

## A DISCRETE LYAPUNOV FUNCTION FOR A CLASS OF LINEAR DIFFERENTIAL EQUATIONS

HAL L. SMITH

**An integer-valued function on  $\mathcal{R}^n$  is shown to decrease along trajectories of certain linear systems of ordinary differential equations.**

**1. Introduction.** Recently, there has been a great deal of success in obtaining very striking results on the global behavior of both finite and infinite dimensional dynamical systems through the use of integer-valued Lyapunov functions which decrease in value along trajectories. Matano [6] uses the lap number (originally discovered by Nickel [7]), a measure of the number of zeros of a solution of a scalar reaction diffusion equation on an interval, and the fact that it does not increase, to show that solutions converge to equilibria. Henry [3] and Angenent [1] exploit this further to obtain the Morse-Smale property for scalar reaction diffusion equations. For functional differential equations, Mallet-Paret [4] shows that a count of the number of zeros of a solution in an interval is an integer valued Lyapunov function for a certain class of equations. He exploits this fact to obtain a Morse decomposition of the global attractor. Smillie [8] uses the fact that the number of sign changes in the components  $\dot{x}_1, \dot{x}_2, \dots, \dot{x}_n$  does not increase with time for cooperative, tridiagonal systems of ordinary differential equations to conclude that all bounded solutions converge. Fusco and Oliva [2] use a similar integer-valued Lyapunov function to identify a class of ordinary differential equations with the Morse-Smale property. The author and Mallet-Paret [5] use such a function to obtain a Poincaré-Bendixson type result for a class of ordinary differential equations in  $\mathcal{R}^n$ ,  $n \geq 2$ . It is clear that these various (and related) integer-valued Lyapunov functions put severe constraints on the global asymptotic behavior of the various systems.

We note that for many of the applications cited above, the crucial point is to have an appropriate integer valued Lyapunov function for the variational equation along an orbit, or for a class of linear systems which contain the variational equation. For then one can argue that the Lyapunov function decreases along the difference of two solutions.

This paper will be devoted to showing that the integer-valued Lyapunov function described in [5] applies to a larger class of linear systems than those described there. Consider the linear system of ordinary differential equations

$$(1.1) \quad \dot{x}_i(t) = \delta_i a_{ii-1}(t)x_{i-1}(t) + a_{ii}(t)x_i(t) + \delta_{i+1} a_{ii+1}(t)x_{i+1}(t), \\ 1 \leq i \leq n, \quad i \text{ modulo } n.$$

Concerning (1.1), we assume hereafter that the  $a_{ij}$  are continuous functions and

$$\delta_i \in \{-1, +1\}, \quad a_{ii-1}(t) \geq 0, \quad a_{ii+1}(t) \geq 0, \\ 1 \leq i \leq n, \quad i \text{ modulo } n.$$

Further assumptions will be described below. Since it will be necessary throughout this paper to identify indices  $i$  modulo  $n$  (e.g.  $0 = n$ ,  $n + 1 = 1$ , etc.), we adopt it as a standing practice and make no further mention of it.

Following [5], define the function  $N$ , taking values in  $\{0, 1, 2, \dots, n\}$  by

$$N(y) = \text{card}\{i: \delta_i y_i y_{i-1} < 0\}, \quad y \in \mathfrak{R}^n, \quad y_i \neq 0 \text{ all } i.$$

It is easy to see that the domain of definition of  $N$  can be extended to the maximal domain  $\mathcal{N}$  on which  $N$  remains continuous where  $\mathcal{N}$  is the open set given by

$$\mathcal{N} = \{y \in \mathfrak{R}^n: \text{if } y_i = 0, \text{ then } \delta_{i+1} \delta_i y_{i+1} y_{i-1} < 0\}.$$

For those  $y$  with  $y_i \neq 0$ ,  $1 \leq i \leq n$ , we have

$$(-1)^{N(y)} = \text{sign} \prod_{i=1}^n \delta_i y_i y_{i-1} = \prod_{i=1}^n \delta_i \equiv \Delta.$$

Hence,  $N$  takes only even values if  $\Delta = +1$  and only odd values if  $\Delta = -1$ . We say (1.1) has positive (negative) feedback if  $\Delta = +1$  ( $\Delta = -1$ ). The significance of this terminology becomes apparent if we make the change of variables  $y_i = \mu_i x_i$  in (1.1) where  $\mu_i \in \{-1, +1\}$ . The resulting system for  $y$  is again of the form (1.1); the only change is that the new  $\bar{\delta}_i$  are given by  $\bar{\delta}_i = \mu_i \mu_{i-1} \delta_i$ . It is not hard to see that the  $\mu_i$  may be chosen so that  $\bar{\delta}_i = +1$ ,  $2 \leq i \leq n$ , and  $\bar{\delta}_1 = \Delta$ . If  $\Delta = +1$ , the resulting system has all positive feedbacks. If  $\Delta = -1$ , the system contains a pair of negative feedbacks.

In addition to the standing assumptions above, we make the following additional assumptions involving the matrix  $A(t) = (a_{ij}(t))$  where  $a_{ij} \equiv 0$  if  $j \notin \{i-1, i, i+1\}$ .

- (I)  $A(t)$  is irreducible for each  $t$ , and
- (V) If  $a_{ii-1}(t)$  ( $a_{ii+1}(t)$ ) vanishes at some  $t = t_0$  then

$$\lim_{t \rightarrow t_0} \frac{a_{ii-1}(t)}{(t - t_0)^{n-2}} = 0.$$

Given the special form of the matrix  $A(t)$ , (I) holds true if and only if one of the following holds for each value of  $t$ :

- (1.2) (a)  $\prod_{i=1}^n a_{ii-1}(t) > 0$ , or
- (b)  $\prod_{i=1}^n a_{ii+1}(t) > 0$ , or
- (c)  $\prod_{i=1}^{n-1} a_{ii+1}(t) \prod_{i=2}^n a_{ii-1}(t) > 0$ .

We remind the reader of our convention that indices are interpreted modulo  $n$  so that  $n + 1 \sim 1$  and  $0 \sim n$ .

The assumption (I) limits the set of functions  $a_{ii-1}(t)$ ,  $a_{ii+1}(t)$ ,  $1 \leq i \leq n$ , which can vanish at a particular value of  $t$ . The assumption (V) on the other hand, describes how  $a_{ii+1}(t)$  may vanish at a point. It requires that if  $a_{ii-1}$  ( $a_{ii+1}$ ) vanishes at  $t_0$  then it must vanish to sufficiently high order. Note that (V) certainly holds if

(V')  $a_{ii-1}(a_{ii+1})$  never vanishes or vanishes identically.

Our main result is the following (compare [2, 5, 8]).

**THEOREM.** *If (1.1) satisfies (I) and (V) and if  $x(t)$  is a nontrivial solution of (1.1) then*

- (i)  $x(t) \in \mathcal{N}$  except possibly for isolated values of  $t$ .
- (ii) If  $x(s) \notin \mathcal{N}$  then  $N(x(s+)) < N(x(s-))$ .

Assertion (i) implies that if  $x(s) \notin \mathcal{N}$  then  $x(t) \in \mathcal{N}$  for  $0 < |t - s| < \epsilon$  for some positive  $\epsilon$ . It follows from the continuity of  $N$  on  $\mathcal{N}$  that  $N(x(t))$  is constant on  $(s - \epsilon, s)$  and on  $(s, s + \epsilon)$ . It will be established that  $N$  decreases by a multiple of two as  $t$  increases through  $t = s$ .

The proof of the theorem will be deferred to §2. We make several remarks below.

The hypothesis that (I) holds is crucial for the theorem to hold.

Consider the constant coefficient system

$$\dot{x} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix} x$$

where  $a_{12} > 0$ ,  $a_{23} > 0$  and  $a_{32} > 0$ . Observe that (I) fails and that  $x(t) = e^{a_{11}t}(1, 0, 0)$  is a solution satisfying  $x(t) \notin \mathcal{N}$  for all  $t$ .

The assumption (V) is probably unnecessary. In §2 it is shown that if  $A(t)$  is analytic in  $t$  then (V) can be dropped entirely. We conjecture that the theorem holds without (V). Observe that if  $n = 3$  and  $A(t)$  is  $C^1$  in  $t$  then (V) holds. Indeed, in this case, if  $a_{ii-1}$  vanishes at  $t = t_0$  then  $\dot{a}_{ii-1}(t_0) = 0$  since  $a_{ii-1}(t) \geq 0$  so  $a_{ii-1}(t) = o(t - t_0)$  as  $t \rightarrow t_0$ .

The theorem implies the invariance of certain families of wedge-like subsets of  $\mathfrak{R}^n$  under the forward flow associated with (1.1). Let  $\Phi(t, s)$  be the fundamental matrix solution of (1.1) satisfying  $\Phi(s, s) = I$ , where  $I$  is the identity matrix. For definiteness, assume that  $\Delta = -1$  and that  $k$  is an odd integer less than  $n$ .

The set

$$H_k = \text{clos}\{x \in \mathfrak{R}^n : \text{there exists a neighborhood } U \text{ of } x \\ \text{such that } N(y) \leq k \text{ for every } y \in U \cap \mathcal{N}\}$$

is a nonempty set with the property that  $\alpha x \in H_k$  whenever  $x \in H_k$  and  $\alpha \neq 0$ . Here,  $\text{clos } B$  denotes the closure of the set  $B$  and  $\text{int } B$  denotes the interior of  $B$ . The theorem implies that

$$\Phi(t, s)(H_k \setminus \{0\}) \subset \text{int } H_k, \quad t > s.$$

This invariance property has strong implications for the Floquet theory for (1.1) if  $A(t)$  is  $T$ -periodic. For example,  $N$  is defined and constant on the set of nonzero vectors belonging to the real part of the direct sum of those generalized eigenspaces of  $\Phi(T + s, s)$  corresponding to Floquet multipliers having a fixed modulus. Moreover, such subspaces can be at most two dimensional (see [5, Lemma 2.2 and Lemma 2.5]).

Our theorem has strong implications for the nonlinear system of differential equations

$$(1.3) \quad \dot{y}_i(t) = f_i(y_{i-1}(t), y_i(t), y_{i+1}(t))$$

in the case that

$$(1.4) \quad \delta_i \frac{\partial f_i}{\partial x_{i-1}} \geq 0, \quad \delta_{i+1} \frac{\partial f_i}{\partial x_{i-1}} \geq 0$$

hold throughout an invariant region for (1.3). Here, we merely note the more obvious implications and refer the reader to [2, 5, 8] for more thorough considerations under different hypotheses.

In order to simplify our exposition, assume that the  $f_i$  are analytic functions and that the Jacobian matrix of the vector field is an irreducible matrix for each  $x$ . If  $y(t)$  and  $\bar{y}(t)$  are distinct solutions of (1.3) then  $x(t) = y(t) - \bar{y}(t)$  is a nontrivial solution of (1.1) where

$$a_{ij}(t) = \int_0^1 \frac{\partial f^i}{\partial y^j}(u_{i-1}(s, t), u_i(s, t), u_{i+1}(s, t)) ds,$$

$$u_i(s, t) = sy_i(t) + (1 - s)\bar{y}_i(t).$$

Observe that  $A(t)$  is analytic and (I) holds. Similarly, if  $y(t)$  is a nonconstant solution of (1.3) then  $\dot{y}(t)$  satisfies the variational equation which has the form (1.1). It follows from the theorem that the function  $t \mapsto N(y(t) - \bar{y}(t))$  is defined except possibly at a finite number of points (no more than  $[n/2] + 1$  points) of the common domain,  $I$ , of  $y(\cdot)$  and  $\bar{y}(\cdot)$ . Points, if any, at which the function is not defined, separate adjacent open subintervals of  $I$  on each of which the function  $N$  is constant and  $N$  decreases by a multiple of two as  $t$  increases through such a point.

If  $y(t)$  and  $\bar{y}(t)$  are defined for  $t \geq 0$  then the theorem implies that  $y(t) - \bar{y}(t) \in \mathcal{N}$  for all large  $t$ , say  $t > t_0$ . This in turn implies that the two curves  $t \rightarrow (t, y_i(t), y_{i+1}(t))$  and  $t \rightarrow (t, \bar{y}_i(t), \bar{y}_{i+1}(t))$  do not meet for each  $i$  and suggest the possibility of a phase plane analysis (see [5, sec. 3]).

**2. Proof of the Theorem.** This section will be devoted to the proof of the theorem. We must show that if  $x(s) \notin \mathcal{N}$  then there exists  $\varepsilon > 0$  such that  $x(t) \in \mathcal{N}$  for  $0 < |t - s| < \varepsilon$  and that (ii) holds. Without loss of generality we may assume that  $s = 0$  for we can always perform a translation in  $t$ . We begin with some definitions similar to those in [8] except that we do not require differentiability.

**DEFINITION.** For  $1 \leq i \leq n$ , define  $k(i) = k$ , a nonnegative integer, if there exists a real number  $p_i \neq 0$  such that

$$x_i(t) = p_i t^k + o(t^k), \quad \text{as } t \rightarrow 0.$$

That is,

$$\lim_{t \rightarrow 0} \frac{x_i(t)}{t^k} = p_i.$$

If there does not exist such a  $k$  and  $p_i$  then  $k(i)$  is not defined. Let

$$P_i = \text{sign } p_i = \begin{cases} +1 & \text{if } p_i > 0, \\ -1 & \text{if } p_i < 0. \end{cases}$$

Observe that  $k(i)$  and  $p_i$  are uniquely defined and that if  $k(i)$  is defined, then  $t = 0$  is an isolated zero of  $x_i(t)$ .

The following lemma will be the key to our proof.

**LEMMA.** *Let  $y(t)$  be the solution of the initial value problem*

$$(2.1) \quad \dot{y}(t) = B(t)y + g(t), \quad y(0) = 0$$

where  $B(t)(g(t))$  is a continuous matrix (vector)-valued function on some neighborhood of  $t = 0$ . Let  $\phi(t)$  be a vector-valued continuous function on some neighborhood of  $t = 0$  and let  $\mu(t)$  be a positive scalar-valued continuous function on a neighborhood of  $t = 0$  such that  $|\phi(t)| \leq \mu(t)$  and

$$\lim_{t \rightarrow 0} \frac{|g(t) - \phi(t)|}{\mu(t)} = 0.$$

Let  $\psi(t) = \int_0^t \phi(s) ds$  and  $\chi(t) = |\int_0^t \mu(s) ds|$ .

Then

$$\lim_{t \rightarrow 0} \frac{|y(t) - \psi(t)|}{\chi(t)} = 0.$$

*Proof.* The solution  $y(t)$  satisfies the integral equation

$$y(t) = \int_0^t g(s) ds + \int_0^t B(s)y(s) ds.$$

Hence for  $t > 0$ ,

$$\begin{aligned} |y(t) - \psi(t)| &\leq \int_0^t |g(s) - \phi(s)| ds + \int_0^t |B(s)||\psi(s)| ds \\ &\quad + \int_0^t |B(s)||y(s) - \psi(s)| ds. \end{aligned}$$

By Gronwall's inequality, for  $t > 0$ ,

$$|y(t) - \psi(t)| \leq \alpha(t) + \int_0^t |B(s)|\alpha(s) \left( \exp \int_0^t |B(u)| du \right) ds$$

where

$$\alpha(t) \equiv \int_0^t |g(s) - \phi(s)| ds + \int_0^t |B(s)||\psi(s)| ds.$$

Hereafter, we consider only  $t > 0$  as a similar argument applies when  $t < 0$ . Now, if  $B_m = \sup|B(s)|$  over some interval containing zero then

$$\begin{aligned} \alpha(t) &\leq \int_0^t |g(s) - \phi(s)| ds + B_m \int_0^t |\psi(s)| ds \\ &\leq \int_0^t \frac{|g(s) - \phi(s)|}{\mu(s)} \mu(s) ds + B_m \int_0^t \int_0^s |\phi(r)| dr ds \\ &\leq \left[ \sup_{0 \leq s \leq t} \frac{|g(s) - \phi(s)|}{\mu(s)} \chi(t) + B_m \int_0^t \int_0^s \mu(r) dr ds \right] \\ &\leq \left[ \sup_{0 \leq s \leq t} \frac{|g(s) - \phi(s)|}{\mu(s)} + B_m t \right] \chi(t). \end{aligned}$$

Hence

$$\lim_{t \rightarrow 0^+} \frac{\alpha(t)}{\chi(t)} = 0.$$

Similarly,

$$\begin{aligned} \int_0^t |B(s)|\alpha(s) ds \exp \int_0^t |B(u)| du &\leq B_m e^{B_m t} \cdot \int_0^t \alpha(s) ds \\ &\leq B_m e^{B_m t} t \alpha(t) \leq B_m e^{B_m t} t \frac{\alpha(t)}{\chi(t)} \chi(t). \end{aligned}$$

Hence

$$\lim_{t \rightarrow 0^+} \frac{1}{\chi(t)} \int_0^t |B(s)|\alpha(s) ds \exp \int_0^t |B(u)| du = 0.$$

This completes our proof. □

An important special case of the lemma occurs when  $g(t) = g_m t^m + o(t^m)$  as  $t \rightarrow 0$ , where  $m$  is a nonnegative integer and  $g_m \in \mathfrak{R}^n$ . Put  $\phi(t) = g_m t^m$  and  $\mu(t) = \mu_0 |t|^m$ ,  $\mu_0 > |g_m|$ . Then the lemma implies that

$$y(t) = \frac{g_m}{m+1} t^{m+1} + o(t^{m+1}).$$

Returning to the proof of the theorem, let  $Z = \{i: x_i(0) = 0\}$  and observe that  $Z$  is a proper subset of  $\{1, 2, \dots, n\}$  since  $x(0) \neq 0$  but  $Z$  is nonempty since  $x(0) \notin \mathcal{N}$ . Partition  $Z$  into pairwise disjoint intervals  $I_1, I_2, \dots, I_l$ ,  $l \geq 1$ , such that each set  $I_j \subset Z$  consists of consecutive integers (mod  $n$ ),  $I_j = \{i+1, i+2, \dots, i+p\}$ ,  $p \geq 1$ , and  $i \notin Z$ ,  $i+p+1 \notin Z$ . Note that such an interval  $I$  may be of the form  $I = \{l, l+1, \dots, n, 1, 2, \dots, r\}$  where  $r < l$ .

The proof of the theorem involves considering the individual intervals  $I_j$  separately and determining their contribution, if any, to the change in the value of  $N$  as  $t$  increases through zero. The simplest type of interval to consider and the most important is the case that  $I = \{j\}$  is a singleton interval (we will drop subscripts on the intervals when no confusion is possible). Since  $x_{j-1}(0)x_{j+1}(0) \neq 0$ , there are two cases. If  $\delta_{j+1}\delta_j P_{j+1}P_{j-1} = -1$ , then the interval  $I$  contributes no change to the value of  $N$  as  $t$  increases through zero (see the definition of  $N$  and recall how  $\mathcal{N}$  is defined). Moreover, by continuity of  $x(t)$ ,  $\delta_{j+1}\delta_j x_{j+1}x_{j-1}(t) < 0$  in a neighborhood of zero. It is important to observe that not every interval  $I_j$ ,  $j = 1, 2, \dots, l$ , can be of this type since  $x(0) \notin \mathcal{N}$ . Suppose now that  $\delta_{j+1}\delta_j P_{j+1}P_{j-1} = +1$ . Then

$$\dot{x}_j(0) = \delta_j a_{jj-1}(0)x_{j-1}(0) + \delta_{j+1} a_{jj+1}(0)x_{j+1}(0) \neq 0$$

since (I) does not permit both  $a_{jj-1}(0)$  and  $a_{jj+1}(0)$  to vanish. Moreover

$$\left. \frac{d}{dt} \right|_{t=0} \delta_j x_j x_{j-1} = a_{jj-1} x_{j-1}^2 + \delta_j \delta_{j+1} a_{jj+1} x_{j-1} x_{j+1} > 0$$

and

$$\left. \frac{d}{dt} \right|_{t=0} \delta_{j+1} x_{j+1} x_j = \delta_{j+1} \delta_j a_{jj-1} x_{j-1} x_{j+1} + a_{jj+1} x_{j+1}^2 > 0.$$

Hence we see that there exists  $\varepsilon > 0$  such that  $x_j(t) \neq 0$  for  $0 < |t| < \varepsilon$  and that the interval  $I = \{j\}$  contributes to a decrease in  $N$  by two as  $t$  increases through zero. Note that this argument did not require that (V) hold.

The remainder of our proof of the theorem is somewhat tedious because we are unable to treat intervals  $I$  in the decomposition of  $Z$  which are not singleton intervals in the generality with which we could deal with singleton intervals. There is a plethora of special cases.

We begin by considering the special case that

$$(2.2) \quad a_{ii-1}(0) > 0, \quad a_{ii+1}(0) > 0, \quad 1 \leq i \leq n.$$

Let  $I$  be an interval in the partition of  $Z$  of even length,  $I = \{j+1, j+2, \dots, j+p\}$ ,  $p = 2q$ ,  $q \geq 1$ . Then we have the following.



*Claim.* (a)  $k(j+r) = r$ ,  $0 \leq r \leq q$ , and  $P_{j+r} = \delta_{j+r}P_{j+r-1}$ ,  $1 \leq r \leq q$ .

(b)  $k(j+p+1-s) = s$ ,  $0 \leq s \leq q$ , and  $P_{j+p+1-s} = \delta_{j+p+1-s+1}P_{j+p+1-s+1}$ ,  $1 \leq s \leq q$ .

*Proof.* Since  $\dot{x}_{j+1}(0) = \delta_{j+1}a_{j+1j}(0)x_j(0) \neq 0$ ,  $k(j+1) = 1$ ,  $p_{j+1} = \delta_{j+1}a_{j+1j}(0)p_j$ , so  $P_{j+1} = \delta_{j+1}P_j$ . Similarly,

$$x_{j+p}(t) = \delta_{j+p+1}a_{j+p, j+p+1}(0)p_{j+p+1}t + o(t)$$

so  $k(j+p) = 1$  and  $P_{j+p} = \delta_{j+p+1}P_{j+p+1}$ . Hence (a) and (b) hold for  $r, s = 1$ . Suppose (a) and (b) holds for  $s, r \leq k < q$ . We will show that (a) and (b) hold for  $s, r \leq k+1$ . Let  $y(t) = (x_{j+k+1}(t), x_{j+k+2}(t), \dots, x_{j+p-k}(t))$ . Then from (1.1) and our induction hypothesis  $y(t)$  satisfies

$$\dot{y} = B(t)y + g(t), \quad y(0) = 0$$

where

$$\begin{aligned} g(t) &= (\delta_{j+k+1}a_{j+k+1, j+k}x_{j+k}, 0, \dots, 0, \\ &\quad \delta_{j+p-k+1}a_{j+p-k, j+p-k+1}x_{j+p-k+1}) \\ &= (\delta_{j+k+1}a_{j+k+1, j+k}(0)p_{j+k}, 0, \dots, 0, \\ &\quad \delta_{j+p-k+1}a_{j+p-k, j+p-k+1}(0)p_{j+p-k+1})t^k + o(t^k). \end{aligned}$$

By the lemma,

$$y(t) = \frac{1}{k+1}t^{k+1}g_k + o(t^{k+1}), \quad t \rightarrow 0$$

where  $g_k$  is the vector coefficient of  $t^k$  in the above expansion of  $g(t)$ .

In particular,  $k(j+k+1) = k+1 = k(j+p-k)$  and

$$\begin{aligned} P_{j+k+1} &= \delta_{j+k+1}P_{j+k}, \\ P_{j+p-k} &= \delta_{j+p-k+1}P_{j+p-k+1}. \end{aligned}$$

This establishes our claim.

As an immediate consequence,  $x_i(t)$ ,  $i \in I$ , have isolated zeros at  $t = 0$ . Also, we have for  $t \neq 0$  sufficiently small

$$\begin{aligned} \text{sign}[\delta_{j+1}x_{j+1}x_j] &= \text{sign}[\delta_{j+1}P_{j+1}P_j t] \\ &= \delta_{j+1}P_{j+1}P_j \text{sign } t \\ &= \delta_{j+1}\delta_{j+1}P_jP_j \text{sign } t = \text{sign } t, \\ \text{sign}[\delta_{j+2}x_{j+2}x_{j+1}] &= \delta_{j+2}P_{j+2}P_{j+1} \text{sign}[t^{2+1}] \\ &= \delta_{j+2}\delta_{j+2}P_{j+1}P_{j+1} \text{sign } t^3 \\ &= \text{sign } t^3, \\ &\vdots \\ \text{sign}[\delta_{j+q}x_{j+q}x_{j+q-1}] &= \delta_{j+q}P_{j+q}P_{j+q-1} \text{sign } t^{2q-1} \\ &= \text{sign } t^{2q-1}, \\ \text{sign}[\delta_{j+q+1}x_{j+q+1}x_{j+q}] &= \delta_{j+q+1}P_{j+q+1}P_{j+q} \text{sign } t^{2q} \\ &= \delta_{j+q+1}P_{j+q+1}P_{j+q}, \\ \text{sign}[\delta_{j+q+2}x_{j+q+2}x_{j+q+1}] &= \delta_{j+q+2}P_{j+q+2}P_{j+q+1} \text{sign } t^{2q-1} \\ &= \delta_{j+q+2}P_{j+q+2}\delta_{j+q+2}P_{j+q+2} \text{sign } t^{2q-1} \\ &= \text{sign } t^{2q-1}, \\ &\vdots \\ \text{sign}[\delta_{j+q+1}x_{j+p+1}x_{j+p}] &= \delta_{j+p+1}P_{j+p+1}P_{j+p} \text{sign } t \\ &= \text{sign } t. \end{aligned}$$

Thus the interval  $I$  of size  $p = 2q$  contributes to a decrease in  $N$  of  $p = 2q$  as  $t$  increases through zero.

Next we consider the case that  $I$  has odd length,  $I = \{j + 1, j + 2, \dots, j + p\}$ ,  $p = 2q + 1$ ,  $q \geq 1$ . Observe that the previous claim still holds. Furthermore,  $y(t) \equiv x_{j+q+1}(t)$  satisfies

$$\dot{y}(t) = a_{j+q+1, j+q+1}(t)y(t) + g(t), \quad y(0) = 0,$$

where  $g(t) = ht^q + o(t^q)$  and

$$h = \delta_{j+q+1}a_{j+q+1, j+q}(0)p_{j+q} + \delta_{j+q+2}a_{j+q+1, j+q+2}(0)p_{j+q+2}.$$

There are three cases to consider. We begin with

$$\text{Case I. } \delta_{j+q+1}P_{j+q}\delta_{j+q+2}P_{j+q+2} = +1.$$

In this case  $h \neq 0$  and by the lemma

$$y(t) = x_{j+q+1}(t) = \frac{h}{q+1}t^{q+1} + o(t^{q+1}).$$

In particular it follows that  $t = 0$  is an isolated zero of  $x_i(t)$  for every  $i \in I$ . We also have  $k(j + q + 1) = q + 1$ ,  $P_{j+q+1} = \delta_{j+q+1}P_{j+q} = \delta_{j+q+2}P_{j+q+2}$ . Hence, for  $t \neq 0$  sufficiently small,

$$\begin{aligned} \text{sign}[\delta_{j+q+1}x_{j+q+1}x_{j+q}] &= \delta_{j+q+1}P_{j+q+1}P_{j+q} \text{sign } t^{2q+1} \\ &= \text{sign } t^{2q+1} \end{aligned}$$

and

$$\begin{aligned} \text{sign}[\delta_{j+q+2}x_{j+q+2}x_{j+q+1}] &= \delta_{j+q+2}P_{j+q+2}P_{j+q+1} \text{sign } t^{2q+1} \\ &= \text{sign } t^{2q+1}. \end{aligned}$$

Thus  $I$  contributes to a decrease in  $N$  of  $2q+2 = p+1$  as  $t$  increases through  $t = 0$  in Case I.

*Case II.*  $\delta_{j+q+1}P_{j+q}\delta_{j+q+2}P_{j+q+2} = -1$  but  $h \neq 0$ .

As in Case I,  $t = 0$  is still an isolated zero of  $x_i(t)$ ,  $i \in I$ . The only change in the argument of Case I is that for  $t \neq 0$ ,  $t$  small,  $\delta_{j+q+1}x_{j+q+1}x_{j+q}$  and  $\delta_{j+q+2}x_{j+q+2}x_{j+q+1}$  have opposite signs. Thus these two terms cancel each other and the change in  $N$  due to interval  $I$  as  $t$  increases through  $t = 0$  is  $p - 1$ .

*Case III.*  $\delta_{j+q+1}P_{j+q}\delta_{j+q+2}P_{j+q+2} = -1$  and  $h = 0$ .

In this case, the "order" of the zero of  $x_{j+q+1}$  at  $t = 0$  is indeterminate. For  $t \neq 0$  and  $t$  sufficiently small, we have

$$\begin{aligned} \text{sign}[\delta_{j+q+2}\delta_{j+q+1}x_{j+q+2}x_{j+q}] \\ = \delta_{j+q+2}\delta_{j+q+1}P_{j+q+2}P_{j+q} \text{sign } t^{2q} = -1. \end{aligned}$$

That is,  $\delta_{j+q+2}\delta_{j+q+1}x_{j+q+2}x_{j+q} < 0$  for  $|t| > 0$ ,  $|t|$  small. Thus, from the definition of  $N$  on  $\mathcal{N}$ ,  $x_{j+q+1}(t)$  has no effect on the value of  $N$  for  $t \neq 0$ ,  $t$  small. As in the previous case, the interval  $I$  contributes a decrease of  $p - 1$  to  $N$  as  $t$  increases through  $t = 0$ .

In summary, when (2.2) holds, we have established that there exists  $\varepsilon > 0$  such if  $0 < |t| < \varepsilon$  then  $x(t) \in \mathcal{N}$ . This follows from our consideration of each type of interval in the partition of  $Z$ . Furthermore, except in the case of a singleton interval  $I = \{j\}$  where  $\delta_{j+1}\delta_jP_{j+1}P_{j-1} = -1$ , all other types of intervals contribute a net decrease in  $N$  by an even integer as  $t$  increases through  $t = 0$ . We observed that the special singleton interval described above contributes no change in  $N$  and further, that not every interval in the partition of  $Z$  could be of this type. It follows that our proof of the theorem is complete in the case that (2.2) holds.

We now drop the assumption that (2.2) holds and consider the case where (I) and (V') hold. Later, we remark on the change in the following argument when (V) replaces (V'). Since (I) holds we may assume that one of (1.2)(a), (b) or (c) holds. We begin by assuming that (1.2)(b) holds and show that there exists  $\varepsilon > 0$  such that  $x(t) \in \mathcal{N}$  for  $0 < |t| < \varepsilon$ . This will establish (i) of the theorem in this case. The argument in case (1.2)(a) holds is very similar and will not be given. We consider the case that (1.2)(c) holds in a separate argument.

Assume that (1.2)(b) holds and let  $I$  be an interval in the partition of  $Z$  with length greater than one,  $I = \{j+1, j+2, \dots, j+p\}$ ,  $p \geq 2$ . Now we know that  $a_{ii+1} > 0$  for every  $i$  by (1.2)(b) but  $a_{ii-1}$  may vanish identically for some  $i \in I$ . If  $a_{ii-1} > 0$  for every  $i \in I$  then, of course, we may argue exactly as when (2.2) held. Thus we assume that  $a_{ii-1}$  vanishes identically for some  $i \in I$ .

Let  $j+r$  be the smallest index in  $I$  such that  $a_{j+r, j+r-1} \equiv 0$ . Letting  $y = (x_{j+r}, \dots, x_{j+p})$ , equation (1.1) implies that  $y$  satisfies an initial value problem (2.1) where

$$g(t) = (0, 0, \dots, 0, \delta_{j+p+1} a_{j+p, j+p+1}(0) x_{j+p+1}(0)) + o(1).$$

Hence the lemma implies that  $k(j+p) = 1$  and  $P_{j+p} = \delta_{j+p+1} P_{j+p+1}$ . Assume that we have established that  $k(j+p-s) = s+1$ ,  $P_{j+p-s} = \delta_{j+p-s+1} P_{j+p-s+1}$  for  $0 \leq s \leq l$ , where  $j+p-l > j+r$ . Again, let  $y = (x_{j+r}, \dots, x_{j+p-l-1})$ . Then  $y$  satisfies an initial value problem (2.1) where  $g(t) = (0, 0, \dots, 0, \delta_{j+p-l} a_{j+p-l-1, j+p-l}(0) p_{j+p-l}) t^{l+1} + o(t^{l+1})$ . By the lemma, it follows that  $k(j+p-l-1) = l+2$  and  $P_{j+p-l-1} = \delta_{j+p-l} P_{j+p-l}$ . Hence we have established that  $k(j+p-s) = s+1$  and  $P_{j+p-s} = \delta_{j+p-s+1} P_{j+p-s+1}$  for  $0 \leq s \leq p-r$ . If  $r = 1$  then we have established that  $t = 0$  is an isolated zero for each  $x_i(t)$ ,  $i \in I$ . If  $r > 1$  then we must consider  $x_i(t)$  for  $i$  in the subinterval  $I' = \{j+1, \dots, j+r-1\}$  of  $I$ . Now we have  $j \notin Z$ ,  $j+r \in Z$  and  $y_{j+r} = p_{j+r} t^{k(j+r)} + o(t^{k(j+r)})$  where  $k(j+r) = p-r+1$  and  $P_{j+r} = \delta_{j+r+1} P_{j+r+1}$ . In addition, we have  $a_{ii-1} > 0$ ,  $a_{ii+1} > 0$  for every  $i \in I'$  by the definition of  $j+r$ . Now we may continue our application of the lemma, setting  $y = (x_{j+1}, \dots, x_{j+r-1})$  and observe that  $y$  satisfies an initial value problem (2.1) where  $g(t) = (\delta_{j+1} a_{j+1j}(0) p_j, 0, 0, \dots, 0) + o(1)$  as  $t \rightarrow 0$ . Thus  $y(t) = g(0)t + o(t)$ , by the lemma, and so  $k(j+1) = 1$  and  $P_{j+1} = \delta_{j+1} P_j$ . If  $1 < k(j+r)$  then we apply the lemma again to  $y = (x_{j+2}, \dots, x_{j+r-1})$  where now  $g(t) = (\delta_{j+2} a_{j+2, j+1} p_{j+1}, 0, 0, \dots, 0)t + o(t)$ . Thus  $k(j+2) = 2$  and

$P_{j+2} = \delta_{j+2}P_{j+1}$ . Clearly, we may continue this application of the lemma, obtaining  $k(j+s) = s$  and  $P_{j+s} = \delta_{j+s}P_{j+s-1}$  so long as  $s \leq k(j+r) = p-r+1$ . If  $p = 2q+1$  is odd and  $j+r \leq j+q+1$  then  $k(j+r) \geq p-q = q+1$  so  $q \leq k(j+r)$ . Hence  $k(j+s) = s$  for  $1 \leq s \leq j+r-1 \leq j+q$ . In this case, then,  $t = 0$  is an isolated zero of  $x_i(t)$  for each  $i \in I$ . However, if  $j+r > j+q+1$  then we have  $k(j+s) = s$ ,  $P_{j+s} = \delta_{j+s}P_{j+s-1}$  for  $1 \leq s \leq p-r+1$  where  $p-r+1 < q+1$ . In particular  $k(j+p-r+1) = p-r+1 = k(j+r)$ . Let  $y = (x_{j+p-r+2}, \dots, x_{j+r-1})$  so  $y$  consists of an odd number of components and note that  $y$  satisfies an initial value problem (2.1) where

$$(2.3) \quad g(t) = (\delta_{j+p-r+2}a_{j+p-r+2, j+p-r+1}(0)p_{j+p-r+1}, \\ 0, \dots, 0, \delta_{j+r}a_{j+r-1, j+r}(0)p_{j+r})t^{p-r+1} + o(t^{p-r+1}).$$

Hence  $k(j+p-r+2) = k(j+r-1) = p-r+2$ . Since  $y$  contains an odd number of components, we may continue the application of the lemma until we have  $k(j+s) = s$ ,  $P_{j+s} = \delta_{j+s}P_{j+s-1}$ , for  $1 \leq s \leq q$ , and  $k(j+p-s) = s+1$ ,  $P_{j+p-s} = \delta_{j+p-s+1}P_{j+p-s+1}$ , for  $0 \leq s \leq q-1$ . Only  $x_{j+q+1}$  remains to be investigated. As usual, set  $y = x_{j+q+1}$  and observe that  $y$  satisfies a scalar initial value problem (2.1) where

$$g(t) = [\delta_{j+q+1}a_{j+q+1, j+q}(0)p_{j+q} + \delta_{j+q+2}a_{j+q+1, j+q+2}(0)p_{j+q+2}]t^q \\ + o(t^q).$$

If the quantity in brackets does not vanish, then the lemma implies that  $k(j+q+1) = q+1$  and we are done since  $t = 0$  is an isolated zero of  $x_i(t)$  for every  $i \in I$ . If the quantity in brackets does vanish then  $\delta_{j+q+1}P_{j+q}\delta_{j+q+2}P_{j+q+2} = -1$  so  $\delta_{j+q+1}\delta_{j+q+2}x_{j+q}(t)x_{j+q+2}(t) = \delta_{j+q+1}\delta_{j+q+2}p_{j+q}p_{j+q+2}t^{2q} + o(t^{2q}) < 0$  in some deleted neighborhood of  $t = 0$ . Thus, although we are ignorant of the order of the zero of  $x_{j+q+1}$  at  $t = 0$ , the index  $j+q+1$  cannot be a cause for  $x(t) \notin \mathcal{N}$ ,  $|t|$  small.

If  $I$  contains an even number of elements  $p = 2q$  and if  $j+r \leq j+q+1$  then  $k(j+r) \geq q$ . Hence  $k(j+s) = s$  for  $0 \leq s \leq j+r-1$  and so again  $t = 0$  is an isolated zero of  $x_i(t)$ ,  $i \in I$ . However, if  $j+r > j+q+1$  then we have  $k(j+s) = s$ ,  $P_{j+s} = \delta_{j+s}P_{j+s-1}$  for  $0 \leq s \leq p-r+1$ . Let  $y = (x_{j+p-r+2}, \dots, x_{j+r-1})$  so  $y$  has an even number of components. Since  $y$  satisfies an initial value problem (2.1) with  $g(t)$  as in (2.3) the lemma gives that  $k(j+p-r+2) = k(j+r-1) = p-r+2$ . Since  $y$  contains an even number of elements, we can continue the application of the lemma

until we obtain  $k(j+s) = s$ ,  $0 \leq s \leq q$ , and  $k(j+p-s) = s+1$ ,  $0 \leq s \leq q-1$ . Thus we find that  $t = 0$  is an isolated zero of  $x_i(t)$  for  $i \in I$ .

In summary, if (1.2)(b) holds, we have shown that each interval  $I$  in the decomposition of  $Z$  with length larger than one (length one intervals were considered quite generally in an earlier argument) has the property that either (a)  $t = 0$  is an isolated zero of  $x_i(t)$  for every  $i \in I$  or (b)  $t = 0$  is an isolated zero of  $x_i(t)$  for every  $i \in I$  except one (the middle one) where  $I$  has odd length, and for this  $i$  it happens that  $\delta_{i+1}\delta_i x_{i+1}x_{i-1} < 0$  in a deleted neighborhood of  $t = 0$ . Hence we may conclude that there exists  $\varepsilon > 0$  such that  $x(t) \in \mathcal{N}$  for  $0 < |t| < \varepsilon$  if (1.2)(b) holds.

Let us now consider the case that (1.2)(c) holds. Again, we focus on the intervals  $I$  making up the partition of  $Z$ . Let  $I$  be such an interval with length exceeding one,  $I = \{j+1, j+2, \dots, j+p\}$ ,  $p \geq 2$ . If  $I \subseteq \{2, 3, \dots, n-1\}$  then  $a_{ii-1}$  and  $a_{ii+1}$  are positive for every  $i \in I$ , by (1.2)(c), so the same arguments apply to  $I$  as in the case that (2.2) held. Hence we assume that either 1 or  $n$  or both belong to  $I$ . If 1 is a left endpoint of  $I$  then  $n \notin I$  and the argument that  $t = 0$  is an isolated zero for  $x_i(t)$  for each  $i \in I$  is exactly as in the previous case (that (1.2)(b) holds) where  $j+r = 1$  (recall  $j+r$  was the minimal element of  $I$  for which  $a_{j+r, j+r-1} \equiv 0$ ). Similarly, if  $n$  is the right-hand endpoint of  $I$ , then one argues that  $t = 0$  is an isolated zero of  $x_i(t)$  for each  $i \in I$ . Hence the only case left to check is that both 1 and  $n$  belong to  $I$ :  $I = \{j+1, j+2, \dots, n, 1, 2, \dots, r\}$ ,  $r < j+1$ . But such an interval can be decomposed into two subintervals,  $I_1 = \{j+1, j+2, \dots, n\}$  and  $I_2 = \{1, 2, \dots, r\}$ , each of which can be treated as described above where  $n$  is a right endpoint or 1 is a left endpoint. One obtains that  $k(j+1) = 1$ ,  $k(j+2) = 2, \dots, k(n) = n-j$  and  $k(r) = 1$ ,  $k(r-1) = 2, \dots, k(1) = r$ . In particular,  $t = 0$  is an isolated zero of  $x_i(t)$  for each  $i \in I$  in this case.

In summary, if (1.2)(c) holds then there exists  $\varepsilon > 0$  such that  $x(t) \in \mathcal{N}$  for  $0 < |t| < \varepsilon$ .

We now have established that (i) of the theorem holds provided (I) and (V') hold. The reader will see that, in our arguments so far, we do not really require  $a_{ij-1}(a_{ii+1})$  to vanish identically if it vanishes. Our arguments require no modification if whenever  $a_{ii-1}$  vanishes, it vanishes to a large enough order that the term  $a_{ii-1}(t)x_{i-1}(t)$  is of higher order than any "competing" term. This will ensure that

when we construct  $g(t)$  in order to apply the lemma, we can ignore  $a_{ii-1}(t)x_{i-1}(t)$ . It is not hard to see that if  $a_{ii-1}(t) = o(t^{n-2})$ , when it vanishes, then it has no effect on the arguments given above. Thus, (i) of the theorem holds provided (I) and (V) hold.

We now complete the proof of the theorem by showing that (ii) holds. This is done by approximating (1.1) by systems (1.1) $_m$ ,  $m = 1, 2, \dots$ , for which (2.2) holds. Recall we have established the theorem in case (2.2) holds. Thus, let  $A^m(t)$  be a sequence of continuous matrix functions having the same structure as  $A(t)$ , namely, (I) and (2.2) hold. Suppose that  $A^m(t) \rightarrow A(t)$  on some open interval containing  $t = 0$ . Let  $x^m(t)$  be the solution of the corresponding equation (1.1) where  $A^m$  replaces  $A$  and  $x^m(0) = x(0)$ ,  $m = 1, 2, \dots$ . Then,  $x^m(t) \rightarrow x(t)$  as  $m \rightarrow \infty$  uniformly on some interval containing  $t = 0$ . Now there exists  $\varepsilon > 0$  such that  $x(t) \in \mathcal{N}$  for  $0 < |t| < \varepsilon$ , since (i) of the theorem holds for  $x(t)$ . Since  $x^m(\pm\varepsilon/2) \rightarrow x(\pm\varepsilon/2)$  as  $m \rightarrow \infty$  and  $\mathcal{N}$  is an open set, it follows that  $x^m(\pm\varepsilon/2) \in \mathcal{N}$  for all large  $m$ . For these  $m$ ,  $N(x^m(-\varepsilon/2)) - N(x^m(\varepsilon/2)) \geq 2$  since  $x^m(0) \notin \mathcal{N}$  and the theorem applies to the  $x^m$ . By continuity of  $N$  on  $\mathcal{N}$ , it follows that  $N(x(-\varepsilon/2)) - N(x(\varepsilon/2)) \geq 2$ . Since  $N(x(t))$  is defined and constant on  $(-\varepsilon, 0)$  and on  $(0, \varepsilon)$ , we have  $N(x(0-)) > N(x(0+))$  as asserted. This concludes the proof of the theorem.

We conclude this section by showing that if  $A(t)$  is analytic then the theorem holds without the hypothesis that (V) hold. It is sufficient to show that (i) of the theorem holds since (ii) may be obtained by approximation as in the proof just completed. Suppose  $x(t)$  is a nontrivial solution of (1.1),  $x(t_0) \notin \mathcal{N}$  and further that there exists a sequence  $t_m \rightarrow t_0$ ,  $t_m \neq t_0$ , such that  $x(t_m) \notin \mathcal{N}$  for  $m = 1, 2, \dots$ . We will reach a contradiction by showing that  $x(t)$  vanishes identically. Since  $x(t_m) \notin \mathcal{N}$ , there exists  $i_m$  such that  $x_{i_m}(t_m) = 0$  and  $\delta_{i_m+1}\delta_{i_m}x_{i_m+1}(t_m)x_{i_m-1}(t_m) \geq 0$  for  $m \geq 1$ . For some  $j$  we have  $i_m = j$  for infinitely many  $m$  so without loss of generality we assume  $i_m = j$  for all  $m \geq 1$ . By analyticity of  $x_j(t)$ , it follows that  $x_j$  vanishes identically. Then

$$0 \equiv \dot{x}_j(t) = \delta_j a_{jj-1}(t)x_{j-1}(t) + \delta_{j+1}a_{jj+1}(t)x_{j+1}(t)$$

and for  $t = t_m$ ,  $m \geq 1$ , both terms in the sum are nonnegative. By analyticity, both  $a_{jj-1}x_{j-1}$  and  $a_{jj+1}x_{j+1}$  must vanish identically and since at least one of  $a_{jj-1}$ ,  $a_{jj+1}$  is positive,  $x_{j-1}$  or  $x_{j+1}$  or both vanish identically. Suppose (1.2)(b) holds so that  $a_{ii+1} > 0$  for all  $i$ .

Then  $x_{j+1}$  vanishes identically. Since  $x_j$  and  $x_{j+1}$  vanish identically, the differential equation forces  $x_{j+2}, x_{j+3}, \dots, x_n, x_1, x_2, \dots, x_{j-1}$  to vanish identically. A similar argument applies if (1.2)(a) or (c) hold. Hence we obtain the contradiction  $x(t)$  vanishes identically.

#### REFERENCES

- [1] S. B. Angenent, *The Morse-Smale property for a semi-linear parabolic equation*, preprint.
- [2] G. Fusco and W. M. Oliva, *Jacobi matrices and transversality*, to appear in Proc. Royal Soc. Edin. Sec. A.
- [3] D. B. Henry, *Some infinite dimensional Morse-Smale systems defined by parabolic partial differential equations*, J. Differential Equations, **59** (1985), 165–205.
- [4] J. Mallet-Paret, *Morse decompositions for delay-differential equations*, J. Differential Equations, **72** (1988), 270–315.
- [5] J. Mallet-Paret and H. L. Smith, *The Poincaré-Bendixson Theorem for monotone cyclic feedback systems*, preprint.
- [6] H. Matano, *Non increase of the lapnumber of a solution for a one dimensional semi-linear parabolic equation*, J. Fac. Sc. Univ. Tokyo, **29** (1982), 401–441.
- [7] K. Nickel, *Gestattaussagen über Lösungen parabolischer differentialgleichungen*, J. Reine Angew. Math., **211** (1962), 78–94.
- [8] J. Smillie, *Competitive and cooperative tridiagonal systems of differential equations*, SIAM J. Math. Anal., **15** (1984), 530–534.

Received October 3, 1988, and in revised form October 3, 1989. Supported by NSF Grant DMS 8521605 and DMS 8722279.

ARIZONA STATE UNIVERSITY  
TEMPE, AZ 85287