natural number i let ρ_i be a unit vector in $\mathrm{Ker}X_A(b_i,a)$. Without loss of generality we may assume that the sequence (ρ_i) is convergent. Put $\rho := \lim_{i\to\infty} \rho_i$. Obviously, $\rho \in \mathrm{Ker}X_A(b,a)$ and $|\rho| = 1$. Define

$$f_i(t) := (X_{A^T}(t,b)^T X'_A(t,a)\rho_i|\sigma) \qquad (i \ge 1)$$

Then $f_i(t)$ is a continuously differentiable real valued function. From $X_{A^T}(b,b) = 0$ we have that $f_i(b) = 0$. From Theorem 2, relation (a), and from $X_A(b_i,a)\rho_i = 0$, we have that

$$f_i(b_i) = (X_{A^T}(b_i,b)^T X'_A(b_i,a)\rho_i|\sigma) =$$

$$(X_{A^T}(b_i,b)^T X'_A(b_i,a)\rho_i - X'_{A^T}(b_i,b)^T X_A(b_i,a)\rho_i|\sigma) =$$
$$-(X_A(b,a)\rho_i|\sigma) = 0.$$

By Rolle's theorem there exists $c_i \in (min\{b,b_i\}, max\{b,b_i\})$ such that $f_i'(c_i) = 0$. This means that

$$-(X_{A^T}(c_i,b)^T A(c_i) X_A(c_i,a)\rho_i|\sigma) + (X'_{A^T}(c_i,b)^T X'_A(c_i,a)\rho_i|\sigma) = 0.$$

Obviously, $\lim_{i\to\infty} c_i = b$, and so, passing to the limit in the last relation, we obtain that $(X'_A(b,a)\rho|\sigma) = 0$, by virtue of $X_{A^T}(b,b) = 0$ and $X'_{A^T}(b,b) = I$. So $X'_A(b,a)\rho \in$ Range$X_A(b,a)$. The vector $X'_A(b,a)\rho$ must be different from zero because $X_A(t,a)\rho$ is a non-trivial solution of (1) vanishing for $t = b$. From (ii) it follows that Range$X'_A(b,a) =$ Range$X_A(b,a)$. **QED**

References

[1] M.A.Krasnosel'skij: *Positive linear system: the method of positive operators*, Heldermann, 1989-VII, p.354.

[2] C.R.Johnson / A.Horn: *Matrix analysis*, Cambridge University Press, Cambridge (1985), p.561.

[3] W.J.Reid: *Sturmian theory for ordinary differential equations*, Springer Verlag, Berlin (1971), p.559.