A Model for the Organic Synchronization of Communications Systems

By M. KARNAUGH

(Manuscript received August 16, 1966)

Organic synchronization is a method for the mutual synchronization of a set of geographically separated clocks. It is applicable to pulse code modulation (PCM) communications networks and to other systems which have similar requirements for synchronism.

After a brief review and history of the problem, a model for organic synchronization is developed. A control-independent study of possible equilibrium solutions is then carried out. A special class of controls is shown to provide asymptotic stability in the limiting case of zero delays. This result leads heuristically to the synthesis of a broad class of nonlinear controls. With these controls, the systems are represented by families of nonlinear differential-functional equations. This model provides a basis for the simulation of organic synchronization. Broad conditions which are mathematically sufficient for the stability of the nonlinear systems are not yet known. The final frequencies of a linear subclass of organic systems, known to be stable, are examined.

I. INTRODUCTION

The timing of the switching actions at each switching center of a pulse code modulated (PCM) communications system is governed by a device called the "local clock." It may consist of a cyclic counter driven by an oscillator. Each cycle of the counter is then one clock cycle.

In a geographically widespread PCM system, the local clocks may be either autonomous or synchronized. This choice should be made with the best possible knowledge of the available technology, as well as consideration of its functional and economic consequences. The choice is clearly a rather basic one, and it may have long term effects upon the evolution of the system.

The time-multiplexed PCM signals arriving at any locality may have arbitrary, and usually scattered, points of origin. Some of them require decoding into a common analog form. In particular, they may be voice signals. A homogeneous, time-multiplexed set of such signals is easily decoded by a common digital-to-analog converter, provided that the transmitted samples have been generated synchronously. A nonsynchronous alternative is to insert extra digits into the signals in order to permit multiplex transmission. Additional equipment is needed to remove these digits and smooth the timing of the demultiplexed samples before or after decoding them.

This paper is only one of a number of studies of system synchronization, and it does not provide a complete solution to the problems touched upon. After a very brief review of some past work in this field, I shall go back to fundamentals to derive a model for organic synchronization. Following this, the sections entitled "Equilibrium Points", "Reduced System Equations", "Controls: Qualitative Discussion", and "System with Zero Delays" provide background for the synthesis of a family of controls which is introduced under the heading, "A Family of Realizable Organic Systems".

The question of the final frequencies of certain linear organic systems is then taken up. Finally, some remarks are made to clarify the stability problem.

II. HISTORY

The synchronization of PCM networks has long been a subject of interest. The question of synchronizing switching centers, in addition to the transmission links, arose in 1956, when the PCM telephone switching experiment, later named Essex,[1] was planned.

The term "organic synchronization", which seems to have been introduced by V. E. Beneš,[2] will be used herein for systems fitting the model to be derived in later sections. The systems treated by Beneš, excepting a certain minor idealization, form a subclass of these systems. This same subclass of systems is discussed in a patent[3] by J. P. Runyon.

Beneš[2] has demonstrated asymptotic stability for his systems, which are linear, under quite interesting conditions. He has also given formulas for the asymptotic system frequency and for the asymptotic relative phases of the oscillators. A. J. Goldstein[4] has given simplified derivations of these formulas.

An alternative mutual synchronization method, called "frequency averaging", has been treated by Beneš and Goldstein.[5] Frequency averaging systems, while stable, lack a frequency determining element. Each oscillator puts out the average of the frequencies received from its neighbors, and the system frequency will wander in the presence of noise. Because of this feature, it does not seem to be very practical, unless it is combined with other techniques.

The transmission of a synchronizing signal from a master oscillator to all other oscillators, which are locked to this signal, is perhaps the simplest approach to synchronization. However, such a system is vulnerable to failure of the master oscillator or failure of a transmission link. Means for mitigating this weakness have been proposed by G. P. Darwin and R. C. Prim.[6] They equip the system with automatic means to reorganize itself in the event of a failure. Unfortunately, this adds considerable complexity to the basically simple method.

Further discussion will be limited to organic systems for synchronization.

B. J. Karafin[7] has carried out some digital computer simulations of organic synchronization of small systems. A. Gersho and B. J. Karafin[8] have simplified the proof of asymptotic stability for Beneš' systems. C. J. Candy and M. Karnaugh[9] have studied organic systems of up to four switching centers by means of an analog simulator. M. B. Brilliant has also studied linear organic systems[10] and has computed transient responses of certain large linear systems.[11]

Linear systems with zero delays have also been studied at the University of Tokyo by T. Saito, H. Fujisaki and H. Inose.[12]

III. THE SYNCHRONIZED NETWORK

Consider a set of $N \geqq 2$ geographically separated pulse code switching centers, interconnected by directed pulse transmission links, as illustrated in Fig. 1 for the case $N = 4$.

All possible links need not be physically provided. The cases of greatest interest are those in which there is a directed path from any center

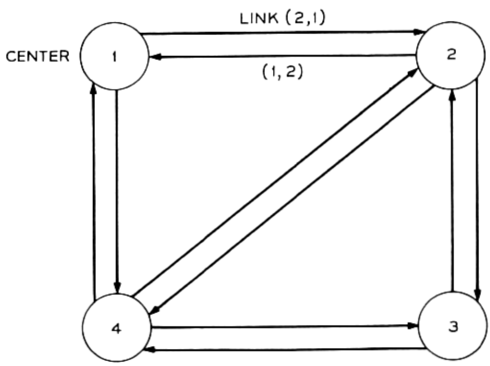

Fig. 1 — A sample network.

to any other center, possibly by way of some intermediate centers. Systems with this property will be called "connected systems".

A connected system of N centers must have at least N links, because at least one link must terminate at each center. When the centers are connected in a unidirectional loop, there are exactly N links. The maximum possible number of links, assuming no duplications, is realized when every ordered pair (i,j) of distinct centers is connected by a link to i from j. This number is $N(N - 1)$. The correspondence between the ordered pair (i,j) and the direction to i from j is a convention which will be followed consistently.

An important component at each center is the local clock which determines the timing of all switching actions at that center. The messages from all other centers arrive in the form of framed pulse codes. These are pulse codes divided into sequences containing equal numbers of digits by means of periodically introduced framing digits. In order for the pulse codes to be correctly processed, a correct phase relationship must exist between the arriving framed code and the local clock.

The desired phase relations are realized by providing a certain amount of buffer storage for each incoming link.[13] Such equipment is illustrated in Fig. 2. The arriving digits are stored in a cyclically addressed discrete memory. They are read out of the memory under control of the local clock and of a circuit which monitors the appearances of the framing digits, so as to be correctly phased.

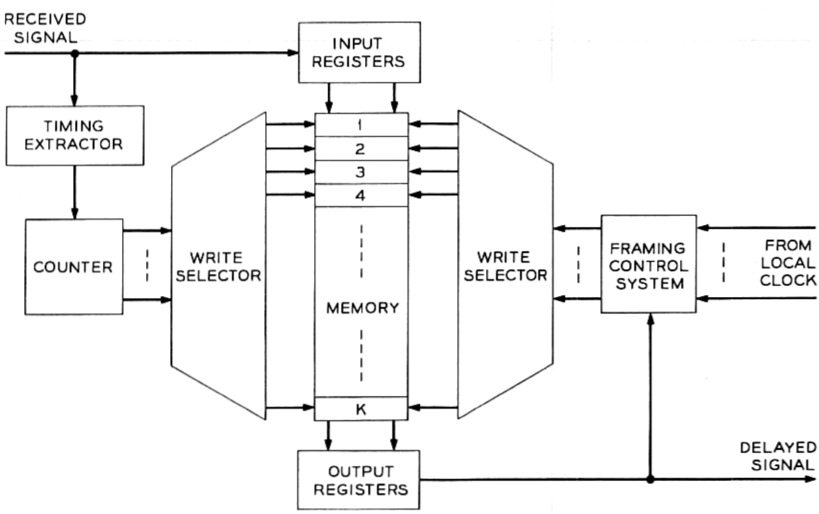

Fig. 2 — Buffer memory.

Other forms of buffer memory which incorporate variable delay lines have also been proposed. These may be acceptable and less costly in some cases.

Under favorable conditions, the arriving signals can be correctly phased by means of the buffers. However, unless the oscillator frequencies are properly controlled, their phase differences will wander beyond any bound. Then, some of the buffer stores will fill up or be emptied, causing erroneous codes to appear at their outputs. It is the primary object of the controls to avoid such malfunctions. The system will be considered to be operating synchronously when no information is being lost in this fashion.

I shall adopt the point of view that it is desired to keep the buffer memories just half full, in which case the system would not be unduly vulnerable to transient disturbances. Controls will be sought which tend to inhibit large deviations from the desired condition. We shall see that these deviations cannot, in general, be reduced to zero. The oscillator control signals will be derived from them.

It should be noted that the transmission delays between centers are variable over some small fractions of their center values. These delays will depend upon the environmental conditions of the propagating media and on message-induced jitter at pulse repeaters.[14] The buffer memories must mop up the delay variations as well as the effects of phase wander of the oscillators.

IV. NOTATION

The single subscripts i, j, k, $\cdots$, refer to the various centers and to the oscillators located at these centers. Their range is the integers, 1, 2, $\cdots$, N. When one of them appears in a statement or equation with no other qualification, the statement or equation holds over the whole range.

It has already been pointed out that the ordered pair (j,i) designates the link to center j from center i. When a statement or equation contains a pair of subscripts with no other qualification, it holds for all pairs (j,i) which designate existing links.

The set of all existing links will be called R. Thus, $(j,i) \in R$ means there is a link to j from i in the system.

Similarly, R_i is the set of links terminating (i.e., receiving) at center i, and S_i is the set of links originating (i.e., sending) at center i. Thus,

$$R = \bigcup_{i=1}^{N} R_i = \bigcup_{i=1}^{N} S_i = \bigcup_{R} \{(j,i)\}.$$

Let M be the number of links in the system. We have seen that

$$N \leqq M \leqq N(N-1).$$

System controls will be supposed in effect for $t \geqq 0$. Prior history of the system provides the initial condition. Statements about functions of t with no other qualification will hold for $t \geqq 0$.

Occasionally, vectors will be used having N singly subscripted components or else M doubly subscripted components. For example, the delay vector $\boldsymbol{\tau}$ has the M components τ_{ji}, $(j,i) \in R$. It will be clear which vector space is meant in each case.

V. PHASE, FREQUENCY, AND DELAY

The local clocks will emit coherent signals. That is, for time intervals which are very long compared to one period of the clock, the output will be approximately periodic. Under these conditions, many formally different definitions of instantaneous frequency will be in good numerical agreement. I shall simply postulate the existence of such continuous functions, $f_i(t)$.

Phases of the oscillators are defined to be

$$p_i(t) = p_i(0) + \int_0^t f_i(s)ds \tag{1}$$

in cycles, and

$$f_i = p_i'. \tag{2}$$

The principal value of the phase is

$$\varphi_i = p_i \text{ modulo } 1 \tag{3}$$

and has the range $0 \leqq \varphi_i < 1$.

The initial condition for the phases will be

$$p_i(0) = \varphi_i(0). \tag{4}$$

The values, $\varphi_i(0)$, are observables of the physical system. In fact, the switching actions at center i are timed according to $\varphi_i(t)$.

If there is a transmission link to center j from center i, the signals transmitted therein will be subject to a time delay $\tau_{ji}(t)$. If a pulse is launched from center i at a time t_1 and received at center j at time t_2, then the delay is defined to be

$$\tau_{ji}(t_2) = t_2 - t_1. \tag{5}$$

Analogously, the phase of the signal received at time t by center j

from center i is defined to be

$$p_{ji}(t) = p_i[t - \tau_{ji}(t)]. \tag{6}$$

It should be recognized that the principal phase of a timing wave recovered from the received pulse code would, in practice, only approximate $\varphi_{ji}(t)$. Errors of a few percent of one pulse period are common in pulse repeaters.[14] This corresponds to a fraction of one percent of a typical frame period.

The frequency of the received signal is obtained by differentiation of (6).

$$f_{ji}(t) = [1 - \tau_{ji}'(t)]\cdot f_i[t - \tau_{ji}(t)]. \tag{7}$$

This equation displays the Doppler shift of frequency caused by variation of τ_{ji}.

The clock frequency at the ith center may be represented in the idealized form

$$f_i(t) = F + E_i + g_i(t) + \eta_i(t). \tag{8}$$

F is the mean center frequency of all clocks in the system, averaged over the time during which the system is observed. E_i is the incremental center frequency of the ith clock, also time averaged. By definition,

$$\sum_{i=1}^{N} E_i = 0. \tag{9}$$

The time function $g_i(t)$ is the contribution of the system controls, while $\eta_i(t)$ is a random noise with zero mean. There will be a symmetrical bound on $g_i(t)$,

$$| g_i(t) | \leqq G, \tag{10}$$

which is supposed to be larger than the other frequency deviations. This is necessary if the controls are to bring all oscillators to the same average frequency. Under realistic conditions of operation,

$$G > \max_i | E_i | + \max_{\substack{i \\ t \geqq 0}} \sigma \left(F \int_t^{t+1/F} \eta_i(s)ds \right) \tag{11}$$

where $\sigma(\cdot)$ is the standard deviation.

VI. FUNCTION OF THE BUFFER MEMORY

The principal phase of the signal arriving at center j from center i will usually not agree with that of the oscillator at j. The purpose of the buffer memory in this link is to delay the signal by an additional time, $d_{ji}(t)$, so that

$$\varphi_j(t) = \varphi_{ji}[t - d_{ji}(t)]. \tag{12}$$

In view of (6), this can be written

$$\varphi_j(t) = \varphi_i(t_{ji}), \tag{13}$$

where

$$t_{ji} = t - d_{ji}(t) - \tau_{ji}[t - d_{ji}(t)]. \tag{14}$$

Taking the right-hand derivative of (13), and because this derivative of the principal phase is always equal to the derivative of the phase, we get

$$f_j(t) = [1 - d_{ji}'(t)] \cdot \{1 - \tau_{ji}'[t - d_{ji}(t)]\} \cdot f_i(t_{ji}), \tag{15}$$

where physical considerations make it clear that

$$|d_{ji}'|, \qquad |\tau_{ji}'| \ll 1.$$

Equation (15) implies the dependence of d_{ji} on f_i, f_j, and τ_{ji}.

Matters can be simplified by shifting attention from the delays in the buffer memories to the numbers of frames, i.e., clock cycles, they contain. The number of cycles in the (j,i) buffer at time t is:

$$y_{ji}(t) = p_{ji}(t) - p_{ji}[t - d_{ji}(t)] \tag{16}$$

and

$$y_{ji}'(t) = f_{ji}(t) - [1 - d_{ji}'(t)] \cdot f_{ji}[t - d_{ji}(t)].$$

Using (7) and (15), this can be put in the form

$$y_{ji}'(t) = f_{ji}(t) - f_j(t), \tag{17}$$

which equates the rate of accumulation of cycles to the rate of arrival minus the rate of removal. In terms of the oscillator frequencies,

$$y_{ji}'(t) = [1 - \tau_{ji}'(t)] \cdot f_i[t - \tau_{ji}(t)] - f_j(t). \tag{18}$$

Suppose the (j,i) buffer has a capacity of $2D_{ji}$ cycles. The normalized state of this buffer is defined to be

$$x_{ji}(t) = D_{ji}^{-1} \cdot [y_{ji}(t) - D_{ji}], \tag{19}$$

which is the fractional deviation of its contents from half its capacity. In terms of this variable, (18) becomes

$$x_{ji}'(t) = D_{ji}^{-1} \cdot [1 - \tau_{ji}'(t)] \cdot f_i[t - \tau_{ji}(t)] - D_{ji}^{-1} \cdot f_j(t). \tag{20}$$

This equation explicitly relates the derivative of the buffer memory state variable, $\mathbf{x}(t)$, to the delays and clock frequencies in the system.

I have remarked that the frequencies should be controlled so as to prevent any buffer memory from emptying or filling. For example, if buffer (j,i) is nearly empty, then we desire $f_i > f_j$ until the situation is sufficiently corrected. On the other hand, if the (j,i) buffer is nearly full, then the inequality is reversed. However, things are complicated by the fact that all buffers associated with the ith center, that is, those in links of the set $R_i \cup S_i$, are affected by a change in f_i.

The system is said to malfunction whenever

$$\max_{(j,i)\epsilon R} | x_{ji}(t) | > 1,$$

that is, whenever the buffer memory state vector leaves the "unit cube".

Defining the nonnegative scalar,

$$r_\infty(\mathbf{x}) = \max_{(j,i)\epsilon R} | x_{ji} |, \tag{21}$$

we see that the system is in a permitted buffer state when

$$r_\infty(\mathbf{x}) \leqq 1. \tag{22}$$

VII. EQUILIBRIUM POINTS

Suppose the system is so controlled that an equilibrium solution to the system equations is possible. That is, in the absence of disturbances,

$$\mathbf{x}'(t) = 0 \quad \text{for} \quad t \geqq 0$$

and there exists a constant f such that

$$f_i(t) = f \quad \text{for} \quad t \geqq 0, \qquad i = 1, 2, \cdots, N.$$

If the system, in or near this state, is disturbed by a change in the network configuration, noise in the oscillators, or changes in some of the delays, then variations in the state of the buffer memories will result. To minimize the chance of malfunction under such disturbances, it seems reasonable to seek an equilibrium in which the buffer state is, in some sense, near $\mathbf{x} = 0$. That is, the buffers are nearly half full.

I shall begin along these lines by seeking the set of equilibrium points which can be reached from arbitrary initial conditions and under any controls whatever. The situation of asymptotic equilibrium to be considered is as follows.

(*i*) $\boldsymbol{\eta}(t) = 0 \quad \text{for} \quad t \geqq 0.$
(*ii*) $\tau_{ji}(t) = \tau_{ji}(0) > 0 \quad \text{for} \quad t \geqq 0, \ (j,i) \epsilon R.$

(*iii*) $\lim_{t\to\infty} f_i(t) = f$ for $i = 1, 2, \cdots, N$.

(*iv*) $\lim_{t\to\infty} [p_i(t) - p(t)] = \tilde{q}_i$ for $i = 1, 2, \cdots, N$,

where

$$p(t) = \frac{1}{N}\sum_{i=1}^{N} p_i(t), \qquad |\tilde{q}_i| < \infty.$$

Let $\mathbf{x} = \lim_{t\to\infty} \mathbf{x}(t)$. The locus of attainable points $\mathbf{x}$ will be examined under the above conditions.

With the delay vector $\boldsymbol{\tau}$ assumed to be constant, (20) has the simple form,

$$x_{ji}' = D_{ji}^{-1}\cdot f_i(t - \tau_{ji}) - D_{ji}^{-1}\cdot f_j(t).$$

Therefore,

$$\begin{aligned} x_{ji} &= x_{ji}(0) + \int_0^\infty x_{ji}'(t)dt \\ &= x_{ji}(0) + D_{ji}^{-1}\int_0^\infty [f_i(t - \tau_{ji}) - f_j(t)]dt. \end{aligned}$$

The integral may be evaluated, using (1) and the conditions (*iii*) and (*iv*) of asymptotic stability. The result is,

$$x_{ji} = D_{ji}^{-1}(\tilde{q}_i - \tilde{q}_j - \tau_{ji}f) + B_{ji}, \tag{23}$$

where

$$B_{ji} = x_{ji}(0) + D_{ji}^{-1}[p_j(0) - p_i(-\tau_{ji})] \tag{24}$$

is a constant which depends upon the initial condition. Equations (23) express the set of buffer memory equilibrium states attainable from a given initial condition in terms of the asymptotic phase differences and the asymptotic system frequency. This set depends upon the initial condition through the parameters, B_{ji}

It is shown in Appendix A that the set of phase differences

$$\{(\tilde{q}_i - \tilde{q}_j) \mid (j,i) \in R\}$$

contains exactly $(N - 1)$ linearly independent elements. There are M components of the asymptotic buffer state vector $\mathbf{x}$, with $M \geqq N$. Therefore, if f were an unconstrained real variable, we see that, as a function of the phase differences and f, $\mathbf{x}$ would range over an N-dimensional linear manifold of its M-dimensional space.

However, the system frequency f must be near F, the average center frequency of the clocks. More specifically,

$$F + \max_i E_i - G < f < F + \min_i E_i + G.$$

This inequality requires that the bound on the magnitude of the frequency control be large enough to reduce the highest frequency below the system frequency and to raise the lowest frequency above the system frequency. Thus, all clocks can be brought to a common frequency, even in the presence of noise. Nevertheless, $G \ll F$ in cases of practical interest.

In a typical application, $G = 10^{-6} F$, so the domain of allowed values of f is a very narrow interval. It may be said that the range of $\mathbf{x}$ is a neighborhood of an $(N - 1)$-dimensional linear manifold. The distance of this manifold from the origin is determined by the initial condition.

In cases of practical interest $\tau_{ji}G \ll 1$, so that

$$p_j(0) - p_i(-\tau_{ji}) \cong \tau_{ji}f + p_j(0) - p_i(0).$$

In such cases, (23) has the approximation

$$x_{ji} \cong x_{ji}(0) + D_{ji}^{-1}[\tilde{q}_i - \tilde{q}_j + p_j(0) - p_i(0)]. \tag{25}$$

From this it is clear that $x_{ji} \cong x_{ji}(0)$ if $\tilde{q}_i - \tilde{q}_j$ equals $p_i(0) - p_j(0)$. Therefore, an asymptotic state vector $\mathbf{x}$ which is, in some sense, small is attainable when the initial state vector $\mathbf{x}(0)$ is small in the same sense.

VIII. REDUCED SYSTEM EQUATIONS

The trajectory of the buffer memory state vector is of central importance to this work. However, the system controls operate directly upon the clock frequencies. For this reason, it will be convenient to shift attention from the M equations (20) to an N-dimensional vector equation for the frequencies. This equation, and its component equations, will be called the "reduced" system equations because $N \leqq M$.

I shall proceed under the assumptions of no frequency noise and constant delays;

$$\boldsymbol{\eta}(t) = 0 \tag{26}$$

$$\boldsymbol{\tau}(t) = \boldsymbol{\tau}(0). \tag{27}$$

Then, (20) can be integrated to the form

$$x_{ji}(t) = D_{ji}^{-1}[p_i(t - \tau_{ji}) - p_j(t)] + B_{ji}, \tag{28}$$

where, as before, B_{ji} depends upon the initial condition and is defined by (24).

Now a simple change of variables is made.

$$q_i(t) = p_i(t) - ft. \tag{29}$$

In vector notation, this is,

$$\mathbf{q}(t) = \mathbf{p}(t) - ft\,\mathbf{1}_N, \tag{30}$$

where $\mathbf{1}_N$ is an N-dimensional column vector with unit coordinates.

If the system is asymptotically stable at the frequency f, then $\mathbf{q}'(t)$ will go asymptotically to zero. Therefore, given any controls, it will be of interest to see whether this condition is realized for any value of f in the allowed domain.

Substitution of (29) in (28) gives

$$x_{ji}(t) = D_{ji}^{-1}[q_i(t - \tau_{ji}) - q_j(t) - \tau_{ji}f] + B_{ji}. \tag{31}$$

In view of (2), (8), (26), and (30),

$$\mathbf{q}'(t) = (F - f)\mathbf{1}_N + \mathbf{E} + \mathbf{g}(t), \tag{32}$$

where $\mathbf{g}(t)$ is the increment to the clock frequencies under system control. In particular, let

$$\mathbf{g}(t) = \mathbf{\Gamma}\{t,\mathbf{x}(\cdot)\}. \tag{33}$$

Each component of $\mathbf{\Gamma}$ must have a realizable dependence upon the buffer memory state vector trajectory. The problem of control synthesis is precisely the problem of finding a suitable form for $\mathbf{\Gamma}$.

The reduced system equations are the differential-functional equations,

$$\mathbf{q}'(t) = (F - f)\mathbf{1}_N + \mathbf{E} + \mathbf{\Gamma}\{t,\mathbf{x}[\mathbf{q}(\cdot), f, \mathbf{B}]\}, \tag{34}$$

with $\mathbf{x}$ defined by (31). The parameter, $\mathbf{B}$, which depends upon the initialization of the system, enters (34) as a parameter in the controls. It is, therefore, not surprising that the system's trajectory and its final state, if any, depend upon its initialization. $\mathbf{B}$ is not entirely arbitrary. This can be seen by applying the condition for phase agreement, (12) at $t = 0$. Some manipulation of (16), (6), (12), (19), and (24) shows that

$$B_{ji} = D_{ji}^{-1}K_{ji} - 1, \tag{35}$$

where K_{ji} is an integer, and

$$K_{ji} = y_{ji}(0) + p_j(0) - p_i(-\tau_{ji}).$$

Given the initial phases, the initial condition can be changed only by integral changes in the numbers of cycles $\mathbf{y}(0)$ stored in the buffer memories.

IX. CONTROLS: QUALITATIVE DISCUSSION

Loosely speaking, it is desired to control the system so as to keep the buffer memory state vector small, in some appropriate sense. More specifically, the vector should be kept away from the faces of the unit cube.

These qualitative considerations will be made more concrete by defining a class of real valued functions $r(\mathbf{x})$ of the buffer memory state vector, which will be called "penalty functions". Each such function will have the following properties:

(*i*) $r(0) = 0$.
(*ii*) $r(-\mathbf{x}) = r(\mathbf{x})$.
(*iii*) $r(\mathbf{x})$ has a continuous gradient $\nabla r(\mathbf{x})$.
(*iv*) $r(\mathbf{x})$ is strictly convex; that is, for any two distinct vectors $\mathbf{x}_1$, $\mathbf{x}_2$ and any real number λ in the open interval $(0, 1)$,

$$r[\lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2] < \lambda r(\mathbf{x}_1) + (1 - \lambda) r(\mathbf{x}_2).$$

(*v*) For any $\mathbf{x}$ such that

$$|\mathbf{x}| = 1, \qquad \lim_{s \to \infty} \frac{s}{r(s\mathbf{x})} = 0.$$

These functions will have a unique minimum at the origin and will go to infinity uniformly on all rays from the origin. When properly chosen, their convex surfaces of constant value may closely approximate the cube surfaces having equal values of $\max_R |x_{ij}|$. This latter quantity, however, does not have a continuous gradient.

The attainable equilibrium points have been shown to lie in a neighborhood of an $(N - 1)$-dimensional linear manifold. The infimum of the values of $r(\mathbf{x})$ over this set is realized at a unique point in its closure. This point is either the origin or else the point of tangency with a surface of constant r. After selecting a suitable penalty function, controls will be sought which bring the buffer state vector near this point.

In attempting to reach this objective, a subclass of penalty functions having the simple form

$$r(\mathbf{x}) = \sum_R u(x_{ij}) \tag{36}$$

will be employed. The function $u(\cdot)$ must have the properties, (*i*) through (*v*), of a penalty function on a one-dimensional real space. A simple example of such a function is

$$u_n(x) = x^{2n}, \qquad n = 1, 2, \cdots .$$

Fig. 3 illustrates the surfaces,

$$\begin{aligned} r_\infty &\equiv \max(x_{12}, x_{21}) = 1 \\ r_1 &\equiv x_{12}^2 + x_{21}^2 = 1 \\ r_2 &\equiv x_{12}^4 + x_{21}^4 = 1 \\ r_3 &\equiv x_{12}^6 + x_{21}^6 = 1. \end{aligned}$$

The last three are penalty functions of the type defined in (36), for a system having just two links.

X. SYSTEM WITH ZERO DELAYS

The family of systems under consideration will have widely varying nonnegative delays for the transmission links. In many cases of interest, the product of maximum loop delay and control bandwidth may be very small compared to unity. In such cases, the extrapolation to zero delays

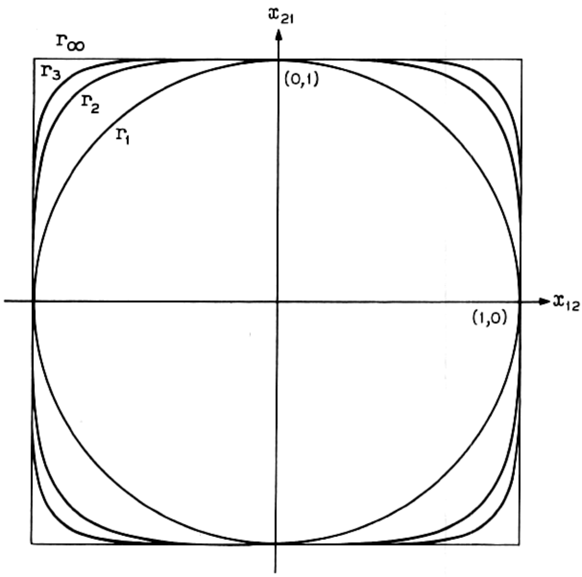

Fig. 3 — Curves of unit value.

may be a useful and illuminating exercise. One such system will be treated here as a step toward the synthesis of a family of controls.

In the zero delay case, let the controls $\mathbf{\Gamma}$ be simply a function of $\mathbf{x}$. This results in a system of ordinary, autonomous differential equations.

$$\mathbf{q}'(t) = (F - f)\mathbf{1}_N + \mathbf{E} + \mathbf{\Gamma}(\mathbf{x}) \tag{37}$$

in which $\mathbf{x}$ is now defined by

$$x_{ji}(t) = D_{ji}^{-1}[q_i(t) - q_j(t)] + B_{ji} . \tag{38}$$

The synthesis of $\mathbf{\Gamma}(\mathbf{x})$ will be based upon a penalty function which takes the form given in (36). One of the desired properties of the system (37) is that it come to rest near the attainable minimum of $r(\mathbf{x})$. It is therefore reasonable to try controls which have a component along the negative gradient of $r(\mathbf{x})$.

Let $\nabla_q r(\mathbf{x})$ be the column vector whose ith component is

$$[\nabla_q r(\mathbf{x})]_i = \frac{\partial}{\partial q_i} r(\mathbf{x}) \tag{39}$$

and let A be any $N \times N$ positive definite matrix. The controls to be considered here are of the form

$$\mathbf{\Gamma}(\mathbf{x}) = -A\nabla_q r(\mathbf{x}). \tag{40}$$

Thus, we are assured that*

$$[-\nabla_q r(\mathbf{x})]^T \cdot \mathbf{\Gamma}(\mathbf{x}) \geqq 0 \tag{41}$$

with equality if and only if $\nabla_q r(\mathbf{x}) = 0$.

Make the linear change of variables,

$$\mathbf{w}(t) = A^{-1}\mathbf{q}(t). \tag{42}$$

Then, the system equations are

$$\mathbf{w}'(t) = A^{-1}[(F - f)\mathbf{1}_N + \mathbf{E}] - \nabla_q r(\mathbf{x}). \tag{43}$$

The inverse of a positive definite matrix is also positive definite. Therefore,

$$\mathbf{1}_N^T A^{-1} \mathbf{1}_N > 0$$

and it is possible to choose f such that

$$\mathbf{1}_N^T A^{-1}[(F - f)\mathbf{1}_N + \mathbf{E}] = 0. \tag{44}$$

---

* The superscript T will indicate the transpose of a vector or matrix.

It is easy to verify that

$$\mathbf{1}_N{}^T \nabla_q r(\mathbf{x}) = 0 \tag{45}$$

in the zero delay system. Having chosen f to satisfy (44), we find that the trajectory of $\mathbf{w}$ must lie in a linear manifold which is orthogonal to $\mathbf{1}_N$. It will now be shown that the system exhibits global asymptotic stability when $\mathbf{w}$ is restricted to this linear manifold. That is, from any starting point in the manifold, the system will ultimately come to rest at a unique point in the manifold.

Consider the following function:

$$L(\mathbf{w}) = -\mathbf{w}^T A^T A^{-1}[(F - f)\mathbf{1}_N + \mathbf{E}] + r(\mathbf{x}). \tag{46}$$

Because of some of the properties of $r(\mathbf{x})$ and the connectedness of the systems under study, it can be shown that $L(\mathbf{w})$ has a unique minimum in every linear manifold orthogonal to $\mathbf{1}_N$. Also, $\nabla_w L(\mathbf{w})$ is zero only at this minimum. The proofs are given in Appendix B.

The time rate of change of $L(\mathbf{w})$ is

$$\dot{L}(\mathbf{w}) = [\mathbf{w}'(t)]^T \nabla_w L(\mathbf{w}).$$

But

$$\begin{aligned} \nabla_w L(\mathbf{w}) &= -A^T A^{-1}[(F - f)\mathbf{1}_N + \mathbf{E}] + A^T \nabla_q r(\mathbf{x}) \\ &= -A^T \mathbf{w}'(t). \end{aligned} \tag{47}$$

Therefore,

$$\dot{L}(\mathbf{w}) = -[\mathbf{w}'(t)]^T A^T \mathbf{w}'(t) = -[w'(t)]^T A \mathbf{w}'(t). \tag{48}$$

In view of the hypothesis that A is positive definite,

$$\dot{L}(\mathbf{w}) \leqq 0,$$

with equality if and only if $w'(t) = 0$. When this occurs,

$$-A^T \mathbf{w}'(t) = \nabla_w L(\mathbf{w}) = 0$$

and the system is at the minimum of $L(\mathbf{w})$. Thus, $L(\mathbf{w})$ is a Liapunov function for the system[15,16] and the system is globally asymptotically stable in the restricted sense mentioned above.

Inasmuch as

$$\mathbf{q}'(t) = A\mathbf{w}'(t),$$

$\mathbf{w}'(t) = 0$ implies $\mathbf{q}'(t) = 0$ and the system of (37) and (40) is also globally asymptotically stable in the linear manifold of its motion.

It is apparent from the definition (30) of $\mathbf{q}(t)$ that the system has the common frequency f when $\mathbf{q}'(t) = 0$. The Liapunov function

$$L(\mathbf{w}) = \hat{L}(\mathbf{q}) = -\mathbf{q}^T A^{-1}[(f - f)\mathbf{1}_N + \mathbf{E}] + r[\mathbf{x}(\mathbf{q})]$$

has its minimum at the equilibrium point, $\mathbf{q}_{\min}$, where

$$\nabla_q \hat{L}(\mathbf{q}_{\min}) = -A^{-1}[(F - f)\mathbf{1} + \mathbf{E}] + \nabla_q r[\mathbf{x}(\mathbf{q}_{\min})] = 0.$$

The equilibrium point is seen to be offset from that point at which $\nabla_q r = 0$ when the mistuning $\mathbf{E}$ of the clocks does not vanish. This should not be surprising, because the controls must compensate for the frequency differences.

Suppose the $(N - 1)$-dimensional subspace orthogonal to $\mathbf{1}_N$ is invariant under A^{-1}. Then the requirement (44) reduces to $f = F$, and the system frequency is the average of the clock center frequencies. It can be shown that when the matrix A has the above property, its column sums are all equal. In particular, if a diagonal matrix has this property, it is the identity matrix, multiplied by a positive scalar.

In this section, I have considered global asymptotic stability, rather than trajectories within the unit cube. Attainment of a suitably bounded trajectory will depend upon the A matrix and the initialization of the system.

XI. A FAMILY OF REALIZABLE ORGANIC SYSTEMS

In the last section, the controls

$$\mathbf{\Gamma}(x) = -A\nabla_q r(\mathbf{x}) = -A\nabla_q \sum_R u(x_{ji})$$

were shown to stabilize the system of (37) with zero delays. A family of controls will now be synthesized so as to be realizable and practical for systems having positive delay.

Equations (38) show that

$$\frac{\partial}{\partial q_k} u(x_{ji}) = D_{ji}^{-1}(\delta_{ik} - \delta_{jk})u'(x_{ji}) \tag{49}$$

using the Kronecker δ notation. Thus, when the matrix A is diagonal, the controls for the clock at center i depend only upon the buffer memory states in links terminating at center i or originating at center i. This is a very desirable simplification, and there seems to be no merit in employing more complicated forms. A more general type of control having this property is

$$\Gamma_k(\mathbf{x}) = \sum_{R_k} a_{kj}u'(x_{kj}) - \sum_{S_k} b_{jk}u'(x_{jk}), \tag{50}$$

where the signs have been chosen to agree with the earlier model when a_{kj}, b_{jk} are positive. The development is quite heuristic at this point, because a proof of stability for the system having the controls (50) with arbitrary positive coefficients is lacking.

The final model is based upon a modified form of the controls (50).

(*i*) To achieve realizability, a delay Δ_{kj} must be imposed on the argument of a control signal from the (j,k)th buffer memory to center k.

(*ii*) The controlled frequency deviations must be limited. For this purpose, I introduce the limiter function $\rho(\cdot)$ such that

$$\begin{aligned} \rho(x) &= x, & |x| &\leqq 1 \\ &= 1, & x &> 1 \\ &= -1, & x &< -1. \end{aligned}$$

(*iii*) For the purpose of reducing system bandwidth, a filter with impulse $h(t)$ and unit dc response may be employed. Let $*$ indicate convolution.

$$\begin{aligned} \Gamma_i\{t,\mathbf{x}(\cdot)\} = G\rho \Big(h(t) * \Big\{\sum_{R_i} a_{ij} u'[x_{ij}(t)] \\ - \sum_{S_i} b_{ji} u'[x_{ji}(t - \Delta_{ij})]\Big\}\Big). \end{aligned} \tag{51}$$

The complete system equations are

$$\mathbf{q}'(t) = (F - f)\mathbf{1}_N + \mathbf{E} + \mathbf{\Gamma}\{t,\mathbf{x}(\cdot)\} + \mathbf{n}(t) \tag{52}$$

$$x_{ji}(t) = D_{ji}^{-1}\{q_i[t - \tau_{ji}(t)] - q_j(t) - \tau_{ji}(t)f\} + B_{ji} \tag{53}$$

$$B_{ji} = x_{ji}(0) + D_{ji}^{-1}\{p_j(0) - p_i[-\tau_{ji}(0)]\}. \tag{54}$$

Equations (53) and (54) have been obtained by integration of (20).

The definition of the delay Δ_{ij} in (51) will depend upon the manner of transmission of the control signal. When the state x_{ji} is transmitted to center i via link (i,j),

$$\Delta_{ij} = \Delta_{ij}(t) = d_{ij}(t) + \tau_{ij}[t - d_{ij}(t)]. \tag{55}$$

This form is particularly awkward because the buffer memory delay, $d_{ij}(t)$, is not one of the canonical variables. However, it can be very closely approximated as follows:

$$d_{ij}(t) \cong D_{ij}[x_{ij}(t) + 1]F. \tag{56}$$

A simpler but cruder approximation is

$$d_{ij}(t) \cong D_{ij}F. \tag{57}$$

In the special case of (51), for which $b_{ji} = 0$, this complication does not arise.

XII. COMPARISON WITH THE SYSTEMS OF BENEŠ

V. E. Beneš[2] has analyzed a class of linear systems having delays and filters. In my notation, these systems obey the equations

$$p_i'(t) = F + E_i + Gh(t) * \sum_{R_i} \tilde{a}_{ij}[p_j(t - \tau_{ij}) - p_i(t)] + \eta_i(t). \tag{58}$$

He considers the delays to be fixed, the systems to be connected systems, and adds the constraints

$$\tilde{a}_{ij} > 0 \quad \text{for} \quad (i,j) \in R_i$$

$$\sum_{R_i} \tilde{a}_{ij} = 1.$$

Then, assuming the noise $\boldsymbol{\eta}(t)$ to be bounded and to go asymptotically to zero, he finds a sufficient condition for global asymptotic stability. This asymptotic stability is defined by

$$\lim_{t\to\infty} p_i'(t) = f, \qquad i = 1, 2, \cdots, N$$

$$\left| \lim_{t\to\infty} [p_i(t) - p_N(t)] \right| < \infty, \qquad i = 1, 2, \cdots, (N - 1).$$

Beneš sufficient condition is that

$$G > 0$$

$$\left| \frac{GH(i\omega)}{i\omega + GH(i\omega)} \right| < 1 \quad \text{for all } \omega \neq 0.$$

$H(s)$ is the Laplace transform of the filters' impulse response, $h(t)$; ω is real radian frequency, i is the imaginary unit, and it is assumed that $H(0) = 1$, as before.

This condition is stricter than that needed to stabilize an ordinary phase controlled oscillator, but it is not too difficult to satisfy. It is also quite remarkable in its independence of the system graph and its delays.

Beneš also gives formulas for the final system frequency and phase differences. These have been rederived more simply by Goldstein,[4] using the final value theorem.

A very direct approach to the final values is to insert them in (58), replacing the convolution with $Gh(t)$ by multiplication with G.

Let

$$\lim_{t\to\infty} p_i'(t) = f$$

$$\lim_{t\to\infty} [p_i(t) - p_N(t)] = \tilde{p}_i$$

and note that

$$\lim_{t\to\infty} [p_i(t - \tau_{ij}) - p_i(t)] = -\tau_{ij}f.$$

Making the appropriate changes in (58),

$$f = F + E_i + G \sum_{R_i} \tilde{a}_{ij}[\tilde{p}_j - \tilde{p}_i - \tau_{ij}f].$$

Because $\tilde{p}_N = 0$ by definition, we now have a set of N linear equations in the N variables f, $\tilde{p}_1, \cdots, \tilde{p}_{N-1}$. Using the notation,

$$\tau_i = \sum_{R_i} \tilde{a}_{ij}\tau_{ij},$$

for an average of the delays in the links to the ith center, the equations assume the simpler form,

$$(1 + G\tau_i)f + G\tilde{p}_i - G \sum_{R_i} \tilde{a}_{ij}\tilde{p}_j = F + E_i . \tag{59}$$

They have been shown by Goldstein to give the following solution for f

$$f = \frac{F + \sum b_i E_i}{1 + G \sum b_i \tau_i} . \tag{60}$$

Here b_i, $i = 1, 2, \cdots, N$, depends only upon the averaging coefficients, $\tilde{a}_{ij}$, $(i,j) \in R$, and

$$b_i \geqq 0$$

$$\Sigma b_i = 1.$$

A glance at (60) shows that the final system frequency is monotone decreasing with the product of the dc gain, G, and an average of all delays in the system. This effect has caused some dismay, but it results from an unrealistic model.

Let us go back to the family of organic systems defined by (51) through (54) and (29). These will be specialized in such a way as to obtain a class of linear systems analogous to that of Beneš. The following steps must be taken.

(*i*) Eliminate the limiter, $\rho(\cdot)$, from the controls, (51).

(*ii*) Let $u(x_{ij}) = \frac{1}{2}x_{ij}^2$, so $u'(x_{ij}) = x_{ij}$.

(*iii*) Use the phase variables,

$$p_i(t) = q_i(t) + ft.$$

(*iv*) Set $b_{ji} = 0$, for all (j,i), in (51).
(*v*) Make the identification

$$\tilde{a}_{ij} = a_{ij}D_{ij}^{-1}$$

and impose the constraint,

$$\sum_{R_i} \tilde{a}_{ij} = 1.$$

(*vi*) Assume the delays to be constant.
With these changes,

$$\begin{aligned} p_i'(t) = F + E_i + Gh(t) * \sum_{R_i} \tilde{a}_{ij}[p_j(t - \tau_{ij}) - p_i(t)] \\ + G \sum_{R_i} \tilde{a}_{ij}[D_{ij}x_{ij}(0) + p_i(0) - p_j(-\tau_{ij})] + \eta_i(t). \end{aligned} \tag{61}$$

This set of system equations differs from those of Beneš, (58), only in the addition of a term which is constant in time, but which depends upon the initial condition. It may be considered to be a modification of the mistuning, E_i, in the treatment of the stability problem. Therefore, the proof given by Beneš of global asymptotic stability under his sufficient condition also applies to (61).

Now let us derive the equations for final values. In doing this, note that the jth oscillator has the natural frequency, $F + E_j$, for $t < 0$, while its frequency has the final value f. Therefore,

$$p_j(-\tau_{ij}) = p_j(0) - \tau_{ij}(F + E_j)$$

$$\lim_{t\to\infty} [p_j(t - \tau_{ij}) - p_j(t)] = -\tau_{ij}f.$$

Now, proceeding as before, (61) leads to

$$\begin{aligned} f = F + E_i + G \sum_{R_i} \tilde{a}_{ij} \cdot (\tilde{p}_j - \tilde{p}_i - \tau_{ij}f) \\ + G \sum_{R_i} \tilde{a}_{ij}[D_{ij}x_{ij}(0) + p_i(0) - p_j(0) + \tau_{ij} \cdot (F + E_j)] \end{aligned}$$

Putting this in a form analogous to (59),

$$\begin{aligned} (1 + G\tau_i)f + G\tilde{p}_i - G \sum_{R_i} \tilde{a}_{ij}\tilde{p}_j = (1 + G\tau_i)F + E_i \\ + G \sum_{R_i} \tilde{a}_{ij}[\tau_{ij}E_j + D_{ij}x_{ij}(0) + p_i(0) - p_j(0)]. \end{aligned} \tag{62}$$

The solution for f, analogous to (60), is

$$f = F + \sum_R c_{ij}E_j + \frac{G \sum_R b_i\tilde{a}_{ij}D_{ij}x_{ij}(0)}{1 + G \sum b_i\tau_i} \tag{63}$$

where the coefficients $\mathbf{c}$ are averaging coefficients defined by

$$c_{ij} = \frac{b_j\tilde{a}_{ji} + Gb_i\tilde{a}_{ij}\tau_{ij}}{\sum_R (b_j\tilde{a}_{ji} + Gb_i\tilde{a}_{ij}\tau_{ij})} \tag{64}$$

and the coefficients $\mathbf{b}$ are the same as before.

The first two terms on the right of (63) give an average of the individual clock frequencies. The last term depends on the initial condition, but it goes to zero as the system delays become large. Thus, the behavior indicated by (60) does not really occur in our model for organic systems.

XIII. SOME REMARKS ON THE STABILITY PROBLEM

The mathematical problem of stability is not yet satisfactorily solved for general organic systems. Two special families of organic systems are now known to be globally asymptotically stable. These are certain nonlinear systems with zero delays and certain linear systems with delays and filters. These very special cases nourish the hope that broader sufficient conditions for stability can be found.

The present section will be devoted to redefinition of the stability problem and a discussion of some necessary conditions for stability.

We have seen that the system will malfunction whenever the M-dimensional buffer state vector $\mathbf{x}(t)$ leaves the unit cube. This leads to the following practical definition of stability.

Definition: For any positive ε, trajectory $\mathbf{x}(t)$ is ε-stable if

$$\max_i | x_i(t) | \leqq \varepsilon \quad \text{for} \quad 0 \leqq t < \infty.$$

Definition: A trajectory is stable if there is an $\varepsilon < 1$ for which it is ε-stable.

The trajectory of an undisturbed organic system will depend upon the system parameters and the initial condition. Therefore, the domain of system stability must be defined in a space having the following coordinates, which appear in (51) through (54). G, $h(\cdot)$, $\mathbf{a}$, $\mathbf{b}$, $u(\cdot)$, $\mathbf{\Delta}$, $\mathbf{D}$, $\mathbf{x}(0)$, $\mathbf{E}$, $\boldsymbol{\tau}$ and $\mathbf{p}(t)$ for $t \leqq 0$. The initialization of the filter states must also be defined.

Several necessary conditions should be kept in mind when testing for stability. First, the asymptotic frequency of the system must lie within the controllable range,

$$F + E_i - G \leqq f \leqq F + E_i + G, \qquad i = 1, 2, \cdots, N.$$

Second, the limit points of the buffer state trajectory must lie within the unit cube. Third, the system must be "connected," in some sense.

The connectedness of the systems deserves further discussion. I have required that the organic systems must be connected in the following sense: that there must exist a directed transmission path from each center to every other center. Beneš has used the same condition. On the other hand, the stability of the special nonlinear systems which I have treated depends upon a weaker condition. Namely, that the nonoriented graph having a branch corresponding to each link must be connected. This condition is used in Appendix B, which is essential to the proof.

An important difference between the systems treated here and those of Beneš, is that the former have the additional control coefficients, $b_{ij} \geqq 0$ for $(i,j) \in R$. Thus, the state $x_{ij}(t)$ of a particular buffer may exert a control over the frequency of the sending center, j, as well as over the receiving center, i.

Intuition suggests a necessary condition for stability based upon the control coefficients, a_{ij}, b_{ij}, which appear in the family of equations, (51). Inasmuch as negative coefficients tend to make the systems unstable, these are assumed to be either positive or zero.

Consider a "control graph" with nodes numbered 1, 2, $\cdots$, N. Let a directed branch exist from node j to node i if and only if $a_{ij} + b_{ji} > 0$. This condition permits the frequency at center i to be influenced by its phase relative to that at center j. Then a necessary condition for system stability is that the control graph shall have a node from which directed paths exist to all other nodes.

Under this "weak" condition, some parts of the system may simply be "slaves" of another part of the system. The "strong" condition that there exist a directed path from each node to each other node precludes this possibility. However, it should be understood that the condition satisfied by the control graph need not be satisfied by the graph whose branches correspond to transmission links. This is the case because each link may give rise to two oppositely directed branches of the control graph. On the other hand, a connected system may lack stability when too few of the control coefficients are positive.

A simple example is provided by the system shown in Fig. 4. The digital transmission links appear in the "system graph" Fig. 4(a). When

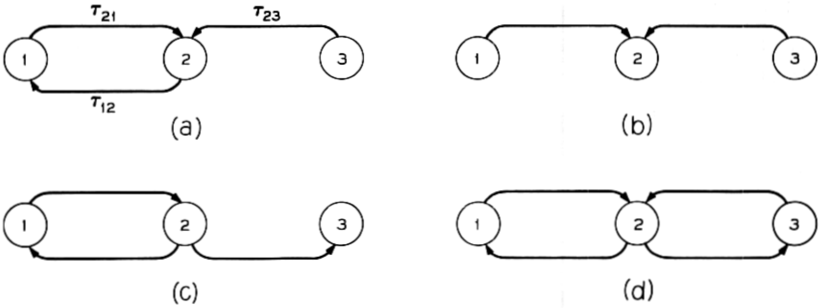

Fig. 4 — A system graph and three control graphs. (a) system graph; (b) inadequately connected control graph; (c) weakly connected control graph; (d) strongly connected control graph.

the only positive coefficients are a_{21}, a_{23}, the control graph of Fig. 4(b) results. This fails to satisfy even the weak condition. It is clear that no means exists for bringing centers 1 and 3 to a common frequency. The control graph in Fig. 4(c) results when a_{12}, a_{21}, b_{23} are the positive coefficients. It must be assumed that a *separate* means for transmitting the narrow band signal, b_{23}, to center 3 exists. This control graph satisfies the weak condition. If the system is stable, center 3 will be a slave to the rest of the system. It will have no influence upon the frequency trajectories of centers 1 and 2. If a_{12}, a_{21}, a_{23}, b_{23} are the positive coefficients, then the control graph of Fig. 4(d) results. This one obeys the strong condition. When the control graph is the same as the system graph, Fig. 4(a), it is weakly connected. In this case, if the system is stable, center 3 determines the common frequency.

XIV. SUMMARY

A class of systems for the mutual synchronization of spatially separated oscillators has been synthesized and a mathematical model for these systems has been presented. The model may be said to be physically realizable in that real systems can be built whose function will very closely approximate the behavior of the model. While no such system hardware has been presented here, a simple hardware analog has been built.[9]

These systems, called "organic synchronization systems," have a possible application to continental or worldwide PCM communications.

A re-examination of the systems treated by Beneš in the light of the newly derived organic model indicates that

(*i*) his stability proof does apply to a particular class of linearized organic systems, and (*ii*) that his formula for the final system frequency must be modified; the suitably modified formula no longer displays a monotone decreasing final frequency with increasing system delays.

APPENDIX A

Asymptotic Phase Differences

The set of all differences of the form, $(\tilde{q}_i - \tilde{q}_j)$, can have at most $(N - 1)$ linearly independent elements. To verify this, consider the $(N - 1)$ elements

$$(\tilde{q}_N - \tilde{q}_{N-1}),\ (\tilde{q}_{N-1} - \tilde{q}_{N-2}),\ \cdots,\ (\tilde{q}_2 - \tilde{q}_1).$$

Suppose $i > j$. Then

$$(\tilde{q}_i - \tilde{q}_j) = (\tilde{q}_i - \tilde{q}_{i-1}) + (\tilde{q}_{i-1} - \tilde{q}_{i-2}) + \cdots + (\tilde{q}_{j+1} - \tilde{q}_j)$$

while

$$(\tilde{q}_j - \tilde{q}_i) = -(\tilde{q}_i - \tilde{q}_j).$$

Hence, any other element of the set of differences can be represented as a linear combination of the selected $(N - 1)$ elements.

When the directed graph, which corresponds to the synchronizing network, is connected the set of differences

$$\{(\tilde{q}_i - \tilde{q}_j) \mid (j,i) \in R\}$$

has at least $(N - 1)$ linearly independent elements. Actually, only the weak, i.e., nonoriented, sense of connectedness of necessary for the proof.

Theorem: Let G be a directed graph with N vertices such that the corresponding nonoriented graph is connected. Associate the N independent real variables, $\tilde{q}_1, \tilde{q}_2, \cdots, \tilde{q}_N$, one to one with the correspondingly indexed vertices. Associate the difference $(\tilde{q}_i - \tilde{q}_j)$ with the edge from vertex i to vertex j, for each edge in G. Let T be any complete tree of G.

Then, there is a set of $(N - 1)$ linearly independent differences associated with edges of T.

Proof: We shall proceed by induction. The theorem clearly is true when $N = 2$.

Suppose the theorem to be true for $N = L - 1$. Now consider G to have L vertices. Since it is connected, it contains a complete tree T, which will have $L - 1$ edges, but all L vertices. It follows that not all vertices can have more than one edge of T incident on them. Let vertex j be an end vertex. Then, only one of the differences associated with the edges of τ contains $\tilde{q}_j$. It follows that this difference, say $\pm(\tilde{q}_i - \tilde{q}_j)$, is linearly independent of the remaining $(L - 2)$ differences. But if vertex j and the edge incident on it are removed from T, a tree having $(L - 1)$ vertices remains. By hypothesis, this contains $(L - 2)$ edges whose associated differences are linearly independent. Hence, there are $(L - 1)$ linearly independent differences associated with the edges of T, and the proof is complete.

Corollary: The differences are unchanged when the average of the $\tilde{q}$'s is subtracted from each of them. Therefore, the theorem applies even when the N real variables are constrained to have zero sum.

APPENDIX B

Properties of $L(\mathbf{w})$

Certain preliminaries concerning convex functions will be necessary here. While they are familiar to mathematicians, other readers may find the following review helpful.

All sets to be considered here are subsets of finite dimensional real linear spaces. All functions will be defined on such sets.

Definition 1: A set of points X is convex if every point of the set

$$\{\lambda \mathbf{x}_2 + (1 - \lambda)\mathbf{x}_1 : \mathbf{x}_1, \quad \mathbf{x}_2 \in X \quad \text{and} \quad 0 \leqq \lambda \leqq 1\}$$

is also a point of X.

Definition 2: A real valued function $f(\cdot)$ defined on a convex set X is a convex function on X if

$$f[\lambda \mathbf{x}_2 + (1 - \lambda)\mathbf{x}_1] \leqq \lambda f(\mathbf{x}_2) + (1 - \lambda)f(\mathbf{x}_1) \tag{65}$$

whenever $\mathbf{x}_1$, $\mathbf{x}_2 \in X$ and $0 \leqq \lambda \leqq 1$.

A convex function $f(\cdot)$ is *strictly* convex if the equality in (65) implies that $\lambda = 0$ or $\lambda = 1$ or $\mathbf{x}_1 = \mathbf{x}_2$.

Theorem 1: If two convex functions are defined on the same convex set, their sum is a convex function on that set. If one of the functions is strictly convex, then the sum is strictly convex.

Theorem 2: If $f(\cdot)$ is a convex function on a convex set X, and if $\alpha \geqq 0$, then $\alpha f(\cdot)$ is a convex function on X. If $f(\cdot)$ is strictly convex, and if $\alpha > 0$, then $\alpha f(\cdot)$ is strictly convex.

Theorem 3: If $\mathbf{c}$ is a fixed vector in an n-dimensional space, and $\mathbf{x}$ is a *variable vector in the same space, and d is a real number, then* $(\mathbf{c}^T \cdot \mathbf{x} + d)$ *is a convex function on the entire space.*

Theorem 4: Let $f(\cdot)$ *be a convex function on an n-dimensional space. Then the sets of points*

$$\{\mathbf{x}: f(\mathbf{x}) \leqq d\}, \qquad \{\mathbf{x}: f(\mathbf{x}) < d\}$$

are convex subsets of the space for every real number, d.

The above theorems are elementary consequences of the definitions, 1 and 2.

Theorem 5: Let $f(\cdot)$ *be a differentiable convex function defined on a convex set X. Let* $\mathbf{x}_1$, $\mathbf{x}_2$ *be distinct points of X. Then the directional derivative,* $(d/d\lambda) f[\lambda \mathbf{x}_2 + (1 - \lambda)\mathbf{x}_1]$, *is an increasing function of* λ *in* $(0,1)$.

Proof: Select any $\lambda_0 > 0$, $\delta > 0$ such that $\lambda_0 + \delta < 1$. Let

$$\mathbf{x}_0 = \lambda_0 \mathbf{x}_2 + (1 - \lambda_0)\mathbf{x}_1, \qquad \mathbf{y} = \mathbf{x}_2 - \mathbf{x}_1.$$

Then

$$(\lambda_0 + \delta)\mathbf{x}_2 + (1 - \lambda_0 - \delta)\mathbf{x}_1 = \mathbf{x}_0 + \delta \mathbf{y}.$$

Now select an ε such that $\varepsilon > 0$ and $\lambda_0 + \delta + \varepsilon < 1$ and apply (65) twice as follows:

$$f(\mathbf{x}_0 + \delta \mathbf{y}) = f\left[\frac{\delta}{\delta + \varepsilon}(\mathbf{x}_0 + \delta \mathbf{y} + \varepsilon \mathbf{y}) + \frac{\varepsilon}{\delta + \varepsilon}\mathbf{x}_0\right]$$

$$\leqq \frac{\delta}{\delta + \varepsilon} f(\mathbf{x}_0 + \delta \mathbf{y} + \varepsilon \mathbf{y}) + \frac{\varepsilon}{\delta + \varepsilon} f(\mathbf{x}_0)$$

$$f(\mathbf{x}_0 + \varepsilon \mathbf{y}) \leqq \frac{\varepsilon}{\delta + \varepsilon} f(\mathbf{x}_0 + \delta \mathbf{y} + \varepsilon \mathbf{y}) + \frac{\delta}{\delta + \varepsilon} f(\mathbf{x}_0).$$

Adding these inequalities and rearranging terms,

$$f(\mathbf{x}_0 + \delta \mathbf{y} + \varepsilon \mathbf{y}) - f(\mathbf{x}_0 + \delta \mathbf{y}) \geqq f(\mathbf{x}_0 + \varepsilon \mathbf{y}) - f(\mathbf{x}_0).$$

Dividing both members by ε and taking the limits as $\varepsilon \to 0$ yields the desired result,

$$\frac{d}{d\lambda} f[\lambda \mathbf{x}_2 + (1 - \lambda)\mathbf{x}_1]\bigg|_{\lambda_0 + \delta} \geqq \frac{d}{d\lambda} f[\lambda \mathbf{x}_2 + (1 - \lambda)\mathbf{x}_1]\bigg|_{\lambda_0}.$$

Corollary: If $f(\cdot)$ *is strictly convex on the convex set X, and if* $\mathbf{x}_1$, $\mathbf{x}_2$ *are distinct points of X, then* $(d/d\lambda) f[\lambda \mathbf{x}_2 + (1 - \lambda)\mathbf{x}_1]$ *is a strictly increasing function of* λ *in* $(0,1)$.

Proof: Suppose that the equality holds in Theorem 5. If the directional derivative is nondecreasing in $[\lambda_0, (\lambda_0 + \delta)]$ and takes equal values at the end points, then it must be constant on this interval. It follows that $f(\cdot)$ varies linearly on the line segment from

$$\mathbf{x} = \lambda_0\mathbf{x}_2 + (1 - \lambda_0)\mathbf{x}_1 \quad \text{to} \quad \mathbf{x} = (\lambda_0 + \delta)\mathbf{x}_2 + (1 - \lambda_0 - \delta)\mathbf{x}_1 .$$

This contradicts the hypothesis that $f(\cdot)$ is strictly convex on X, and the proof is complete.

Theorem 6: Let $f(\cdot)$ be a strictly convex differentiable function defined on a convex set D. Let C be a closed and bounded convex set in the interior of D. Then,

(*i*) *$f(\cdot)$ assumes its minimum value over C at a unique point of C,*
(*ii*) *$f(\cdot)$ has a vanishing gradient at no more than one point of C, and*
(*iii*) *$f(\cdot)$ assumes its minimum value over C at an interior point of C if and only if the gradient of $f(\cdot)$ vanishes at that point.*

Proof:

(*i*) The hypotheses imply that $f(\cdot)$ is a continuous function and that C is a compact set. It follows that $f(\cdot)$ assumes its minimum value over C, $f_{\min}$, at some point of C. Now suppose that $\mathbf{x}_1$, $\mathbf{x}_2$ are distinct points of C such that

$$f(\mathbf{x}_1) = f(\mathbf{x}_2) = f_{\min} .$$

Then the strict convexity of $f(\cdot)$ implies that

$$f\left(\frac{\mathbf{x}_1}{2} + \frac{\mathbf{x}_2}{2}\right) < f_{\min}.$$

This is a contradiction of the hypothesis that $f_{\min}$ is the least value of $f(\cdot)$ over C.

(*ii*) Suppose the gradient of f vanishes at two distinct points of C, $\mathbf{x}_1$ and $\mathbf{x}_2$. Consider the directional derivative of $f(\cdot)$ along $(\mathbf{x}_2 - \mathbf{x}_1)$. By hypothesis, this derivative vanishes at $\mathbf{x}_1$. Because $f(\cdot)$ is strictly convex, it is a strictly increasing function of position along $(\mathbf{x}_2 - \mathbf{x}_1)$. Therefore, it is greater than zero at $\mathbf{x}_2$, which contradicts the hypothesis that the gradient vanishes at $\mathbf{x}_2$.

(*iii*) Suppose $f(\mathbf{x}_0) = f_{\min}$ and $\mathbf{x}_0$ is an interior point of C. All points in a neighborhood of $\mathbf{x}_0$ are in C. If the gradient of $f(\cdot)$ does not vanish at $\mathbf{x}_0$, then there are points in this neighborhood, along the negative gradient from $\mathbf{x}_0$, at which $f(\cdot)$ assumes smaller values. This contradicts the hypothesis that $f_{\min}$ is the least value of $f(\cdot)$ in C.

Now suppose that the gradient of $f(\cdot)$ vanishes at $\mathbf{x}_0$, an interior point of C. Let $\mathbf{x}_1$ be any other point of C and consider the directional derivative of $f(\cdot)$ along $(\mathbf{x}_1 - \mathbf{x}_0)$. This vanishes at $\mathbf{x}_0$ by hypothesis. The strict convexity of $f(\cdot)$ implies that it is strictly increasing from $\mathbf{x}_0$ to $\mathbf{x}_1$. Therefore, $f(\mathbf{x}_1) > f(\mathbf{x}_0)$ and $f(\cdot)$ has its minimum over C at $\mathbf{x}_0$.

This completes the proof.

In considering the properties of $L(\mathbf{w})$, an additional result will be needed concerning the penalty function $r(\cdot)$. Its properties are numbered (i) through (v). Note that the convergence in (v) was not assumed to be uniform over the unit sphere. This property will now be deduced.

Because $r(\cdot)$ is a strictly convex function on the M-dimensional vector space, it is also strictly convex on any subspace. In particular, $r(s\mathbf{x}_1)$ is a strictly convex function of s for any $\mathbf{x}_1$ on the unit sphere.

Then, for any real number P, however large,

$$e(s,\mathbf{x}_1) \equiv r(s\mathbf{x}_1) - sP$$

is a strictly convex function of s. Let E_s be the set of vectors, $\mathbf{x}$, on the unit sphere for which

$$e(s,\mathbf{x}) \leqq 0.$$

For any fixed s, E_s is a closed subset of the unit sphere because $e(s,\mathbf{x})$ is continuous. It follows that E_s is a compact set.

The corollary to Theorem 5 tells us that along any ray from the origin, i.e., for s going from zero to infinity, the derivative of $e(s,\mathbf{x}_1)$ is strictly monotone increasing. From this it can be seen that $\mathbf{x}_1 \notin E_{s_0}$ implies that $\mathbf{x}_1 \notin E_s$ for $s > s_0$. Therefore, the sets E_s decrease as s increases from zero to infinity.

The intersection of a class of compact, decreasing, nonempty sets is nonempty. Therefore, if

$$\bigcap_{s=0}^{\infty} E_s = 0,$$

it is clear that there exists an s_0 such that $E_s = 0$ for $s \geqq s_0$. In this case,

$$r(s\mathbf{x}) > sP \quad \text{for all} \quad s \geqq s_0(P)$$

independent of $\mathbf{x}$ on the unit sphere.

On the other hand, if

$$\bigcap_{s=0}^{\infty} E_s \neq 0,$$

then there exists a unit vector $\mathbf{x}_1$, such that

$$r(s\mathbf{x}_1) \leqq sP \quad \text{for all} \quad s \text{ in } (0, \infty).$$

This contradicts hypothesis (v) concerning the penalty functions.

Thus, we have seen that $r(s\mathbf{x})/s$ goes to infinity with s uniformly for all $\mathbf{x}$ on the unit sphere.

Now consider the function $L(\mathbf{w})$, defined by (46), with $\mathbf{w}$ restricted to an $(N - 1)$-dimensional linear manifold orthogonal to $\mathbf{1}_N$. Let

$$\mathbf{w} = \mathbf{v} + \alpha \mathbf{1}_N$$

where α is a real number and $\mathbf{v}$ is restricted to the $(N - 1)$-dimensional linear subspace orthogonal to $\mathbf{1}_N$.

In view of (38) and (42), $L(\mathbf{w})$ can be put in the form,

$$L(\mathbf{w}) = \tilde{L}(\mathbf{v}) = \mathbf{c}^T \cdot \mathbf{v} + d + r(KA\mathbf{v} + \mathbf{b}). \tag{66}$$

Here, $\mathbf{c}$ is a fixed N-dimensional vector, d is a scalar constant, and $\mathbf{b}$ is a fixed M-dimensional vector, with $M \geqq N$. The fixed $N \times N$ matrix A is positive definite. It can be seen from the discussion in Appendix A and (38) that the fixed $M \times N$ matrix K is of rank $(N - 1)$ for connected systems. Its null space is spanned by $\mathbf{1}_N$.

Using the convexity of $r(\cdot)$,

$$r(\tfrac{1}{2}KA\mathbf{v}) \leqq \tfrac{1}{2}r(-\mathbf{b}) + \tfrac{1}{2}r(KA\mathbf{v} + \mathbf{b})$$

$$r(KA\mathbf{v} + \mathbf{b}) \geqq 2r(\tfrac{1}{2}KA\mathbf{v}) - r(-\mathbf{b}).$$

Using this in (66),

$$L(\mathbf{w}) = \tilde{L}(\mathbf{v}) \geqq 2r(\tfrac{1}{2}KA\mathbf{v}) + \mathbf{c}^T \cdot \mathbf{v} + d - r(-\mathbf{b}).$$

The right-hand member is dominated by its first term as $|\mathbf{v}|$ becomes large, uniformly over the subspace orthogonal to $\mathbf{1}_N$. Therefore, we can find a sphere of sufficiently large radius so that

$$\tilde{L}(\mathbf{v}) > \tilde{L}(0)$$

for all $\mathbf{v}$ on its surface. Then the minimum value of $\tilde{L}(\mathbf{v})$ over this sphere is not assumed on the boundary.

Now it will be shown that $\tilde{L}(\mathbf{v})$ is a strictly convex function of $\mathbf{v}$ on the subspace orthogonal to $\mathbf{1}_N$. Let

$$\mathbf{x}_1 = KA\mathbf{v}_1 + \mathbf{b}$$

$$\mathbf{x}_2 = KA\mathbf{v}_2 + \mathbf{b}.$$

The properties of K and A are such that $\mathbf{v}_1 \neq \mathbf{v}_2$ implies $\mathbf{x}_1 \neq \mathbf{x}_2$. This permits us to apply the strict convexity condition,

$$r[KA\lambda\mathbf{v}_2 + KA(1-\lambda)\mathbf{v}_1 + \mathbf{b}]$$
$$= r[\lambda\mathbf{x}_2 + (1-\lambda)\mathbf{x}_1] < \lambda r(\mathbf{x}_2) + (1-\lambda)r(\mathbf{x}_1)$$
$$= \lambda r(KA\mathbf{v}_2 + \mathbf{b}) + (1-\lambda)r(KA\mathbf{v}_1 + \mathbf{b}).$$

This is sufficient to establish the strict convexity of $\tilde{L}(\mathbf{v})$.

Inasmuch as the strictly convex function $\tilde{L}(\mathbf{v})$ takes on its minimum value over every large sphere at an interior point, its gradient vanishes uniquely at that point.

The above statement applies to the restriction of $\tilde{L}(\mathbf{v})$ to an $(N-1)$-dimensional space. However, we know that $\mathbf{w}'(t)$ vanishes along $\mathbf{1}_N$. Equation (47) then shows that

$$\nabla_w L(\mathbf{w}) = -A^T\mathbf{w}'(t)$$

vanishes along a direction which is not orthogonal to $\mathbf{1}_N$. It follows that the unrestricted gradient of $L(\mathbf{w})$ vanishes at a unique point of every linear manifold orthogonal to $\mathbf{1}_N$.

REFERENCES

1. Vaughan, H. E., Research Model for Time-Separation Integrated Communication, B.S.T.J., *38*, July, 1959, pp. 909–932.
2. Beneš, V. E., unpublished work.
3. Runyon, J. P., Reciprocal Timing of Time Division Switching Centers, U. S. Patent No. 3,050,586, August 21, 1962.
4. Goldstein, A. J., unpublished work.
5. Beneš, V. E. and Goldstein, A. J., unpublished work.
6. Darwin, G. P. and Prim, R. C., Synchronization in a System of Interconnected Units, U. S. Patent No. 2,986,723, May, 1961.
7. Karafin, B. J., unpublished work.
8. Gersho, A. and Karafin, B. J., Mutual Synchronization of Geographically Separated Oscillators, B.S.T.J., this issue, pp. 1689–1704.
9. Candy, C. J. N. and Karnaugh, M., to be published.
10. Brilliant, M. B., The Determination of Frequency in Systems of Mutually Synchronized Oscillators, B.S.T.J., this issue, pp. 1737–1748.
11. Brilliant, M. B., Dynamic Response of Systems of Mutually Synchronized Oscillators, B.S.T.J., to be published.
12. Saito, T., Fujisaki, H., and Inose, H., unpublished work.
13. Byrne, C. J., Karnaugh, M., and Scattaglia, J. V., Retiming of Digital Signals with a Local Clock, 1960 Northeast Electronic Res. and Eng. Meeting, NEREM Record, November, 1960.
14. Byrne, C. J., Karafin, B. J., and Robinson, D. B., Jr., Systematic Jitter in a Chain of Digital Regenerators. B.S.T.J., *42,* November, 1963, pp. 2679–2714.
15. LaSalle, J. and Lefschetz, S., *Stability by Liapunov's Direct Method,* Academic Press, New York, 1961.
16. Hahn, W., *Theory and Application of Liapunov's Direct Method,* Prentice-Hall, Englewood Cliffs, New Jersey, 1963.