Reprinted from JOURNAL OF DIFFERENTIAL EQUATIONS Vol. 5, No. 3, May 1969

Printed in Belgium

On the Global Limits and Oscillations of a System of Nonlinear Differential Equations Describing a Flow on a Probabilistic Network

STEPHEN GROSSBERG*

Department of Mathematics,
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received September 22, 1967; revised April 3, 1968

PART I

1. INTRODUCTION

This paper considers various aspects of the global limiting and oscillatory behavior of the following system of nonlinear differential equations.

$$\dot{x}_i(t) = -\alpha x_i(t) + \beta \sum_{k=1}^{3} x_k(t)\, y_{ki}(t) + I_i(t), \tag{1}$$

$$y_{jk}(t) = z_{jk}(t) \left[\sum_{m=1}^{3} z_{jm}(t) \right]^{-1}, \tag{2}$$

$$\dot{z}_{jk}(t) = -u z_{jk}(t) + \beta x_j(t)\, x_k(t), \qquad j \neq k, \qquad (3)^{(*)}$$

and

$$z_{jj}(t) = 0, \tag{4}$$

where $i, j, k = 1, 2, 3$ and β is a positive number.

The system (*) arises as a special case of a nonstationary prediction theory, or learning theory, whose goal is to discuss the prediction of individual events, in a fixed order, and at prescribed times ([1], [2], [3]). In this theory, (*) describes a machine $\mathscr{M}$ subjected to inputs $I = (I_1, I_2, I_3)$ by an experimenter $\mathscr{E}$ who records the outputs $X = (x_1, x_2, x_3)$ created thereby. $\mathscr{E}$ has only the inputs I and the outputs X at his disposal with which to describe (*), and in terms of these variables (*) takes the form

$$\dot{X}(t) = -\alpha X(t) + B(X_t)\, X(t) \qquad + I(t),$$

* The preparation of this work was supported in part by the National Science Foundation (GP 9003) and the Office of Naval Research (N00014-67-A-0204-0016).

where $B(X_t)$ is a matrix of nonlinear functionals of $X(w)$ evaluated at all past times $w \in [-\tau, t]$ with entries

$$B_{ij}(t) = \frac{z_{ji}(0) + \beta \int_0^t e^{uv} x_j(v)\, x_i(v)\, dv}{z_{ji}(0) + z_{jk}(0) + \beta \int_0^t e^{uv} x_j(v)[x_i(v) + x_k(v)]\, dv},$$

$$\{i, j, k\} = \{1, 2, 3\},$$

and

$$B_{ii}(t) = 0,$$

$t \geqslant 0$. The machine $\mathscr{M}$ therefore obeys a functional-differential equation, and $B(X_t)$ contains the "memory" of $\mathscr{M}$.

The mechanism of $\mathscr{M}$ can be described graph-theoretically in the following way. Let G be a directed probabilistic graph [4] with vertices $V = \{v_1, v_2, v_3\}$, directed edges $E = \{e_{ij} : i, j = 1, 2, 3\}$, and weight function $\varphi(e_{ij}) = \frac{1}{2}(1 - \delta_{ij})$. That is, each of the three vertices v_i of G sends out a directed edge e_{ij} with weight $\frac{1}{2}$ to every other vertex v_j, $j \neq i$; thus G is a *complete* 3-graph. On the other hand, no vertex v_i sends a directed edge e_{ii} to itself (that is, $\varphi(e_{ii}) = 0$); thus G is a complete 3-graph *without loops.*

$\mathscr{M}$ can be interpreted as a flow over G [5] by letting $x_k(t)$ denote the state of a process at vertex v_k at time t, and $y_{ki}(t)$ denote the state of a process at the arrowhead of e_{ki} at time t. At each time t, v_k emits a flow of magnitude $\beta x_k(t)$ which travels instantaneously along the edge e_{ki}, reaches the arrowhead of e_{ki}, and thereupon activates the process $y_{ki}(t)$. As a result, a quantity of size $\beta x_k(t)\, y_{ki}(t)$ is immediately released by the arrowhead and reaches v_i. The total input to v_i at time t from all vertices is the sum $\beta \sum_{k=1}^{3} x_k(t)\, y_{ki}(t)$ of these individual quantities, and (1) states that $x_i(t)$ responds at a rate proportional to this total input. The other terms in (1), namely $-\alpha x_i(t)$ and $I_i(t)$, describe, respectively, a spontaneous decay process at v_i and the input delivered by $\mathscr{E}$ to v_i at time t.

$y_{ki}(t)$ is the ratio of functions $z_{kj}(t)$, as in (2). $z_{kj}(t)$ *cross-correlates* the values $\beta x_k(t)$ and $x_j(t)$ in the sense of (3). $\beta x_k(t)$ is interpreted as the flow received at the arrowhead of e_{kj} from v_k at time t, and $z_{kj}(t)$ cross-correlates this flow with the value $x_j(t)$ of the contiguous vertex v_j.

Part I studies (*) mathematically for appropriate experimental inputs $I_i(t)$. Part II interprets these mathematical results in learning and prediction theoretic terms.

(*) can be interpreted prediction-theoretically only when all its initial data, other than for $z_{ii}(0) = 0$, are nonnegative. Moreover, all inputs I_i must be continuous and nonnegative in $[0, \infty)$. Whenever these conditions are satisfied, one readily proves [3] that (*) has a unique, continuously differentiable, and *nonnegative* solution in $(0, \infty)$. Moreover, if $z_{ij}(0) > 0$, then $z_{ij}(t) > 0$, and if $x_k(0) > 0$, then $x_k(t) > 0$ for $t \geqslant 0$. Throughout the following discussion these constraints on initial data and inputs will always be imposed.

2. Main Theorem

Our main theorem discusses the following situation. Suppose that (*) is given at time $t = 0$ with *uniform* initial data; that is, $x_i(0) = \gamma > 0$ and $z_{jk}(0) = \delta(1 - \delta_{jk})$, where $\delta > 0$, and $i, j, k = 1, 2, 3$. Let I_1, I_2 and I_3 be any continuous and nonnegative inputs which are positive only in a finite interval $[0, T]$. What is then the limiting and oscillatory behavior of (*) in (T, ∞)?

Since (*) is symmetric in the indices j and k and, by hypothesis, $z_{jk}(0) = z_{kj}(0)$, it is clear that $z_{jk}(t) = z_{kj}(t)$ for $t \geqslant 0$. It therefore suffices to consider the following situation, where we have replaced time $t = T$ by $t = 0$ for convenience of exposition:

(ω_1) the inputs I_i are identically zero in $[0, \infty)$, $i = 1, 2, 3$; that is, (*) is *input-free.*

(ω_2) the initial data satisfies $x_i(0) > 0$; $z_{jk}(0) = z_{kj}(0) > 0$, $j \neq k$; and $z_{jj}(0) = 0$, for all $i, j, k = 1, 2, 3$.

Concerning any such situation, the following theorem holds describing the ratios $y_{jk}(t)$ and $X_i(t) = x_i(t)[\sum_{k=1}^{3} x_k(t)]^{-1}$.

Theorem 1. *Let (*) be input-free and suppose β is positive. Then for any positive initial data satisfying $z_{jk}(0) = z_{kj}(0)$, $j, k = 1, 2, 3$, the following conclusions hold.*

(1) (*limiting behavior*). *All the ratios X_i and y_{jk} have limits $Q_i = \lim_{t\to\infty} X_i(t)$ and $P_{jk} = \lim_{t\to\infty} y_{jk}(t)$ as $t \to \infty$, which satisfy the system of equations*

$$\tfrac{1}{2} \geqslant Q_i = Q_j P_{ji} + Q_k P_{ki}, \qquad \{i, j, k\} = \{1, 2, 3\}.$$

In particular,

$$\lim_{t\to\infty} x_i(t)\, e^{(\alpha-\beta)t} = Q_i \sum_{k=1}^{3} x_k(0).$$

(2) (*oscillatory behavior*). *For all indices $\{i, j, k\} = \{1, 2, 3\}$, the functions $f_{ij} = x_i - x_j$, $g_{kij} = z_{ki} - z_{kj}$, $h_{jki} = x_j z_{ki} - x_i z_{kj}$, and $\dot{y}_{ij}$ change sign at most once. f_{ij} and g_{kij} do not change sign at all if $f_{ij}(0)\, g_{kij}(0) \geqslant 0$, while h_{jki}*

and $\dot{y}_{ij}$ *do not change sign at all if moreover* $f_{ij}(0)\, h_{jki}(0) \geqslant 0$. *Furthermore,* $f_{ij}(0)\, g_{kij}(0) > 0$ *implies* $f_{ij}(t)\, g_{kij}(t) > 0$ *for all* $t \geqslant 0$, *and if moreover* $f_{ij}(0)\, h_{jki}(0) > 0$ *then also* $f_{ij}(t)\, h_{jki}(t) > 0$ *for all* $t \geqslant 0$.

(3) (*uniqueness of limits*). *If the coefficients satisfy the inequality,* $\sigma \equiv u + 2(\beta - \alpha) > 0$, *then* $Q_i = \frac{1}{3}$ *and* $P_{jk} = \frac{1}{2}(1 - \delta_{jk})$, $i, j, k = 1, 2, 3$.

(4) (*nonuniqueness of limits*). *If* $\sigma < 0$, *then*

$$| P_{jk} - y_{jk}(0)| \leqslant 2 \log(1 + K_j/| \sigma |),$$

where

$$K_j = \frac{\beta(\sum_{k=1}^{3} x_k(0))^2}{4 \sum_{k=1}^{3} z_{jk}(0)} > 0.$$

The following remarks help to visualize the geometrical meaning of this theorem, say for the case $\sigma > 0$ in which uniqueness of limits holds.

(a) (2) shows, for example, that if $x_i(0) > x_j(0)$ and $z_{ki}(0) > z_{kj}(0)$, then $x_i(t) > x_j(t)$ and $z_{ki}(t) > z_{kj}(t)$ for all $t \geqslant 0$. Thus if a common ordering occurs in corresponding edges and vertices at $t = 0$, then this ordering propagates through time (i.e., it is a "geometrical" property of the graph). $z_{ki}(t) > z_{kj}(t)$ is equivalent to $y_{ki}(t) > \frac{1}{2}$. Since $\dot{y}_{ki}$ changes sign at most once, y_{ki} has a graph of either the form (A) or (B) given in Figure 1.

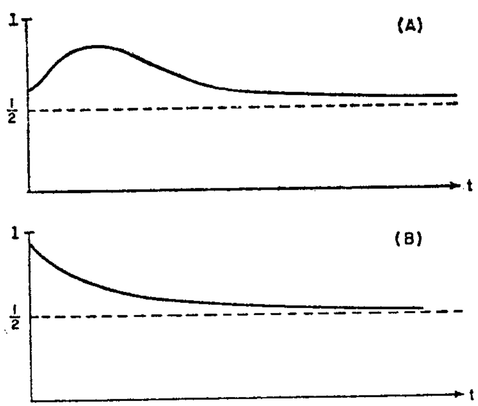

FIG. 1

(B) is guaranteed if, moreover, $x_j(0)\, z_{ki}(0) > x_i(0)\, z_{kj}(0)$. Thus after at most one start in the wrong direction due to an unfortunate choice of initial data, y_{ki} settles monotonically to its limit $\frac{1}{2}$ but does not reach this limit in finite time since $f_{ij}(0)\, g_{kij}(0) > 0$ implies $f_{ij}(t)\, g_{kij}(t) > 0$. This strong control on

(*)'s oscillations is important to our prediction theory, since (*) can be called upon at *any* time by the experimentalist controlling the inputs to reproduce in outputs the ordering of $y_{ki}(t)$ values induced by prior inputs.

(b) The hypothesis $\alpha > \beta$ is not essential to the proof. For example, σ is positive whenever $u > 2 \mid \alpha - \beta \mid$. Because $x = \sum_{k=1}^{3} x_k$ obeys the equation $\dot{x} = (\beta - \alpha)\, x$, (1) implies that $\lim_{t\to\infty} x_i(t) = 0$ if $\alpha > \beta$, that $\lim_{t\to\infty} x_i(t) = \infty$ if $\alpha < \beta$, and that $\sum_{k=1}^{3} x_k(t) =$ constant if $\alpha = \beta$. In all these cases, the ratios

$$X_i(t) = \frac{x_i(t)}{x_1(t) + x_2(t) + x_3(t)}$$

approach $\frac{1}{3}$ as $t \to \infty$ if $\sigma > 0$ and $\beta > 0$.

(c) The hypothesis $z_{ij}(0) = z_{ji}(0)$ is obviously equivalent to the hypothesis $z_{ij}(t) = z_{ji}(t)$, $t \geqslant 0$, even when (*) is not input-free. That is, the "reversibility" of the weights z_{ij} is also a "geometrical" property of the graph.

(d) The condition $\sigma > 0$ is not superfluous to proving, as stated in (3), that $Q_i = \frac{1}{3}$ and $P_{jk} = \frac{1}{2}(1 - \delta_{jk})$. Indeed, by (4), whenever $\sigma < 0$ and

$$| y_{jk}(0) - \tfrac{1}{2} | > 2 \log(1 + K_j/| \sigma |), \qquad \text{then} \quad P_{jk} \neq \tfrac{1}{2}.$$

Proof. Theorem 1 is proved in five major steps. In step (I), we make a change of variables that replaces (*) by a system of nonlinear integro-differential equations in terms of the new variables X_i and y_{jk}. In step (II), we make another change of variables that reveals a surprisingly linear substructure underlying (*). Working directly with these equations we can prove all the oscillatory properties stated in (2) as well as the existence of the limits P_{jk}.

Step (III) uses the existence of all P_{jk} to prove the existence of all limits Q_i by treating the equations found in (I) as a system of linear equations in the unknown variables X_i with the almost constant coefficients y_{jk}. In step (IV), we use the existence of all Q_i and P_{jk} to compute the possible values of these limits using various algebraic properties of (*). Then we apply some special facts found in (II) concerning the manner in which these limits are attained to show that the limits are unique if $\sigma > 0$. Step (V) proves part (4) of the theorem using an equation derived in Step (I).

The following lemma describes our first change of variables.

LEMMA 1. *Let (*) be input-free with arbitrary positive initial data. Then the probability distributions* X_i *and* y_{jk} *satisfy the integro-differential equations*

$$\dot{X}_i = \beta(-X_i + X_j y_{ji} + X_k y_{ki}), \qquad \{i, j, k\} = \{1, 2, 3\}, \tag{5}$$

and

$$\dot{y}_{jk} = G_j\left(\frac{X_k}{1 - X_j} - y_{jk}\right), \qquad j \neq k, \tag{6}$$

where

$$G_j = \frac{d}{dt}\log\left(\gamma_j + \int_0^t e^{\sigma v} X_j(1 - X_j)\,dv\right) \tag{7}$$

and

$$\gamma_j = \frac{\sum_{k=1}^3 z_{jk}(0)}{\beta(\sum_{k=1}^3 x_k(0))^2}.$$

Proof. Let $x = \sum_{k=1}^3 x_k$. Then $X_i = x_i/x$ and

$$\dot{X}_i = \frac{1}{x}\left(\dot{x}_i - x_i\frac{\dot{x}}{x}\right). \tag{8}$$

Summing over i in (1) readily shows, using positivity, that x satisfies the equation

$$\dot{x} = (\beta - \alpha)\,x. \tag{9}$$

Substituting (1) and (9) into (8), we find

$$\begin{aligned}\dot{X}_i &= \frac{1}{x}(-\alpha x_i + \beta x_j y_{ji} + \beta x_k y_{ki} - x_i(\beta - \alpha)),\\ &= \beta(-X_i + X_j y_{ji} + X_k y_{ki}),\end{aligned}$$

which is (5).

(6) has the following derivation. By (3),

$$\begin{aligned}\dot{z}_{jk} &= -uz_{jk} + \beta x_j x_k\\ &= -uz_{jk} + \beta x^2 X_j X_k.\end{aligned}$$

Letting $z^{(j)} = z_{ji} + z_{jk}$, $\{i, j, k\} = \{1, 2, 3\}$, we therefore find

$$\begin{aligned}\dot{z}^{(j)} &= -uz^{(j)} + \beta x^2 X_j(X_i + X_k)\\ &= -uz^{(j)} + \beta x^2 X_j(1 - X_j).\end{aligned} \tag{10}$$

In integral form this equation is

$$z^{(j)}(t) = e^{-ut}[z^{(j)}(0) + \beta\int_0^t e^{uv} x^2 X_j(1 - X_j)\,dv], \qquad t \geqslant 0.$$

Since by (9),

$$x(t) = x(0)\,e^{(\beta-\alpha)t}, \tag{11}$$

this equation becomes

$$z^{(j)}(t) = e^{-ut}[z^{(j)}(0) + \beta x^2(0)\int_0^t e^{\sigma v} X_j(1 - X_j)\,dv], \tag{12}$$

where $\sigma = u + 2(\beta - \alpha)$. Differentiating $y_{jk} = z_{jk}/z^{(j)}$, we find

$$\dot{y}_{jk} = \frac{1}{z^{(j)}}\left(\dot{z}_{jk} - z_{jk}\frac{\dot{z}^{(j)}}{z^{(j)}}\right). \tag{13}$$

Substituting (3) and (10) into (13) gives

$$\begin{aligned}\dot{y}_{jk} &= \frac{1}{z^{(j)}}\left[-uz_{jk} + \beta x^2 X_j X_k - z_{jk}\left(-u + \frac{\beta x^2 X_j(1 - X_j)}{z^{(j)}}\right)\right]\\ &= \frac{\beta x^2 X_j}{z^{(j)}}[X_k - y_{jk}(1 - X_j)]\\ &= \frac{\beta x^2 X_j(1 - X_j)}{z^{(j)}}\left[\frac{X_k}{1 - X_j} - y_{jk}\right].\end{aligned}$$

Letting $G_j = \beta x^2 X_j(1 - X_j)/z^{(j)}$, it remains only to show that G_j has the form given in (7) to complete the proof of (6). By (9),

$$G_j = \frac{\beta x^2(0)\,e^{2(\beta-\alpha)t} X_j(1 - X_j)}{z^{(j)}},$$

and by (12),

$$\begin{aligned}G_j &= \frac{\beta x^2(0)\,e^{2(\beta-\alpha)t} X_j(1 - X_j)}{e^{-ut}[z^{(j)}(0) + \beta x^2(0)\int_0^t e^{\sigma v} X_j(1 - X_j)\,dv]}\\ &= \frac{e^{\sigma t} X_j(1 - X_j)}{\gamma_j + \int_0^t e^{\sigma v} X_j(1 - X_j)\,dv},\end{aligned}$$

which indeed equals (7).

(II) To prove that all limits P_{jk} exist, we transform (5)–(7) still further to exhibit new unknown variables which obey equations that have a surprisingly linear form. We will find two matrix functions $\| X_{ij}(t)\|$ and $\| Y_{ij}(t)\|$ of new unknowns such that each triple (X_{ij}, Y_{ij}, y_{ij}), $i \neq j$, will have the properties of the triple (X, Y, y) in the following basic lemma.

LEMMA 2. *Let the real valued-functions $X(t)$, $Y(t)$, and $y(t)$ satisfy the following system of differential equations.*

$$\dot{X} = aX + bY \tag{14}$$

$$\dot{Y} = cX + dY \tag{15}$$

$$(Y - X)^{\cdot} = e(Y - X) + fY \tag{16}$$

$$\dot{y} = g(Y - X), \tag{17}$$

where the functions a, b, c, d, e, f, and g are continuous, and in addition the functions b, c, f, g are positive. Then the functions X, Y, Y − X, and $\dot{y}$ *change sign at most once. X and Y do not change sign at all if* $X(0)\,Y(0) \geq 0$, *while Y − X and* $\dot{y}$ *do not change sign at all if* $X(0)\,Y(0) \geq 0$ *and* $(Y - X)(0)\,Y(0) \geq 0$. *Moreover,* $X(0)\,Y(0) > 0$ *implies* $X(t)\,Y(t) > 0$ *for all* $t \geq 0$, *while* $X(0)\,Y(0) > 0$ *and* $(Y - X)(0)\,Y(0) > 0$ *imply* $(Y - X)(t)\,Y(t) > 0$ *for all* $t \geq 0$. *If in addition y is bounded, then* $\lim_{t\to\infty} y(t)$ *exists.*

The conclusions of the lemma concerning the sign changes of the functions X, Y, and $Y - X$ can be conveniently pictured in the (X, Y) plane as in Figure 2, where the arrows indicate the direction of the (X, Y) point through

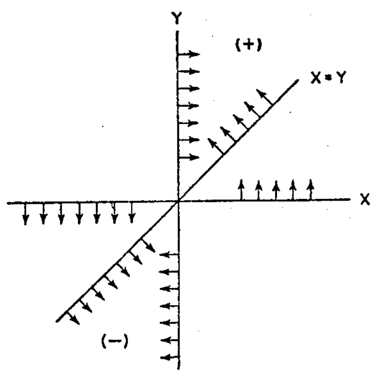

FIG. 2

time. If the (X, Y) point starts in the region $(+)$, then $y(t)$ is always monotone increasing, while if it starts in the region $(-)$, then $y(t)$ is always monotone decreasing.

Proof. The vector field pictured in Figure 2 is obviously nontrivial on the indicated critical lines. Also, by uniqueness, the trajectory cannot arrive at 0 from another point. The lemma's conclusions are therefore clear concerning X, Y, and $Y - X$. They can also easily be seen by integrating (14)–(16) using an exponential change of variable and then invoking nonnegativity of solutions; positivity is inessential.

Consider $\dot{y}$ in (17). Since g is positive, $\dot{y}$ changes sign at most once and not at all if $X(0)\,Y(0) \geq 0$ and $(Y - X)(0)\,Y(0) \geq 0$. We can therefore find a T such that $t \geq T$ implies $y(t)$ is a monotonic function. Thus if

$y(t)$ is also bounded, then $\lim_{t\to\infty} y(t)$ exists. This completes the proof of Lemma 2.

Lemma 2 will prove the existence of all limits P_{jk} if we can find matrix functions $\| X_{jk}(t) \|$ and $\| Y_{jk}(t) \|$ with the properties of (14)–(17). We define these matrix functions in terms of the probability distributions

$$y_{jk} = \frac{z_{jk}}{z_{jk} + z_{ji}}$$

and the analogously defined probability distributions

$$x_{jk} = \frac{x_k}{x_k + x_i}, \qquad \text{where} \qquad \{i, j, k\} = \{1, 2, 3\}.$$

This we do by setting $Y_{jk} = \frac{1}{2} - y_{jk}$ and $X_{jk} = \frac{1}{2} - x_{jk}$. Supposing that these functions satisfy (14)–(17), they also suffice to prove part (2) of the theorem, because of the following identities:

$$X_{jk} = \frac{f_{ik}}{2(x_i + x_k)},$$

$$Y_{jk} = \frac{g_{jik}}{2(z_{ji} + z_{jk})}$$

and

$$Y_{jk} - X_{jk} = \frac{h_{kji}}{(x_i + x_k)(z_{ji} + z_{jk})}$$

where $x_i + x_k > 0$ and $z_{ji} + z_{jk} > 0$. We now proceed to show that the matrix functions $\| X_{jk}(t) \|$ and $\| Y_{jk}(t) \|$ defined above do indeed have the desired properties. In the following discussion, only the triple (X_{21}, Y_{21}, y_{21}) will be considered. Our conclusions will carry over immediately to all triples (X_{ij}, Y_{ij}, y_{ij}) with $i \neq j$ by simply permuting indices.

(IIa) We seek an equation like (14) for X_{21} and Y_{21}. This equation is

$$\dot{X}_{21} = -A_{21}X_{21} + B_{21}Y_{21}, \tag{18}$$

where

$$A_{21} = -\frac{\dot{X}_2}{1 - X_2} + \beta\left(1 + \frac{y_{13} + y_{31}}{2}\right)$$

and

$$B_{21} = \beta\left[\frac{X_2}{1 - X_2} + \frac{y_{13}}{2}\left(\frac{z^{(2)}}{z^{(3)}}\right)\right].$$

B_{21} is obviously positive, as Lemma 2 requires. (18) is derived as follows.

Letting $i = 3$ and then 1 in (5) and subtracting the two equations gives

$$(X_3 - X_1)^{\cdot} = \beta[-(X_3 - X_1) + X_2(y_{23} - y_{21}) + X_1 y_{13} - X_3 y_{31}]$$
$$= \beta[-(X_3 - X_1) + X_2(y_{23} - y_{21})$$
$$+ (X_1 - X_3) y_{13} + X_3(y_{13} - y_{31})],$$

or

$$(X_3 - X_1)^{\cdot} = \beta[-(1 + y_{13})(X_3 - X_1) + X_2(y_{23} - y_{21})$$
$$+ X_3(y_{13} - y_{31})]. \tag{19}$$

The left hand side of (19) is an antisymmetric function of the indices 3 and 1. We now seek an expression for the right hand side of (19) which is also antisymmetric in these indices. Permuting the indices 3 and 1 in (19) gives

$$(X_1 - X_3)^{\cdot} = \beta[-(1 + y_{31})(X_1 - X_3) + X_2(y_{21} - y_{23}) + X_1(y_{31} - y_{13})]. \tag{20}$$

Subtract (20) from (19) and divide the resulting equation by 2. Then

$$(X_3 - X_1)^{\cdot} = -H_{31}(X_3 - X_1) + \beta X_2(y_{23} - y_{21}) + \frac{\beta(1 - X_2)}{2}(y_{13} - y_{31}), \tag{21}$$

where

$$H_{31} = \beta\left(1 + \frac{y_{13} + y_{31}}{2}\right) = H_{13}.$$

The right hand side of (21) is clearly antisymmetric in the indices 1 and 3.

Using (21) we seek an equation for the derivative of X_{21} in terms of X_{21} and Y_{21}. Since

$$X_{21} = \tfrac{1}{2} - x_{21}$$
$$= \tfrac{1}{2} - x_1/(x_1 + x_3) = (x_3 - x_1)/2(x_1 + x_3)$$
$$= (x_3 - x_1)/2(x - x_2) = (X_3 - X_1)/2(1 - X_2), \tag{22}$$

we differentiate X_{21} to find

$$\dot{X}_{21} = [(X_3 - X_1)/2(1 - X_2)]^{\cdot}$$
$$= (X_3 - X_1)\frac{\dot{X}_2}{2(1 - X_2)^2} + \frac{1}{2(1 - X_2)}(X_3 - X_1)^{\cdot},$$

which by (21) and (22) can be written as

$$\dot{X}_{21} = \frac{\dot{X}_2}{1 - X_2} X_{21} + \frac{1}{2(1 - X_2)}\Big[-H_{31}(X_3 - X_1) + \beta X_2(y_{23} - y_{21})$$
$$+ \frac{\beta(1 - X_2)}{2}(y_{13} - y_{31})\Big].$$

By another application of (22) and a rearrangement of terms, we arrive at

$$\dot{X}_{21} = -\left(H_{31} - \frac{\dot{X}_2}{1 - X_2}\right) X_{21} + \frac{\beta X_2}{2(1 - X_2)}(y_{23} - y_{21}) + \frac{\beta}{4}(y_{13} - y_{31}). \tag{23}$$

Consider the term $[\beta X_2/2(1 - X_2)](y_{23} - y_{21})$ in (23). Since $y_{23} = 1 - y_{21}$ and $Y_{21} = \frac{1}{2} - y_{21}$,

$$\frac{\beta X_2}{2(1 - X_2)}(y_{23} - y_{21}) = \frac{\beta X_2}{1 - X_2} Y_{21}.$$

Letting $A_{21} = H_{31} - \dot{X}_2/(1 - X_2)$, we can therefore rewrite (23) as

$$\dot{X}_{21} = -A_{21}X_{21} + \frac{\beta X_2}{1 - X_2} Y_{21} + \frac{\beta}{4}(y_{13} - y_{31}). \tag{24}$$

Consider the term $(\beta/4)(y_{13} - y_{31})$ of (24) in the light of the hypothesis $z_{ij} = z_{ji}$. Since

$$y_{13} - y_{31} = \frac{z_{13}}{z_{12} + z_{13}} - \frac{z_{31}}{z_{31} + z_{32}}$$
$$= \frac{z_{13}}{z_{21} + z_{13}} - \frac{z_{13}}{z_{13} + z_{23}} = \frac{z_{13}(z_{23} - z_{21})}{(z_{21} + z_{13})(z_{13} + z_{23})}$$
$$= \frac{z_{13}z^{(2)}}{z^{(1)}z^{(3)}}(y_{23} - y_{21}),$$

this term equals

$$\frac{\beta}{4}(y_{13} - y_{31}) = \frac{\beta}{2} y_{13}\frac{z^{(2)}}{z^{(3)}} Y_{21}. \tag{25}$$

Substituting (25) into (24) gives

$$\dot{X}_{21} = -A_{21}X_{21} + B_{21}Y_{21} \tag{18}$$

where

$$B_{21} = \beta\left[\frac{X_2}{1 - X_2} + \frac{y_{13}}{2}\left(\frac{z^{(2)}}{z^{(3)}}\right)\right].$$

B_{21} is clearly positive.

(IIb) We now seek an equation like (15) for Y_{21} and X_{21}. This equation is

$$\dot{Y}_{21} = -G_2 Y_{21} + G_2 X_{21}, \tag{26}$$

where G_2 is given in (7) and is thus positive. (26) is derived as follows. By (6),

$$\dot{y}_{21} = G_2\left(\frac{X_1}{1-X_2} - y_{21}\right).$$

Since $Y_{21} = \frac{1}{2} - y_{21}$,

$$\dot{Y}_{21} = G_2\left(-\frac{X_1}{1-X_2} + \frac{1}{2} - Y_{21}\right),$$

and since

$$X_{21} = \tfrac{1}{2} - \frac{x_1}{x_1 + x_3}$$
$$= \tfrac{1}{2} - \frac{x_1}{x - x_2} = \tfrac{1}{2} - \frac{X_1}{1 - X_2},$$
$$\dot{Y}_{21} = G_2(X_{21} - Y_{21}). \tag{26}$$

(IIc) We now seek an equation like (16) for $Y_{21} - X_{21}$ and Y_{21}. This equation is

$$(Y_{21} - X_{21})^{\cdot} = -E_{21}(Y_{21} - X_{21}) + F_{21} Y_{21}, \tag{27}$$

where

$$E_{21} = G_2 + A_{21} + \beta(y_{31} - y_{13})\, Y_{21}$$

and

$$F_{21} = \beta[y_{31}(y_{13} + y_{23}) + y_{21} y_{13}].$$

F_{21} is obviously positive, as Lemma 2 requires.

The strategy for deriving this equation is simple. We compute equations for $\dot{X}_{21}$ and $\dot{Y}_{21}$ by factoring out as many expressions $Y_{21} - X_{21}$ as possible. Then we combine these equations by subtraction. Since the equation for $\dot{Y}_{21}$, namely (26), is already in the desired form, we need only compute an equation for $\dot{X}_{21}$. We already have equation (18) for $\dot{X}_{21}$, and the new equation is obtained by merely rearranging terms in (18). By (18),

$$\dot{X}_{21} = -A_{21} X_{21} + B_{21} Y_{21}.$$

Using the normalization conditions $X_1 + X_2 + X_3 = 1$ and $y_{ij} + y_{ik} = 1$, $\{i, j, k\} = \{1, 2, 3\}$, we will be able to rewrite (18) as

$$\dot{X}_{21} = -D_{21}(X_{21} - Y_{21}) - F_{21} Y_{21} \tag{28}$$

where F_{21} is *positive*. Subtracting (28) from (26) and setting $E_{21} = G_2 + D_{21}$ then gives

$$(Y_{21} - X_{21})^{\cdot} = -E_{21}(Y_{21} - X_{21}) + F_{21} Y_{21} \tag{29}$$

It therefore remains only to show that (28) can be found.

Letting $K_{21} = B_{21} - A_{21}$, (18) can be written as

$$\dot{X}_{21} = -A_{21}(X_{21} - Y_{21}) + K_{21} Y_{21}. \tag{30}$$

Consider K_{21}. Since

$$B_{21} = \beta\left[\frac{X_2}{1 - X_2} + \frac{y_{13}}{2}\left(\frac{z^{(2)}}{z^{(3)}}\right)\right]$$

and

$$A_{21} = \frac{-\dot{X}_2}{1 - X_2} + \beta\left(1 + \frac{y_{13} + y_{31}}{2}\right),$$

where

$$\dot{X}_2 = \beta(-X_2 + X_1 y_{12} + X_3 y_{32}),$$

it follows readily that

$$K_{21} = \beta\left(J_{21} + \frac{y_{13}}{2}\left(\frac{z^{(2)}}{z^{(3)}}\right)\right), \tag{31}$$

where

$$J_{21} = \frac{X_1 y_{12} + X_3 y_{32}}{1 - X_2} - \left(1 + \frac{y_{13} + y_{31}}{2}\right).$$

We apply the normalization conditions to J_{21} and rearrange terms.

$$J_{21} = \frac{X_1(1 - y_{13}) + X_3(1 - y_{31})}{1 - X_2} - \left(1 + \frac{y_{13} + y_{31}}{2}\right)$$
$$= \frac{X_1 + X_3}{1 - X_2} - 1 - y_{13}\left(\tfrac{1}{2} + \frac{X_1}{1 - X_2}\right) - y_{31}\left(\tfrac{1}{2} + \frac{X_3}{1 - X_2}\right)$$
$$= -y_{13}\left(\tfrac{1}{2} + \frac{X_1}{1 - X_2}\right) - y_{31}\left(\tfrac{1}{2} + \frac{X_3}{1 - X_2}\right).$$

We now evaluate the terms $\frac{1}{2} + X_1/(1 - X_2)$ and $\frac{1}{2} + X_3/(1 - X_2)$ in this expression. Since $X_{21} = \frac{1}{2} - X_1/(1 - X_2)$ and $Y_{21} = \frac{1}{2} - y_{21}$, we find

$$\tfrac{1}{2} + \frac{X_1}{1 - X_2} = \tfrac{1}{2} + y_{21} + \left(\frac{X_1}{1 - X_2} - y_{21}\right)$$
$$= \tfrac{1}{2} + y_{21} + (Y_{21} - X_{21}),$$

and

$$\tfrac{1}{2} + \frac{X_3}{1 - X_2} = \tfrac{1}{2} + \frac{1 - X_2 - X_1}{1 - X_2}$$
$$= \tfrac{3}{2} - \frac{X_1}{1 - X_2} = \tfrac{3}{2} - y_{21} + (X_{21} - Y_{21}).$$

Thus

$$J_{21} = (y_{13} - y_{31})(X_{21} - Y_{21}) - y_{13}(\tfrac{1}{2} + y_{21}) - y_{31}(\tfrac{3}{2} - y_{21}). \tag{32}$$

Substituting (31) and (32) into (30) and rearranging terms gives

$$\dot{X}_{21} = -\tilde{D}_{21}(X_{21} - Y_{21}) - \tilde{F}_{21} Y_{21}, \tag{33}$$

where

$$\tilde{D}_{21} = A_{21} + \beta(y_{31} - y_{13})\, Y_{21}$$

and

$$\tilde{F}_{21} = \beta \left[y_{13}(\tfrac{1}{2} + y_{21}) + y_{31}(\tfrac{3}{2} - y_{21}) - \frac{y_{13}}{2}\left(\frac{z^{(2)}}{z^{(3)}}\right)\right].$$

We now show that $\tilde{F}_{21}$ is positive and thus that the identifications $\tilde{D}_{21} = D_{21}$ and $\tilde{F}_{21} = F_{21}$ can be made to give (28). For this computation, both the normalization conditions and the constraints $z_{ij} = z_{ji}$ are needed.

The terms in $(1/\beta)\, \tilde{F}_{21}$ can be rearranged to read

$$\frac{1}{\beta}\tilde{F}_{21} = K_{21} + y_{21}y_{13} + y_{31}(\tfrac{3}{2} - y_{21}),$$

where

$$K_{21} = \frac{y_{13}}{2}\left(1 - \frac{z^{(2)}}{z^{(3)}}\right)$$
$$= \frac{y_{13}}{2}\left(\frac{z_{31} - z_{21}}{z^{(3)}}\right) = \frac{y_{13}}{2}\left(y_{31} - \frac{z_{21}}{z^{(3)}}\right)$$
$$= \tfrac{1}{2}\left(y_{13}y_{31} - \frac{z_{13}z_{21}}{z^{(1)}z^{(3)}}\right) = \tfrac{1}{2}\left(y_{13}y_{31} - \frac{z_{12}z_{31}}{z^{(1)}z^{(3)}}\right)$$
$$= \tfrac{1}{2} y_{31}(y_{13} - y_{12}) = y_{31}(\tfrac{1}{2} - y_{12}).$$

Thus

$$\frac{1}{\beta}\tilde{F}_{21} = y_{31}(\tfrac{1}{2} - y_{12}) + y_{21}y_{13} + y_{31}(\tfrac{3}{2} - y_{21})$$
$$= y_{31}(2 - y_{12} - y_{21}) + y_{21}y_{13}$$
$$= y_{31}(y_{13} + y_{23}) + y_{21}y_{13}.$$

$\tilde{F}_{21}$ is manifestly positive when it is written in this way, and we have therefore derived an equation of the form given in (28). This completes the derivation of an equation such as (16), for which

$$E_{21} = G_{21} + D_{21} = G_2 + \tilde{D}_{21} = G_2 + A_{21} + \beta(y_{31} - y_{13})\, Y_{21}$$

and

$$F_{21} = \tilde{F}_{21} = \beta[y_{31}(y_{13} + y_{23}) + y_{21}y_{13}].$$

(IId) It remains only to produce an equation like (17) for $\dot{y}_{21}$ in terms of $Y_{21} - X_{21}$. This equation follows readily from (26). Since $y_{21} = \tfrac{1}{2} - Y_{21}$,

$$\dot{y}_{21} = G_2(Y_{21} - X_{21}), \tag{34}$$

where $G_2 > 0$. Equations (18), (26), (27), and (34) correspond to equations (14)–(17). By merely permuting indices, we can in the same way derive equations for all indices $\{i, j, k\} = \{1, 2, 3\}$. Thus all the assertions of (2) in Theorem 1 are immediate. Moreover, since y_{jk} is a bounded function, $P_{jk} = \lim_{t\to\infty} y_{jk}(t)$ exists for all $j \neq k$.

(III) We now use the existence of all P_{jk} to prove the existence of all $Q_i = \lim_{t\to\infty} X_i(t)$ by treating (5) as a *linear* system of equations in the unknown variables X_i with the almost constant coefficients y_{jk}. The first step in this treatment is to reduce (5) from a 3-by-3 system of equations to a 2-by-2 system by utilizing the normalization condition $X_1 + X_2 + X_3 = 1$. After doing this, the proof becomes straightforward.

By (5),

$$\dot{X}_1 = \beta(-X_1 + X_2 y_{21} + (1 - X_1 - X_2)\, y_{31})$$
$$= -\beta(1 + y_{31})\, X_1 + (y_{21} - y_{31})\, X_2 + y_{31}.$$

Letting $m_{ij} = y_{ij} - P_{ij}$, this becomes

$$\dot{X}_1 = -\beta(1 + P_{31} + m_{31})\, X_1 + \beta(P_{21} - P_{31} + m_{21} - m_{31})\, X_2 + \beta y_{31}.$$

Similarly,

$$\dot{X}_2 = -\beta(1 + P_{32} + m_{32})\, X_2 + \beta(P_{12} - P_{32} + m_{12} - m_{32})\, X_1 + \beta y_{32}.$$

In terms of the vector $X = \binom{X_1}{X_2}$, these equations can be written in matrix form as

$$\dot{X} = (A + B(t))\, X + C(t), \tag{35}$$

where

$$A = \beta \begin{pmatrix} -(1 + P_{31}) & P_{21} - P_{31} \\ P_{12} - P_{32} & -(1 + P_{32}) \end{pmatrix}$$

$$B(t) = \beta \begin{pmatrix} -m_{31}(t) & m_{21}(t) - m_{31}(t) \\ m_{12}(t) - m_{32}(t) & -m_{32}(t) \end{pmatrix},$$

and

$$C(t) = \beta \begin{pmatrix} y_{31}(t) \\ y_{32}(t) \end{pmatrix}.$$

Since $\lim_{t\to\infty} y_{ij}(t) = P_{ij}$, we conclude that $\lim_{t\to\infty} B(t) = 0$ and

$$\lim_{t\to\infty} C(t) = \beta \begin{pmatrix} P_{31} \\ P_{32} \end{pmatrix}.$$

Since $X(t)$ is bounded, $\lim_{t\to\infty} B(t)\, X(t) = 0$ and (35) can be written in the form

$$\dot{X} = AX + f(t) \tag{36}$$

where

$$f(t) = B(t)\, X(t) + C(t)$$

and

$$\lim_{t\to\infty} f(t) = \beta \begin{pmatrix} P_{31} \\ P_{32} \end{pmatrix}.$$

We now show that the eigenvalues of the matrix A have negative real parts. From this and the existence of the limit $\lim_{t\to\infty} f(t)$, it follows from (36) by elementary arguments [6] that the limits $Q_i = \lim_{t\to\infty} X_i(t)$ exist for all $i = 1, 2, 3$. To show that the eigenvalues λ_1 and λ_2 of A have negative real parts, we need only the elementary formulas $\lambda_1\lambda_2 = \det A$ and $\lambda_1 + \lambda_2 = \operatorname{tr} A$. $\det A$ is positive since

$$\begin{aligned} \det A &= \beta^2[(1 + P_{31})(1 + P_{32}) - (P_{21} - P_{31})(P_{12} - P_{32})] \\ &= \beta^2[1 + P_{31} + P_{32} + P_{21}(P_{32} - P_{12}) + P_{31}P_{12}] \\ &= \beta^2[2 + P_{21}(P_{32} - P_{12}) + P_{31}P_{12}] \\ &\geqslant \beta^2 > 0. \end{aligned}$$

Tr A is negative since

$$\operatorname{tr} A = -\beta(1 + P_{31} + 1 + P_{32}) = -3\beta < 0.$$

The eigenvalues λ_1 and λ_2 are either conjugate complex numbers or they are both real. In the former case, $\operatorname{Re}\lambda_1 = \operatorname{Re}\lambda_2 = \frac{1}{2}\operatorname{tr} A < 0$. In the latter case, λ_1 and λ_2 have the same sign since $\lambda_1\lambda_2 > 0$. This sign is negative since $\lambda_1 + \lambda_2 < 0$. This completes the proof of the existence of the limits Q_i.

(IV) Having established the existence of all limits Q_i and P_{jk}, we now show that $Q_i = \frac{1}{3}$ and $P_{jk} = \frac{1}{2}(1 - \delta_{jk})$ if $\sigma = u + 2(\beta - \alpha) > 0$. The first step in this proof is to show that the following relations hold.

$$Q_i \leqslant \tfrac{1}{2}, \tag{37}$$

$$Q_i = Q_jP_{ji} + Q_kP_{ki}, \tag{38}$$

and

$$P_{jk} = \frac{Q_k}{1 - Q_j}, \qquad j \neq k, \qquad \text{if} \quad Q_j > 0. \tag{39}$$

(37) and (38) will be seen not to depend on the hypothesis $\sigma > 0$.

The inequality (37) follows from the inequality

$$\begin{aligned} \dot{X}_i &= \beta(-X_i + X_jy_{ji} + X_ky_{ki}) \\ &\leqslant \beta(-X_i + X_j + X_k) \\ &= 2\beta(\tfrac{1}{2} - X_i). \end{aligned}$$

since then $\dot{X}_i < 0$ whenever $X_i > \frac{1}{2}$.

The equations (38) follow directly by letting $t \to \infty$ in (5) and using the existence of all limits Q_i and P_{jk} to conclude that $\lim_{t\to\infty} \dot{X}_i(t)$ exists and equals $\beta(-Q_i + Q_jP_{ji} + Q_kP_{ki})$. Since X_i is bounded, $\lim_{t\to\infty} \dot{X}_i(t) = 0$. (38) is now immediate since $\beta > 0$.

Equation (39) can be derived as follows if $\sigma > 0$. By (37) and the hypothesis $Q_j > 0$, we find $0 > Q_j(1 - Q_j)$. Thus by (7),

$$\begin{aligned} \lim_{t\to\infty} G_j(t) &= \frac{\lim_{t\to\infty} X_j(1 - X_j)}{\lim_{t\to\infty}(\gamma_j e^{-\sigma t} + \int_0^t e^{-\sigma(t-v)} X_j(1 - X_j)\, dv} \\ &= Q_j(1 - Q_j)/Q_j(1 - Q_j)/\sigma \\ &= \sigma > 0. \end{aligned}$$

Letting $t \to \infty$ in (6) therefore shows that $\lim_{t\to\infty} \dot{y}_{jk}(t)$ exists and equals

$$\sigma\left(\frac{Q_k}{1 - Q_j} - P_{jk}\right).$$

Since y_{jk} is bounded, $\lim_{t\to\infty} \dot{y}_{jk}(t) = 0$. Since $\sigma > 0$, (39) follows immediately.

Using (37)–(39), we now show that the possible values of Q_i and P_{jk} can be grouped into two cases if $\sigma > 0$.

Case 1. If all $Q_i > 0$, then $Q_i = \frac{1}{3}$ and $P_{jk} = \frac{1}{2}(1 - \delta_{jk})$. This is proved by substituting (39) into (38). Then

$$Q_i = Q_i\left(\frac{Q_j}{1 - Q_j} + \frac{Q_k}{1 - Q_k}\right).$$

Since $Q_i > 0$,

$$1 = \frac{Q_j}{1 - Q_j} + \frac{Q_k}{1 - Q_k}.$$

This is true for all $j \neq k$. Thus

$$\frac{Q_1}{1-Q_1} = \frac{Q_2}{1-Q_2} = \frac{Q_3}{1-Q_3},$$

which leads immediately to the identities $Q_1 = Q_2 = Q_3$. Since moreover $\sum_{k=1}^{3} Q_k = 1$, all $Q_i = \frac{1}{3}$. (39) now implies that

$$P_{jk} = \frac{\frac{1}{3}}{1-\frac{1}{3}} = \frac{1}{2}$$

if $j \neq k$.

Case 2. If one $Q_i = 0$, then $Q_j = Q_k = \frac{1}{2}$, $P_{jk} = P_{kj} = 1$, and $P_{ji} = P_{ki} = 0$. (No more than one $Q_i = 0$ since then some $Q_k = 1$ which contradicts (37)). By (38),

$$0 = Q_i = Q_j P_{ji} + Q_k P_{ki}.$$

Since all limits are nonnegative

$$0 = Q_j P_{ji} = Q_k P_{ki}.$$

Since $Q_j > 0$ and $Q_k > 0$, $0 = P_{ji} = P_{ki}$ and $1 = P_{jk} = P_{kj}$. By (38),

$$Q_j = Q_i P_{ij} + Q_k P_{kj} = Q_k P_{kj} = Q_k.$$

Since $1 = \sum_{m=1}^{3} Q_m = Q_j + Q_k$, we conclude that $Q_j = Q_k = \frac{1}{2}$.

To complete the proof that $Q_i = \frac{1}{3}$ and $P_{jk} = \frac{1}{2}(1 - \delta_{jk})$ if $\sigma > 0$, we now show that Case 2 cannot arise by employing special facts from (II) concerning the manner in which the functions y_{ji} and y_{ki} approach their limits. Suppose Case 2 holds. By (II) there exists a $T > 0$ such that y_{ji} and y_{ki} are monotonic functions for $t \geqslant T$. y_{ji} and y_{ki} are also nonnegative functions and their limits P_{ji} and P_{ki} are zero by hypothesis. Thus y_{ji} and y_{ki} are monotonically decreasing functions for $t \geqslant T$; that is $\dot{y}_{ji} \leqslant 0$ and $\dot{y}_{ki} \leqslant 0$ for $t \geqslant T$. By (6) this means that $y_{ji} \geqslant X_i/(1 - X_j)$ and $y_{ki} \geqslant X_i/(1 - X_k)$ for $t \geqslant T$. We use these inequalities to estimate $\dot{X}_i$ for large t. In particular, we will be able to find a $T_1 \geqslant T$ such that $\dot{X}_i \geqslant (\beta/2) X_i > 0$ for $t \geqslant T_1$. Since X_i is positive, X_i can therefore never achieve a zero limit. But $Q_i = 0$, by hypothesis. This contradiction shows that only Case 1 can arise.

To establish the desired estimate for X_i, consider (5) for $t \geqslant T$. Then

$$\begin{aligned}\dot{X}_i &= \beta(-X_i + X_j y_{ji} + X_k y_{ki}) \\ &= \beta \left[\left(\frac{X_j}{1-X_j} + \frac{X_k}{1-X_k} - 1 \right) X_i + X_j \left(y_{ji} - \frac{X_i}{1-X_j} \right) \right. \\ &\quad \left. + X_k \left(y_{ki} - \frac{X_i}{1-X_k} \right) \right] \\ &\geqslant \beta \left(\frac{X_j}{1-X_j} + \frac{X_k}{1-X_k} - 1 \right) X_i .\end{aligned}$$

By the hypothesis $Q_j = Q_k = \frac{1}{2}$, it follows that

$$\lim_{t\to\infty} \frac{X_j}{1-X_j} = \lim_{t\to\infty} \frac{X_k}{1-X_k} = 1.$$

Thus there surely exists a T_1 such that

$$\dot{X}_i \geqslant \frac{\beta}{2} X_i \qquad \text{for} \quad t \geqslant T_1 .$$

We have hereby proved parts (1), (2), and (3) of the theorem.

(V) Part (4) of the theorem follows from simple estimates on (6). By (6),

$$\begin{aligned}| y_{jk}(t) - y_{jk}(0)| &\leqslant \int_0^t | \dot{y}_{jk} | \, dv \\ &\leqslant 2 \int_0^t G_j \, dv \\ &= 2 \int_0^t \frac{d}{dv} \log \left(\gamma_j + \int_0^v e^{\sigma w} X_j (1 - X_j) \, dw \right) dv \\ &= 2 \log \left(1 + \frac{1}{\gamma_j} \int_0^t e^{-|\sigma| v} X_j (1 - X_j) \, dv \right) \\ &\leqslant 2 \log \left(1 + \frac{1}{4\gamma_j} \int_0^t e^{-|\sigma| v} \, dv \right) \\ &\leqslant 2 \log \left(1 + \frac{1}{4\gamma_j |\sigma|} \right).\end{aligned}$$

Letting $t \to \infty$ shows that

$$| P_{jk} - y_{jk}(0)| \leqslant 2 \log \left(1 + \frac{K_j}{|\sigma|} \right), \tag{40}$$

where $K_j = 1/4\gamma_j$. This completes the proof of Theorem 1.

Note that other estimates such as (40) can be derived from the equation

$$y_{jk}(t) = \frac{y_{jk}(0) + \frac{1}{\gamma_j}\int_0^t e^{\sigma v} X_j X_k \, dv}{1 + \frac{1}{\gamma_j}\int_0^t e^{\sigma v} X_j(1 - X_j) \, dv}$$

which follows from (6) and (7); for example, if $\sigma < 0$ then

$$P_{jk} - y_{jk}(0) \leqslant \frac{1}{4\gamma_j \,|\, \sigma \,|}.$$

3. The Probabilities $x_{ij}(t)$

Theorem 1 shows that $y_{ij}(t)$ approaches its limit $\frac{1}{2}$ monotonically when $\sigma > 0$ except possibly for one peak in its graph. We now discuss the oscillations of the analogous probabilities $x_{ij}(t)$ as $t \to \infty$, since these probabilities describe the relative size of the outputs $x_j(t)$ and $x_k(t)$ as $t \to \infty$. We consider only the time interval $[T_2, \infty)$ after all functions X_{jk}, Y_{jk}, and $Y_{jk} - X_{jk}$ have made their single sign change. We will show that there are only two possible avenues of approach for x_{ij} to its limit $\frac{1}{2}$ in this time interval.

For example, suppose that $y_{ij}(T_2) > \frac{1}{2}$. Then one of the two graphs in Figure 3 holds.

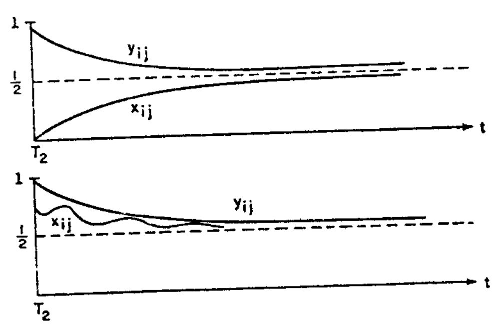

Fig. 3

That is, either $x_{ij}(T_2) < \frac{1}{2}$ and x_{ij} increases *monotonically* to $\frac{1}{2}$, or $x_{ij}(T_2) \geqslant \frac{1}{2}$ and the oscillations of x_{ij}, if any, are squeezed between y_{ij} and $\frac{1}{2}$ as $y_{ij} \to \frac{1}{2}$.

If $y_{ij}(T_2) < \frac{1}{2}$, the symmetric situation prevails. We prove these alternatives in the following corollary.

Corollary 1. *There exists a T_2 such that exactly one of the following alternatives holds for all $t \geqslant T_2$ if $\sigma > 0$:*

1) $y_{ij} \geqslant x_{ij} \geqslant \frac{1}{2}$
2) $\frac{1}{2} \geqslant x_{ij} \geqslant y_{ij}$
3) $y_{ij} \geqslant \frac{1}{2} \geqslant x_{ij}$ *and x_{ij} is monotone increasing.*
4) $x_{ij} \geqslant \frac{1}{2} \geqslant y_{ij}$ *and x_{ij} is monotone decreasing.*

Proof. By (I), X_{ij}, Y_{ij}, and $Y_{ij} - X_{ij}$ do not change sign for $t \geqslant T$ if T is chosen sufficiently large. For example, we can have $X_{ij} \geqslant 0$ for $t \geqslant T$. This is the same as $x_{ij} \leqslant \frac{1}{2}$ for $t \geqslant T$. Similarly $Y_{ij} \geqslant 0$ for $t \geqslant T$ gives $y_{ij} \leqslant \frac{1}{2}$ for $t \geqslant T$. And $Y_{ij} - X_{ij} \geqslant 0$ for $t \geqslant T$ gives $x_{ij} \geqslant y_{ij}$ for $t \geqslant T$. By examining all possible inequalities in this way, it follows that the relative magnitudes of x_{ij}, y_{ij}, and $\frac{1}{2}$ are fixed for $t \geqslant T$.

We also know that y_{ij} is a monotonic function for $t \geqslant T$ and that $P_{ij} = \frac{1}{2}$. Thus if (say) $y_{ij}(T) \geqslant \frac{1}{2}$, then y_{ij} decreases monotonically to $\frac{1}{2}$ for $t \geqslant T$. This means $\dot{y}_{ij} \leqslant 0$ for $t \geqslant T$. Since $\dot{y}_{ij} = G_i(x_{ij} - y_{ij})$, $y_{ij} \geqslant x_{ij}$ for $t \geqslant T$. This situation admits two possible subcases. Either $y_{ij} \geqslant x_{ij} \geqslant \frac{1}{2}$ for $t \geqslant T$, or $y_{ij} \geqslant \frac{1}{2} \geqslant x_{ij}$ for $t \geqslant T$. These are cases (1) and (3) in the statement of the corollary. Cases (2) and (4) arise if $y_{ij}(T) \leqslant \frac{1}{2}$. Obviously cases (1)–(4) exhaust all the possibilities.

In cases (1) and (2), x_{ij} is bounded by $\frac{1}{2}$ and by y_{ij} for $t \geqslant T$ while y_{ij} converges monotonically to $\frac{1}{2}$. Thus x_{ij} is forced into an ever smaller interval as t increases and its oscillations, if any, become smaller and smaller. We shall now show that in cases (3) and (4), no oscillations whatsoever occur in x_{ij} if t is taken sufficiently large. Since $\dot{x}_{ij} = -\dot{X}_{ij}$, it suffices to show that for all large t, $\dot{X}_{ij}$ is either nonpositive or nonnegative.

Consider case (3) for specificity. By (18),

$$\dot{X}_{21} = -A_{21}X_{21} + B_{21}Y_{21}$$

where $B_{21} > 0$ and

$$A_{21} = \frac{-\dot{X}_2}{1 - X_2} + \beta\left(1 + \frac{y_{13} + y_{31}}{2}\right).$$

Suppose we can show that there is a $T_2 \geqslant T$ such that A_{21} is positive for $t \geqslant T_2$. Since we are in case (3), $X_{21} = \frac{1}{2} - x_{21} \geqslant 0$ and $Y_{21} = \frac{1}{2} - y_{21} \leqslant 0$, for $t \geqslant T_2$. Thus $\dot{x}_{21} = -\dot{X}_{21} \geqslant 0$ for $t \geqslant T_2$ and x_{21} increases monotonically to $\frac{1}{2}$. An identical argument shows that x_{21} decreases monotonically to $\frac{1}{2}$ in case (4). We now show that such a T_2 exists.

$$\begin{aligned} A_{21} &= \beta\left[\frac{X_2 - X_1 y_{12} - X_3 y_{32}}{1 - X_2} + 1 + \frac{y_{13} + y_{31}}{2}\right] \\ &= \beta\left[\frac{1 - X_1 y_{12} - X_3 y_{32}}{1 - X_2} + \frac{y_{13} + y_{31}}{2}\right] \\ &\geqslant \beta\left[\frac{1 - \frac{1}{2}(X_1 + X_3) - X_1(y_{12} - \frac{1}{2}) - X_3(y_{32} - \frac{1}{2})}{1 - X_2}\right] \\ &= \beta\left[\frac{1}{1 - X_2} - \frac{1}{2} - \frac{X_1(y_{12} - \frac{1}{2}) - X_3(y_{32} - \frac{1}{2})}{X_1 + X_3}\right] \\ &\geqslant \beta(\tfrac{1}{2} - | y_{12} - \tfrac{1}{2} | - | y_{32} - \tfrac{1}{2} |). \end{aligned}$$

Since $\lim_{t\to\infty} y_{12}(t) = \lim_{t\to\infty} y_{32}(t) = \frac{1}{2}$, we can obviously choose a T_2 such that

$$A_{21} \geqslant \frac{\beta}{4} > 0 \qquad \text{for} \quad t \geqslant T_2,$$

and the proof is complete.

4. A Related Graph with Loops

We now consider a graph which has the same local geometrical properties as (*), but whose global behavior as a flow differs from that of (*). (*) describes a flow over a complete 3-graph without loops. Each vertex of this graph receives two flow arrows and sends out two flow arrows. Another complete graph exists which also has this property, namely the complete 2-graph with loops, which obeys the equations

$$\dot{x}_i(t) = -\alpha x_i(t) + \beta \sum_{m=1}^{2} x_m(t)\, y_{mi}(t),$$

$$y_{ik}(t) = z_{jk}(t)[z_{jj}(t) + z_{jk}(t)]^{-1} \qquad (**)$$

and

$$\dot{z}_{jk}(t) = -u z_{jk}(t) + \beta x_j(t)\, x_k(t),$$

for all $i, j, k = 1, 2$ [3]. An observer sitting on a vertex in either the complete 3-graph without loops or the complete 2-graph with loops cannot tell from the immediate geometry which graph he is in. Nonetheless, the dynamics of the two graphs are dramatically different, as the following theorem concerning the ratios $y_{jk}(t)$ and $X_i(t) = x_i(t)[x_1(t) + x_2(t)]^{-1}$ shows.

Theorem 2. *Let (**) be given with arbitrary nonnegative initial data and arbitrary positive coefficient β. Then*

(1) *the limits $Q_i = \lim_{t\to\infty} X_i(t)$ and $P_{jk} = \lim_{t\to\infty} y_{jk}(t)$ exist and obey the equations*

$$Q_i P_{ij} = Q_j P_{ji}, \qquad \{i, j\} = \{1, 2\}.$$

(2) *the functions $f_{jk} = X_k - y_{jk}$ and $\dot{y}_{jk}$ change sign at most once and not at all if $f_{jk}(0) f_{kk}(0) \leqslant 0$, $\{j, k\} = \{1, 2\}$.*

(3) *the initial data can be chosen so that the limits Q_i and P_{jk} satisfy the equations $P_{ji} = Q_i$, where the limits Q_1, Q_2 can form an arbitrary probability distribution.*

If moreover $\sigma \equiv u + 2(\beta - \alpha) > 0$, then

(4) *the limits Q_i and P_{jk}* always *satisfy the equations $P_{ji} = Q_i$.*

Theorem 2 differs dramatically from Theorem 1 because the probabilities Q_i and P_{jk} are not uniquely determined in Theorem 2 even if $\sigma > 0$. Adding the assumption $z_{12}(0) = z_{21}(0)$ in no way changes this situation, since the relative size of the initial values $z_{11}(0)$ and $z_{22}(0)$ is not affected by this condition. Theorems 1 and 2 together show that a vertex "knows" whether or not the flow it receives comes from another vertex or itself, in the sense that the limiting behavior of its vertex function differs in the two cases.

Proof. The strategy of the proof is essentially the same as that of Theorem 1. We therefore exhibit only the relevant formulas and presuppose familiarity with previous arguments to immediately draw conclusions from these formulas. We also assume that all initial data are positive unless otherwise stated, since all other cases can be handled with ease once this case is understood.

(I) The first step in the proof is to derive equations for the ratios X_i and y_{jk}. These are readily seen to be the following by the usual manipulations:

$$\dot{X}_i(t) = \beta[-X_i(t) + X_i(t)\, y_{ii}(t) + X_j(t)\, y_{ji}(t)], \qquad \{i, j\} = \{1, 2\} \quad (40)$$

and

$$\dot{y}_{ji}(t) = B_j(t)[X_i(t) - y_{ji}(t)], \qquad j, i = 1, 2 \qquad (41)$$

where

$$B_j = \frac{\beta x_j x}{z^{(j)}}, \qquad x = x_1 + x_2, \qquad \text{and} \qquad z^{(j)} = z_{jj} + z_{ji}.$$

Since $X_1 + X_2 = 1$, (41) gives

$$\dot{X}_i = \beta[X_i(y_{ii} - X_i) + X_j(y_{ji} - X_i)]. \tag{42}$$

(II) Using these equations, we show that the limits P_{ij} exist and establish some properties of these limits. Subtracting (41) from (40) gives

$$\dot{f}_{ji} = -(\beta X_j + B_j) f_{ji} - \beta X_i f_{ii} \tag{43}$$

and by renaming indices

$$\dot{f}_{ii} = -(\beta X_i + B_i) f_{ii} - \beta X_j f_{ji} \tag{44}$$

where $f_{uv} = X_v - y_{uv}$, and $\{i, j\} = \{1, 2\}$. From (43) and (44) it is obvious by the positivity of βX_i and βX_j, and the argument used in Lemma 2 that if $f_{ji}(t_0) < 0$ and $f_{ii}(t_0) > 0$, then $f_{ji}(t) < 0$ and $f_{ii}(t) > 0$ for all $t \geqslant t_0$. Similarly, $f_{ji}(t_0) > 0$ and $f_{ii}(t_0) < 0$ implies $f_{ji}(t) > 0$ and $f_{ii}(t) < 0$ for all $t \geqslant t_0$. The same facts hold when the strict inequalities are replaced by weak inequalities. Moreover $f_{ii}(t_0) = f_{ji}(t_0) = 0$ implies $f_{ji}(t) = f_{ii}(t) = 0$ for all $t \geqslant t_0$. It is therefore obvious that the functions f_{ji} and f_{ii} change sign at most once, and not at all if $f_{ji}(0) f_{ii}(0) \leqslant 0$. Also, $f_{ji}(t)$ and $f_{ii}(t)$ are identically zero if $f_{ji}(0) = f_{ii}(0) = 0$.

By (41) and the positivity of B_j, $\dot{y}_{ji}$ changes sign at most once and not at all if $f_{ji}(0) f_{ii}(0) \leqslant 0$. Thus y_{ji} is a monotonic function for large t. Since y_{ji} is also bounded and continuous, $P_{ji} = \lim_{t\to\infty} y_{ji}(t)$ exists.

Moreover $\dot{y}_{ij}(t) = 0$ if $f_{ji}(0) = f_{ii}(0) = 0$. In this case $y_{ji}(t)$ is a constant, and since $f_{ji}(t) = f_{ii}(t) = 0$, $X_i(t) = y_{ji}(t) = y_{ii}(t) =$ constant. In particular, $Q_i = P_{ji} = P_{ii}$.

(III) We now use the fact that the limits P_{jk} exist to show that the limits Q_i exist. Since $X_1 + X_2 = 1$, it suffices to prove the existence of Q_1. By (40),

$$\begin{aligned}\dot{X}_1 &= \beta((y_{11} - 1) X_1 + (1 - X_1) y_{21}) \\ &= \beta(-(y_{12} + y_{21}) X_1 + y_{21}),\end{aligned}$$

which has the integral form

$$X_1(t) = \exp\left(-\int_0^t U_1(\xi)\, d\xi\right) \times \left[X_1(0) + \int_0^t \frac{y_{21}(v)}{y_{21}(v) + y_{12}(v)} \frac{d}{dv} \exp\left(\int_0^v U_1(\xi)\, d\xi\right) dv\right].$$

where $U_1 = \beta(y_{12} + y_{21})$ and y_{21} are positive and have finite limits as $t \to \infty$. It is therefore obvious that Q_1 exists.

To show that the equations $Q_i P_{ij} = Q_j P_{ji}$ hold, note that (40) can be written as

$$\begin{aligned}\dot{X}_i &= \beta[(y_{11} - 1) X_i + X_j y_{ji}] \\ &= \beta(-X_i y_{ij} + X_j y_{ji}).\end{aligned}$$

Since all the limits Q_i and P_{jk} exist, the limit $\lim_{t\to\infty} \dot{X}_i(t)$ also exists and equals $\beta(-Q_i P_{ij} + Q_j P_{ji})$. Since X_i is bounded, $\lim_{t\to\infty} \dot{X}_i(t) = 0$ and thus $Q_i P_{ij} = Q_j P_{ji}$.

(IV) We know from the case $X_1(0) = y_{11}(0) = y_{21}(0)$ and $X_2(0) = y_{22}(0) = y_{12}(0)$ that any probability distributions $Q_1 = P_{11} = P_{21}$ and $Q_2 = P_{22} = P_{12}$ can arise as limits. We now show that only limits Q_i and P_{jk} subject to these constraints can arise for $\sigma > 0$, and that these limits are positive if the initial data is positive. This we prove in two cases, in which we again always assume that all initial data is positive.

Case 1. Suppose $Q_1 = 0$. Then $Q_2 = 1$, and $0 = Q_1 P_{12} = Q_2 P_{21} = P_{21}$. Since $y_{21}(t)$ is a positive and monotonic function for large t and $P_{21} = 0$, $y_{21}(t)$ is a monotone *decreasing* function for large t. By (41) and the positivity of B_2, $y_{21}(t) \geqslant X_1(t)$ for large t. Consider (40). Then

$$\begin{aligned}\dot{X}_1 &= \beta(-X_1 y_{12} + X_2 y_{21}) \\ &= \beta[X_2(y_{21} - X_1) + X_1(X_2 - y_{12})] \\ &\geqslant \beta X_1(X_2 - y_{12}).\end{aligned}$$

Two possibilities arise. Either $X_2 - y_{12} \geqslant 0$ for all large t, or $X_2 - y_{12} \leqslant 0$ for all large t. In the former case, $\dot{X}_1 \geqslant 0$ for all large t. Since X_1 is positive, we conclude that $Q_1 > 0$, which is contradiction. Suppose $X_2 - y_{12} \leqslant 0$ for all large t. Since $X_2 - y_{12} = 1 - X_1 - 1 + y_{11} = y_{11} - X_1$, then $y_{11} - X_1 \leqslant 0$ for all large t. By (41) $\dot{y}_{11} \geqslant 0$ and y_{11} is monotone increasing for all large t. Thus $y_{11} \leqslant X_1 \leqslant y_{21}$ for all large t, where y_{11} is a monotone increasing and positive function. Hence $P_{21} > 0$, which is a contradiction.

We have hereby shown that $Q_1 > 0$. Similarly $Q_2 > 0$.

Case 2. Suppose $Q_1 > 0$ and $Q_2 > 0$. Since $Q_1 > 0$,

$$\begin{aligned}\lim_{t\to\infty} B_1(t) &= \lim_{t\to\infty} \frac{X_1(t)}{\gamma_1 e^{-\sigma t} + \int_0^t e^{-\sigma(t-\xi)} X_1(\xi)\, d\xi} \\ &= \frac{Q_1}{Q_1/\sigma} \\ &= \sigma \\ &> 0.\end{aligned}$$

By (41), $\lim_{t\to\infty} \dot{y}_{12}(t)$ exists and equals $\sigma(Q_2 - P_{12})$. Since y_{12} is bounded, $\lim_{t\to\infty} \dot{y}_{12}(t) = 0$, and thus $Q_2 = P_{12}$. In a similar fashion we find $Q_2 = P_{22} = P_{12}$ and $Q_1 = P_{11} = P_{21}$, which concludes the proof.

The way in which the common limit $Q_i = P_{ii} = P_{ji}$ is approached by X_i, y_{ii}, and y_{ji} for large t can be spelled out very precisely. Exactly two kinds of alternatives exist in an interval $[T, \infty)$ if T is chosen sufficiently large. These alternatives are graphed in Figure 4. Thus either X_i approaches

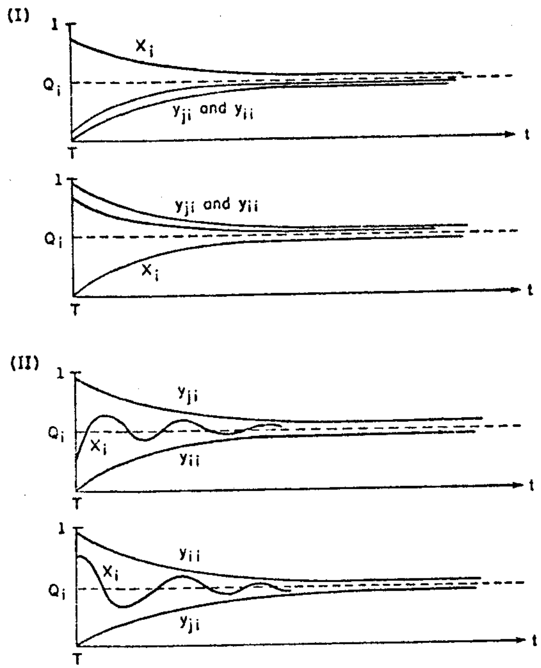

FIG. 4

its limit monotonically in the opposite sense from the monotonic approach of y_{ji} and y_{ii}, or the oscillations of X_i, if any, are pinched between y_{ji} and y_{ii} as they approach the common limit Q_i in opposite senses. These alternatives are proved in the following corollary.

COROLLARY 2. *One of the following alternatives holds for each triple* (X_i, y_{ji}, y_{ii}), $\{i, j\} = \{1, 2\}$.

(1) $y_{ji} \geq X_i \geq y_{ii}$, y_{ji} *is monotone decreasing, and* y_{ii} *is monotone increasing for all* $t \geq 0$,

(2) $y_{ii} \geq X_i \geq y_{ji}$, y_{ii} *is monotone decreasing, and* y_{ji} *is monotone increasing for all* $t \geq 0$.

(3) $X_i \geq y_{ji}$, $X_i \geq y_{ii}$, y_{ji} *and* y_{ii} *are monotone increasing, and* X_i *is monotone decreasing for all* $t \geq 0$.

(4) $y_{ji} \geq X_i$, $y_{ii} \geq X_i$, y_{ji} *and* y_{ii} *are monotone decreasing, and* X_i *is monotone increasing for all* $t \geq 0$.

(5) *Either* (3) *or* (4) *holds for all* t, *or becomes* (1) *or* (2) *for all large* t.

In all cases, the common limit $Q_i = P_{ji} = P_{ii}$ *lies within the interval* $[m_i, M_i]$, *where*

$$m_i = \min\{X_i(0), y_{ji}(0), y_{ii}(0)\} \qquad \text{and} \qquad M_i = \max\{X_i(0), y_{ji}(0), y_{ii}(0)\}.$$

Proof. (1) is a translation of two facts. Firstly $f_{ji} \leq 0 \leq f_{ii}(t)$ for all $t \geq 0$. Secondly $\dot{y}_{ii} = B_i f_{ii} \geq 0$ and $\dot{y}_{ji} = B_j f_{ji} \leq 0$ for all $t \geq 0$. (2) is proved in a similar way. (3) says $f_{ji}(t) \geq 0$ and $f_{ii}(t) \geq 0$ for all $t \geq 0$. By (42), $\dot{X}_i = -\beta(X_i f_{ii} + X_j f_{ji}) \leq 0$ for all $t \geq 0$. (4) is the same situation as (3) with all inequalities reversed. Theorem 2 says that either (3) or (4) hold, or one of the functions $f_{ji}(t)$ and $f_{ii}(t)$ eventually changes sign. This is case (5).

The following corollary can be proved in the same way that Part (4) of Theorem 1 was proved.

COROLLARY 3. *For arbitrary positive initial data and any* σ,

$$| P_{jk} - y_{jk}(0)| \leq 2 \log \left(1 + \frac{1}{\gamma_j} \int_0^\infty e^{\sigma v} X_j \, dv\right),$$

where

$$\gamma_j = \frac{z^{(j)}(0)}{\beta x^2(0)} > 0.$$

In particular, when $\sigma < 0$,

$$| P_{jk} - y_{jk}(0)| \leq 2 \log \left(1 + \frac{1}{\gamma_j | \sigma |}\right).$$

Thus taking $\sigma < 0$ and $|\sigma| \gg 0$ forces P_{jk} to lie very close to $y_{jk}(0)$. We therefore find that a complete 2-graph with loops can remember its initial data both when $\sigma > 0$ and when $\sigma < 0$.

Remark (*Average Output* vs. *Individual Outputs*). In both (*) and (**), when these graphs are input-free, the average output $x = 1/n \sum_{k=1}^{n} x_k$ obeys the linear equation

$$\dot{x} = (\beta - \alpha)\, x.$$

Nonetheless, the individual outputs x_i of these graphs obey nonlinear equations and exhibit substantially different qualitative behavior.

Part II

5. Learning Theoretic Interpretation

The learning theory described by the system (*) provides a mathematical description of the following kind of experiment. An experimenter $\mathscr{E}$, confronted by a machine $\mathscr{M}$, presents $\mathscr{M}$ with a list of "letters", "spatial patterns", or more generally "events" to be learned. Suppose, for example, that $\mathscr{E}$ wishes to teach $\mathscr{M}$ the list of letters AB, or to predict the event B, given the event A. $\mathscr{E}$ does this by presenting A and then B to $\mathscr{M}$ several times. To find out if $\mathscr{M}$ has learned the list as a result of these list presentations, the letter A alone is then presented to $\mathscr{M}$. If $\mathscr{M}$ responds with the letter B, and $\mathscr{M}$ does this whenever A alone is said, then we have good evidence that $\mathscr{M}$ has indeed learned the list AB. Thus $\mathscr{M}$ learns to predict the event B whenever the event A occurs as a result of repeated presentations of the list AB.

In order to translate into formal terms the intuitive idea of presenting a list AB of events to $\mathscr{M}$, we assign one vertex of $\mathscr{M}$ to each distinct symbol of an event. If for example we are given three symbols A, B, and C, then we assign v_1 to A, v_2 to B, and v_3 to C. Given this assignment of symbols to vertices, suppose that $\mathscr{E}$ wishes to teach $\mathscr{M}$ to predict B given A. He indicates to $\mathscr{M}$ that B is the "correct" successor of A by repeating the desired sequence AB several times. AB is presented to M by perturbing the vertices v_1 and v_2 which stand for A and B, respectively, by inputs. Each presentation of a symbol to $\mathscr{M}$ at a time t_0 is represented by an "input pulse" to the corresponding vertex with "onset time" t_0. An *input pulse* is a continuous and nonnegative function J which is positive in a finite interval. The *onset time* of J is $\inf\{t : J(t) > 0\}$. For example, if A and B are presented to $\mathscr{M}$ r times at a periodic rate with A occurring at times $t = 0$, $w + W$, $2(w + W), \ldots, (r - 1)(w + W)$, and B occurring at times $t = w$, $2w + W$, $3w + 2W, \ldots, rw + (r - 1) W$, then

$$I_1(t) = \sum_{k=0}^{r-1} J_1(t - k(w + W)),$$

$$I_2(t) = \sum_{k=0}^{r-1} J_2(t - w - k(w + W)),$$

and

$$I_3(t) = 0,$$

where J_1 and J_2 are input pulses with onset time zero.

By (1), an input $I_i(t)$ corresponding to presentation by $\mathscr{E}$ of an experimental "symbol" or "stimulus" to $\mathscr{M}$ perturbs $x_i(t)$, which is therefore heuristically thought of as a "stimulus trace".

The function $y_{jk}(t)$ in (1) and (2) is thought of as the "associational strength" from v_j to v_k at time t since it controls the size of an output from v_k due to an isolated experimental input to v_j (i.e., the size of the prediction created by the presentation of a symbol). The definition of $y_{jk}(t)$ aims at guaranteeing that $y_{jk}(t)$ will grow only if the symbols corresponding to v_j and v_k are often consecutively presented to $\mathscr{M}$.

6. The Memory of a Graph

We will be able to interpret Theorems 1 and 2 as statements concerning how $\mathscr{M}$ remembers what it has been taught by $\mathscr{E}$. These statements are made rigorous using the following colorful language. If the associational strength $y_{jk}(t)$ changes very little during a time interval $[T_0, T_1]$, we say that $\mathscr{M}$ *remembers* the association from v_j to v_k during these times. If, moreover, $y_{ij}(t) = y_{ik}(t) = \frac{1}{2}$, $\{i, j, k\} = \{1, 2, 3\}$, and $x_1(t) = x_2(t) = x_3(t)$ during a time interval $[T_0, T_1]$, we say that $\mathscr{M}$ is in a state of *maximal ignorance* during these times, since all associational strengths $y_{jk}(t)$ with $j \neq k$ are equal, and all vertices have equal stimulus traces due to prior inputs. If all associational strengths $y_{jk}(t)$ with $j \neq k$ converge to $\frac{1}{2}$ as $t \to \infty$ and all stimulus traces are asymptotically equal as $t \to \infty$, then $\mathscr{M}$ *forgets* what it has learned from $\mathscr{E}$, since it returns eventually to a state of maximal ignorance. Any nonnegative and continuous input to $\mathscr{M}$ that eventually becomes zero can be interpreted as a learning experiment performed by $\mathscr{E}$ on $\mathscr{M}$ which ends after a finite amount of "practice" time.

Theorem 1 discusses any $\mathscr{M}$ which is initially in a state of maximal ignorance after it undergoes an arbitrary learning experiment that ends after a finite amount of practice time. The study of $\mathscr{M}$ after practice ceases is thus the study of how well $\mathscr{M}$ remembers what it has learned from the experiment.

Theorem 1(1) shows that all associational strengths and the relative and absolute sizes of the stimulus traces have limits as $t \to \infty$. By Theorem 1(3), $\mathscr{M}$ forgets what it has learned whenever $\sigma > 0$. By Theorem 1(4), $\mathscr{M}$ remembers what it has learned with an arbitrarily good accuracy if $\sigma < 0$ and $|\sigma|$ is taken sufficiently large. By Theorem 1(2), the associational strengths $y_{jk}(t)$ approach their limits essentially monotonically.

Theorem 2 shows that a complete 2-graph with loops will not forget its associations $y_{jk}(t)$ even if $\sigma > 0$. Thus, changing the geometry of a graph –in the present case by adding loops–can change the way in which the graph remembers its past experiences.

The following sections interpret Theorem 1 in several ways to bring to the reader's attention some suggestive heuristic relationships between the behavior of (*) and more familiar mathematical structures.

7. Connection with Finite Markov Chains

The system of equations

$$Q_i = Q_j P_{ji} + Q_k P_{ki}, \qquad \{i, j, k\} = \{1, 2, 3\}, \tag{42}$$

describes an "equilibrium" or "stationary" state that (*) approaches as $t \to \infty$. When for example $\sigma < 0$ and $|\sigma| \gg 0$, the probabilities $y_{jk}(t)$ move very little, by (40). Nonetheless, the probabilities $X_i(t)$ adjust themselves as much as is required to satisfy (42).

The equations (42) suggest a relationship between (*) and the theory of finite Markov chains [7] if we interpret P_{ij} as the stationary probability of going from state i to state j, and Q_k as the probability of being in state k. We denote this Markov chain by $G(\infty)$, and for every $t \geqslant 0$ define a Markov chain $G(t)$ with transition probabilities $y_{ij}(t)$. Then (*) describes a nonlinear process whose transition probabilities $y_{ij}(t)$ fluctuate as $t \to \infty$ but ultimately approach stationary transition probabilities P_{ij} if the initial process $G(0)$ is "spatially homogeneous" (i.e., has uniform initial data) and the process is perturbed by inputs I_i only over a finite time interval. If we realize each Markov chain $G(t)$, $0 \leqslant t \leqslant \infty$, as a probabilistic graph [4] with weight function φ_t such that $\varphi_t(e_{ij}) = y_{ij}(t)$, then (*) can alternatively be described as a nonlinear mechanism which continuously deforms one probabilistic graph $G(t_0)$ into a future probabilistic graph $G(t_1)$, $t_1 > t_0$. Theorem 1 hereby becomes a kind of "homotopy" theorem [8].

8. A Relationship between Measurement, Linearity, and Reversibility

We can give Theorem 1 another heuristic interpretation which is quite suggestive of situations familiar from statistical mechanics [9]. We do this when (*) is given at $t = 0$ with $\sigma > 0$ and uniform initial data (i.e., $z_{ij}(0) = \delta > 0$, $i \neq j$, and $x_i(0) = \gamma > 0$, $i = 1, 2, 3$). v_1 is perturbed by inputs within the finite time interval $(T_1, T_2]$, where $0 < T_1 < T_2 < \infty$, and (*) is input-free in (T_2, ∞).

Clearly, $x_i(t) = \frac{1}{3}x(t)$ for $t \in [0, T_1]$, where $x = \sum_{k=1}^{3} x_k$. Thus $\dot{x}_i(t) = (\beta - \alpha)\, x_i(t)$ for $t \in [0, T_1]$, so that the output from each v_i is *linear*. Moreover, $z_{ij}(t)$ is independent of i and j, $i \neq j$, for $t \in [0, T_1]$, so $y_{ij}(t) = \frac{1}{2} = y_{ji}(t)$. That is, the flow from v_i to v_j and from v_j to v_i is *globally reversible* ("globally" because y_{ij} depends on all indices i, j, k, and controls the size of the flow between vertices).

In $(T_1, T_2]$, the output from (*) is obviously *nonlinear*. Moreover $y_{12}(t) \neq y_{21}(t)$, so that the flow between v_1 and v_2 is *globally irreversible*. By contrast, $z_{ij} = z_{ji}$ so that the flow is still *locally reversible* ("locally" because z_{ij} depends only on indices i and j).

In (T_2, ∞), (*) is input-free and $z_{ij}(T_2) = z_{ji}(T_2)$. By Theorem 1, $\lim_{t\to\infty} x_1(t)/\frac{1}{3}x(t) = 1$, where $\frac{1}{3}x$ obeys a linear equation. That is, the output of each v_i is *eventually linear* once again. Moreover, $\lim_{t\to\infty} y_{ij}(t)/y_{ji}(t) = \frac{1}{2}/\frac{1}{2} = 1$. That is, the flows within (*) are *eventually globally reversible* once again.

The input to v_1 for $t \in (T_1, T_2]$ is again interpreted as a measurement performed by an experimenter $\mathscr{E}$ studying (*). This example therefore illustrates that a measurement can transform a linear and globally reversible system into a nonlinear and globally irreversible system, but that linearity and global reversibility are gradually restored as the effect of the measurement wears off. The measurement does not affect local reversibility.

When $\sigma < 0$, we again begin with a linear, and both locally and globally reversible system. Again this system remains locally reversible throughout all inputs. Again this system becomes eventually linear since all Q_i exist. But we no longer can assert that the system is eventually globally reversible. Moreover, the alternative between eventual global reversibility and irreversibility depends on a parameter σ which cannot be directly measured by $\mathscr{E}$. Also this alternative does not correspond to an alternative between linearity and nonlinearity of the system, since in both cases the system is eventually linear.

9. Output Size vs. Memory

Only the case $\alpha > \beta$ in Theorem 1 has prediction theoretic interest since then the outputs approach zero as $t \to \infty$ if $\mathscr{E}$ does not perturb $\mathscr{M}$. Also the constraint $u > 0$ is imposed to guarantee that present inputs to $\mathscr{M}$ have a greater importance than past inputs. A major concern of our prediction theory is to study those properties of $\mathscr{M}$ which $\mathscr{E}$ cannot directly measure in terms of inputs and outputs when $\alpha > \beta$ and $u > 0$.

For example, Theorem 1 shows that there is little connection between the absolute magnitude of the outputs produced by $\mathscr{M}$ (when $\mathscr{E}$ does not perturb $\mathcal{M}$) and $\mathcal{M}$'s "memory". This is because Theorem 1 describes an alternative depending on the sign of σ which decides the limiting behavior of the probabilities $y_{jk}(t)$ as $t \to \infty$, whereas the outputs $x_i(t)$ converge exponentially to zero as $t \to \infty$ whenever $\alpha > \beta$. The condition $\sigma > 0$ can be guaranteed by choosing $u > 2(\alpha - \beta) > 0$. Then (*) "forgets" its past as $t \to \infty$, since the probabilities $X_i(t)$ and $y_{jk}(t)$ eventually return to a state of "maximal ignorance"; that is, $Q_i = \frac{1}{3}$ and $P_{jk} = \frac{1}{2}(1 - \delta_{jk})$ are uniformly distributed. When $\sigma < 0$, or $0 < u < 2(\alpha - \beta)$, the graph "remembers" its past at least partially as $t \to \infty$, since the maximal deviation of P_{jk} from $y_{jk}(0)$ decreases as $|\sigma|$ increases.

In both cases, after $\mathcal{E}$ ceases to perturb $\mathcal{M}$, the outputs x_j approach zero. A plausible inference is that both graphs are forgetting $\mathcal{E}$'s prior inputs as the effect of inputs on outputs wears off. This is, however, true only when $\sigma > 0$.

10. Rate of Excitation and Recovery vs. Memory

Another limitation on $\mathcal{E}$'s information concerning $\mathcal{M}$ follows from the fact that the rate with which $\mathcal{M}$ reacts to and recovers from $\mathcal{E}$'s inputs is not closely related to $\mathcal{M}$'s "memory". The parameter β is the rate at which each vertex function x_i is "excited" by an input from another vertex, and α gives the rate of x_i decay as the effect of this input wears off. u is the rate at which the "cross-correlation" $\beta x_j(t)\, x_k(t)$ of pulses $\beta x_j(t)$ and $x_k(t)$ at the arrowhead of e_{jk} wears off.

Let $\alpha = \beta$. If also $\alpha \gg 0$, then the rate of excitation and of decay at the vertices is very large and of the same size. Such a system has the virtue that its response to inputs is rapid and does not introduce large "inertial" effects. If moreover $u > 0$, then also $\sigma > 0$ and $\mathcal{M}$ eventually forgets everything. We can therefore conceive situations in which $\mathcal{M}$ responds as quickly as we please to inputs and in which its cross-correlations die away as slowly as we please, and still $\mathcal{M}$ eventually forgets everything that the inputs have taught to it!

By contrast we can conceive situations in which $\mathcal{M}$'s excitation and decay rates are large and of the same order of magnitude (say $\beta = \frac{19}{20}\alpha \gg 0$) and can still guarantee that $\mathcal{M}$ remembers its past as well as we please by letting the cross-correlations decay sufficiently slowly (say $0 < u \ll \frac{1}{10}\alpha = 2(\alpha - \beta)$ and thus $\sigma < 0$ with $|\sigma| \gg 0$).

These examples show clearly that the absolute response rate of $\mathcal{M}$ to $\mathcal{E}$'s inputs is not closely related to $\mathcal{M}$'s memory. $\mathcal{E}$'s dilemma is enhanced by the existence of a graph, namely a complete graph with loops (Section 4), which remembers certain initial data when $\sigma < 0$ and $|\sigma| \gg 0$.

References

1. Grossberg, S., Nonlinear difference-differential equations in prediction and learning theory. *Proc. Nat'l. Acad. Sci.* 58(4), 1329–1334 (1967).
2. Grossberg, S., Global ratio limit theorems for some nonlinear functional-differential equations (I), (II). *Bull. Amer. Math. Soc.* 74(1), 95–105.
3. Grossberg, S., A prediction theory for some nonlinear functional-differential equations (I). *J. Math. Anal. Applics.* 21(3), 643–694 (1968).
4. Busacker, R. G. and Saaty, T. L., "Finite Graphs and Networks." McGraw-Hill, New York, 1953.
5. Ford, L. R., Jr. and Fulkerson, D. R., "Flows in Networks." Princeton Press, New Jersey, 1962.
6. Coddington, E. A. and Levinson, N., "Theory of Ordinary Differential Equations." McGraw-Hill, New York, 1955.
7. Kemeny, J. G. and Snell, J. L., "Finite Markov Chains." Van Nostrand, 1960.
8. Hu, S-T., "Homotopy Theory." Academic Press, New York, 1959.
9. Kac, M., "Probability and Related Topics in Physical Sciences." Wiley (Interscience), New York, 1959.

PRINTED IN BRUGES, BELGIUM, BY THE ST. CATHERINE PRESS LTD