The Determination of Frequency in Systems of Mutually Synchronized Oscillators

By M. B. BRILLIANT

(Manuscript received September 7, 1966)

The synchronization of large systems of geographically separated oscillators is of considerable practical interest for pulse code modulation (PCM) switching. This study examines the factors that determine the frequency at which such a system operates, considering both the procedure by which it is set up and the topology of system interconnections. A necessary and sufficient connectivity condition is established.

I. INTRODUCTION

The synchronization of large systems of geographically separated oscillators is of considerable practical interest for pulse code modulation (PCM) switching. Synchronization could be achieved by establishing a single master oscillator, with every other cscillator slaved either directly to the master or to another oscillator that is slaved directly or indirectly to the master. However, the system would then be vulnerable to failure of a single link or a single oscillator. An alternative called "mutual synchronization" would permit the oscillators to determine the system frequency jointly and to exchange synchronization information over redundant paths. However, the complexity of the system raises questions concerning the factors that determine the system frequency as well as system stability and dynamic response.

A broad sufficient condition for the stability of mutually synchronized systems was first established by Beneš.[1] This condition has recently been rederived by a different method, for a slightly more general system, by Gersho and Karafin.[2] The model used in both these studies was oversimplified so that it gave a paradoxical result for the system frequency at equilibrium. A model that corrected this oversimplification, by considering the received signal phases observed at each oscillator station at the initial moment when all oscillator controls are put into operation, was first devised by Runyon.[3] A corrected model based on the same principle, but differing in detail, has recently been independently derived by Karnaugh.[4] In all these studies, it was assumed that the system was so interconnected that every oscillator transmitted timing information either directly or indirectly to every other oscillator. This condition has been generally assumed to be necessary for mutual synchronization, and has been proved sufficient by Gersho and Karafin.[2]

This paper will generalize the foregoing results in two ways that appear to be significant for practical applications. In the first place, a mathematical model will be described that allows the synchronized system to be set up by less drastic methods than the simultaneous closure of all control paths at $t = 0$. In the second place, a weaker connectivity condition, which is satisfied by systems in which only some of the oscillators participate in frequency determination, will be proved necessary and sufficient for synchronization.

The practical consequence of these generalizations is that a system with a single master oscillator can be regarded as a special case within the general class of mutually synchronized systems, and a locked oscillator synchronized to a remote source can be regarded as a special case of an oscillator station in a mutual synchronization system. Between the extremes of a system with no slaves and a system in which all stations but one are slaves, a variety of hierarchical organizations may be envisioned. However, the description of particular configurations is beyond the scope of this article. The model developed here also provides the flexibility by which new stations can be added to an existing system, and the system frequency can be adjusted after synchronism has been established.

II. THE MATHEMATICAL MODEL

The system is assumed to consist of N oscillators, or "clocks," numbered $i = 1, \cdots, N$. Each oscillator has its own free-running frequency f_i, at which it would operate in the absence of a control input. Each oscillator accepts a control input that causes its frequency to deviate from the free-running frequency by an amount proportional to the control input. For concreteness, the control input will be referred to as a voltage, although it may, in practice, take other forms. Thus, the instantaneous frequency of the ith oscillator, which will be expressed simply as the rate of change of phase $p_i'(t)$, will, in general, be different from the free-running frequency f_i.

In applications to switched PCM networks, each oscillator controls the timing of a digital signal which is assumed to be organized with a fixed number of pulses per frame. It will be convenient to measure phase in frames of the digital signal, and frequency in frames per second. Each station sends a digital signal, controlled by its local oscillator, to a number of other stations, and this signal conveys timing information. To simplify the description, it will be assumed that all these signals are sent in the same phase, and this will be taken to define the phase of the local clock. However, the model could easily be adapted to the case in which each signal is sent in some arbitrary but fixed phase with respect to the local clock.

The transmission delay from the jth station to the ith will be designated as τ_{ij}. Thus, the phase of the signal received at the ith station from the jth is $p_j(t - \tau_{ij})$. The phase is defined principally by a regular pattern of framing pulses. The pulses between the framing pulses carry information, and are therefore different in successive frames. Since successive frames are distinguishable, the cyclic ambiguity inherent in the measurement of the phase of sinusoidal signals is not inherent in the digital case.

Thus it is possible to measure, at the ith station, the phase difference $p_j(t - \tau_{ij}) - p_i(t)$ between the received signal from the jth station and the local clock. This phase difference will be called the "observed phase" of the jth signal at the ith station. In the Beneš[1] model, used also by Gersho and Karafin,[2] the control voltage at the oscillator consists only of components proportional to the observed phases. However, as Gersho and Karafin[2] pointed out, if all the clocks are in phase all the observed phases will be negative, and every clock will be made to run slower than its free-running frequency. In the present model, a fixed reference phase r_{ij} will be subtracted from each observed phase, this reference phase preferably being equal to the phase difference one would expect to observe. If the observed phase of each signal is equal to the reference phase, no control voltage is applied to the oscillator, which then runs at its free-running frequency.

Historically,[1] the concept of mutual synchronization evolved in terms of phase averaging. Thus, the observed phases of the received signals were respectively multiplied by nonnegative averaging coefficients a_{ij},

$$\sum_{j=1}^{N} a_{ij} = 1, \tag{1}$$

to form an average phase difference between the local clock and the signals received from its neighbors. The average observed phase may then be multiplied by a nonnegative factor λ_i, having the dimensions of inverse time, to determine the frequency displacement of the local clock. This basic notation has been continued in subsequent studies and will be used here. Thus, in the present model, the system equations must be

$$p_i'(t) = f_i + \lambda_i \sum_{j=1}^{N} a_{ij}[p_j(t - \tau_{ij}) - p_i(t) - r_{ij}], \qquad i = 1, \cdots, N. \tag{2}$$

The reference phases can be absorbed into the free-running frequency term by defining a reference frequency

$$v_i = f_i - \lambda_i \sum_{j=1}^{N} a_{ij} r_{ij}, \qquad i = 1, \cdots, N. \tag{3}$$

The system equations can now be written as

$$p_i'(t) = v_i + \lambda_i \sum_{j=1}^{N} a_{ij}[p_j(t - \tau_{ij}) - p_i(t)], \qquad i = 1, \cdots, N. \tag{4}$$

These equations have formally reverted to those of the Beneš model,[1,2] in which the reference phases do not appear. However, while the equations are the same, their application is different, since v_i is not the free-running frequency, but is normally greater than the free-running frequency, because the reference phases r_{ij} are normally negative. The reference phases may, in fact, be identified with the initial-condition terms of Runyon[3] and Karnaugh,[4] so that (4) covers their models as well as the Beneš model.

The dynamic response of the system can be modified by using a filter in each control system. Multiplication by λ_i is then replaced by convolution with the impulse response $h_i(t)$ of a filter whose zero-frequency gain is

$$\int_0^\infty h_i(t)\,dt = \lambda_i, \qquad i = 1, \cdots, N. \tag{5}$$

This has been done in all the referenced studies. Gersho and Karafin[2] also added a variable term to v_i, replacing it formally by $v_i(t)$, to represent the effects of transient disturbances. The system equations therefore become, in the most general form to be used here,

$$p_i'(t) = v_i(t) + h_i(t) * \sum_{j=1}^{N} a_{ij}[p_j(t - \tau_{ij}) - p_i(t)], \qquad i = 1, \cdots, N. \tag{6}$$

where the asterisk (*) denotes convolution. Neither of these changes affects the equilibrium frequency.

It has been assumed that the filter gains and averaging coefficients are all nonnegative,

$$\lambda_i \geqq 0, \tag{7}$$

$$a_{ij} \geqq 0. \tag{8}$$

The connectivity of the network depends on which of these coefficients are zero. If the ith station does not receive from the jth, a_{ij} is zero, except when the ith station does not receive from any other station, in which case (1) forbids all a_{ij} to be zero and λ_i must be zero, and the a_{ij} are then arbitrary. It is understood that if the ith station in fact receives a digital signal from the jth, but uses it only as a medium of communication and does not use its observed phase in controlling its clock, it will be said that the ith station "does not receive from" the jth.

III. THE INITIATION OF SYNCHRONOUS OPERATION

Previous studies have assumed that the system is placed in synchronous operation at $t = 0$ by simultaneously closing all the switches at each station that connect the control voltages to the oscillators. It is assumed that before $t = 0$ all oscillators are operating at their free-running frequencies, and have been running for a sufficiently long time so that, in spite of transmission delays, all stations are receiving signals on all links by the time the switches are closed. Closure of each switch will, in general, cause an immediate change of frequency at every station, and prediction of the frequency at which the system finally will settle down would be a matter of practical importance.

In practice it may be preferable to assemble the system in more leisurely fashion — one station at a time — checking for proper operation after each station is connected before connecting the next one. One might, for example, realize the reference phases r_{ij} as manually controlled bias voltages. When a new station is to be connected into the system, the first connection will be made at the new station, from one of the phase detectors to the input of the clock control filter. This connection will synchronize the new station with the system as a slave station, and adjustment of the corresponding reference phase can be used to establish any desired phase relation between it and the rest of the system. When each subsequent connection is made from a phase detector to a clock control filter, the associated reference phase is adjusted so as to null the voltage across the switch at the moment when it is closed. There is then no discontinuous change in frequency at any time during the connection process.

If the system were built up in this way, starting from one station as the initial system, and if there were no drifts in either free-running frequencies or transmission delays, the final equilibrium frequency of the system would be the free-running frequency of the first station. In this case the equilibrium frequency could be predicted without any calculation. In any case, the system frequency can be deliberately changed after initiation of synchronous operation by adjusting the bias voltages.

The equation for system frequency is still useful as a means of predicting the effects of drifts in the free-running frequencies and the transmission delays. However, serious questions can in principle arise with regard to the applicability of the general system equations (6). These equations represent a system with invariant connectivity, represented by invariant averaging coefficients a_{ij} and gains λ_i, while the actual system connectivity has been a function of time. Karnaugh[4] and Gersho and Karafin[2] have answered these questions for their models under the particular initiation procedure they assumed. The answer will now be extended to cover the present model for arbitrary initiation procedure.

I shall take the point of view that there is some specifiable moment t_0 at which the system has been completely assembled, so that the a_{ij} and λ_i are invariant for $t \geqq t_0$, and that we need only to predict the future behavior of the system, for specified disturbances in $v_i(t)$ and drifts in τ_{ij}, having full knowledge of the past behavior of the system. For the purpose of this discussion, if the transmission delays τ_{ij} are to be allowed to change, they should be considered as having been written as $\tau_{ij}(t)$.

The system equations (6) are actually integrodifferential equations, since the convolution symbol $(*)$ implies an integration. The initial conditions on which the solution of this equation depends are the entire history of the phase variables $p_i(t)$, to the extent that this history determines the state of the filters. The output of each filter for $t \geqq t_0$ can be considered as the sum of two components: a transient term determined by the state of the filter, which is in turn determined by the input for $t < t_0$, and a term representing the response to inputs for $t \geqq t_0$. The transient terms can be calculated from the known filter inputs for $t < t_0$ and included in the $v_i(t)$ terms. Equations (6), with these terms included in $v_i(t)$, with the filters considered quiescent at $t = t_0$, and with the correct initial values of $p_i(t_0)$, will then give a correct description of the behavior of the system for $t \geqq t_0$.

This argument is included here only to establish the validity of (6) in principle. In practical calculations, estimates of the effects of transient disturbances would normally assume an equilibrium state as the initial condition.

IV. EQUILIBRIUM STATES

Gersho and Karafin[2] determined the equilibrium frequency of the system as a limiting value derived by means of the final value theorem for Laplace transforms. Karnaugh[4] used a simpler method, claiming for it only heuristic value. The following approach claims rigorous validity for the simpler method.

The first step in the analysis of the system will be the determination of its equilibrium states, without regard for whether they are stable or unstable equilibria. These can be determined by assuming that the system has been placed in some state, and that it will not change state spontaneously; any state that satisfies these conditions is an equilibrium state. We determine in this step whether the equilibrium state is unique. The second step is to determine whether the system can respond to any transient excitation with components that do not approach zero with increasing time; this step determines the stability of the system. The linearity of the system now implies that if the equilibrium state is unique, and the transient response approaches zero, the system will always approach the equilibrium state in the absence of a disturbance.

It will, in fact, be found that the equilibrium state is not unique, because the system equations include the phases only in phase difference terms, and an arbitrary common constant can be added to every phase variable without changing the phase differences. There is, therefore, a continuum of equilibrium states, all of which are equivalent for practical purposes in that they have the same phase differences and the same system frequency. Because of this equivalence, the system will be considered stable if, after a transient disturbance, it approaches any equilibrium state, not necessarily the one it occupied before the disturbance. This requires only that the transient components of the phase differences approach zero, while the transient components of the phases may approach arbitrary limits.

This section will deal only with the first step: the identification, including determination of conditions for existence, of equilibrium states. The stability of the equilibrium states can be assured by the sufficient condition studied by Gersho and Karafin;[2] their proof remains valid under the weaker connectivity condition shown here to be necessary and sufficient, the statement that at least one M_{i1} is positive sufficing to replace their statement that all M_{i1} are positive.

The only equilibrium states to be considered here are those in which all clock frequencies are constant at a common value; if they are constant, but at different values, synchronism has not been established. It will also be required that the existence of such a state should not depend on the values of the free-running frequencies; if it does, the system is not self-synchronizing, but is synchronous only if the clocks are adjusted by means external to the system.

The instantaneous frequencies $p_i'(t)$ are, therefore, set equal to an equilibrium frequency denoted simply by f. Then

$$p_i(t) = ft + P_i, \qquad i = 1, \cdots, N. \tag{9}$$

Equation (4), which suffices even in the most general linear case represented by (6) for the description of the steady state, becomes

$$f = v_i + \lambda_i \sum_{j=1}^{N} a_{ij}(P_j - P_i - f\tau_{ij}), \qquad i = 1, \cdots, N, \tag{10}$$

or, in more symmetrical form,

$$\sum_{j=1}^{N} \lambda_i(\delta_{ij} - a_{ij})P_j = v_i - f(1 + \lambda_i\tau_i), \qquad i = 1, \cdots, N, \tag{11}$$

where

$$\tau_i = \sum_{j=1}^{N} a_{ij}\tau_{ij} \tag{12}$$

and δ_{ij} is the Kronecker delta, equal to unity for $i = j$ and zero otherwise. The set of equations (11) looks as though it could be solved for the P_j in terms of arbitrary v_i and f, but it cannot, because the matrix of coefficients on the left is singular. This will be stated and proved as a theorem.

Theorem I: Let L denote the diagonal matrix with diagonal elements λ_i, let A denote the averaging matrix with elements a_{ij}, and let I denote the identity matrix. Then the matrix $L(I - A)$, with elements $\lambda_i(\delta_{ij} - a_{ij})$, has rank less than its order N.

Proof: By (1), the sum of the elements in any row is zero. Therefore, the sum of all columns is a column of zeros. Therefore the matrix is singular and its rank is less than its order, Q.E.D.

It is advantageous at this point to choose one P_i arbitrarily as a reference for the others. With P_1 as reference, we change to the phase difference variables

$$Q_j = P_j - P_1, \qquad j = 2, \cdots, N. \tag{13}$$

The equations (11) then become

$$\sum_{j=2}^{N} \lambda_i(\delta_{ij} - a_{ij})Q_j = v_i - f(1 + \lambda_i\tau_i), \qquad i = 1, \cdots, N. \quad (14)$$

If the term in f is transposed to the left side, we get the set of N equations

$$f(1 + \lambda_i\tau_i) + \sum_{j=2}^{N} \lambda_i(\delta_{ij} - a_{ij})Q_j = v_i\,, \qquad i = 1, \cdots, N, \quad (15)$$

which we expect to be able to solve for the N unknowns f and Q_j , $j = 2, \cdots, N$, for arbitrary v_i .

If we formally solve for f by determinants, and expand each determinant in terms of the elements of the first column and their cofactors, the result is

$$f = \frac{\sum_{i=1}^{N} b_i v_i}{\sum_{i=1}^{N} b_i(1 + \lambda_i\tau_i)}, \qquad (16)$$

where b_i is the cofactor (signed minor) of the element in the first column, ith row, of the matrix $L(I - A)$. The following theorem shows that the arbitrary choice of P_1 as the reference for phase differences, and the definition of b_i in terms of the first column, makes no difference in the result.

Theorem II: Let M_{ij} be the cofactor of the (i, j)th element of $L(I - A)$; then $M_{ij} = M_{ik}$ for all i, j, k, that is, all cofactors of elements in the same row are equal, and hence $M_{ij} = b_i$ for all $i, j = 1, \cdots, N$.

Proof: If the rank of $L(I - A)$ is less than $N - 1$ then all M_{ij} are zero and the theorem is satisfied. If the rank is $N - 1$, the matrix equation $L(I - A)x = 0$, where x is an N element column matrix, has only one independent solution. It is known from (1) that a solution exists in which all components are equal, and this must now be true of any solution. It can also be shown that the cofactors of any single row of the matrix $L(I - A)$ must be a solution (see, for example, Guillemin[5]), hence all cofactors of elements in a row must be equal, $M_{ij} = M_{ik}$, hence $M_{ij} = M_{i1} = b_i$, Q.E.D.

Since b_i can now be defined without reference to any particular column, the single-index notation is justified. Since we expect that increasing the free-running frequency of any oscillator will never decrease the equilibrium frequency f, we should expect all the b_i to be nonnegative. The following theorem verifies this expectation.

Theorem III: The cofactors of elements of the matrix $L(I - A)$ are nonnegative, $b_i \geqq 0$, $i = 1, \cdots, N$.

Proof: If a matrix is diagonally dominated, i.e., if every diagonal element is greater in magnitude than the sum of the magnitudes of all other elements in the same row, it is easily shown (Appendix I, Gersho and Karafin[2]) that it must be nonsingular. Consider the matrix $L(I - \varepsilon A)$, $0 \leqq \varepsilon \leqq 1$. The cofactors of its diagonal elements are continuous functions of ε. For $\varepsilon = 0$ they are all unity, hence positive. For $0 < \varepsilon < 1$ the cofactors are the determinants of diagonally dominated submatrices, hence nonzero, so that they cannot pass through zero, and must remain positive. Hence, as $\varepsilon \to 1$ they cannot approach negative limits. But as $\varepsilon \to 1$ they approach the values b_i, hence $b_i \geqq 0$, Q.E.D.

The formal solution (16) is valid if and only if the matrix of coefficients on the left side of (15) is nonsingular; that is, if and only if the denominator of (16) is nonzero. But, since b_i, λ_i, and τ_i are all nonnegative, this is equivalent to the condition that at least one b_i be positive. The following definitions and theorems relate this algebraic condition to the connectivity properties of the network.

Definition: The jth station is said to send to the ith, or equivalently, the ith station is said to receive from the jth, if $\lambda_i a_{ij}$ is positive.

Definition: The jth station is said to send directly or indirectly to the ith, or equivalently, the ith station is said to receive directly or indirectly from the jth, if there exists a chain (ordered set) of stations such that the first is station j, the second receives from j, each receives from the one before, and the last is station i.

Theorem IV: If the kth station does not transmit directly or indirectly to all other stations then $b_k = 0$.

Proof: Let A_{kk} be the submatrix formed by deleting the kth row and column of $L(I - A)$. Let S_k be the set of indices of all stations that do not receive directly or indirectly from the kth. By hypothesis S_k is nonempty; choose $i \epsilon S_k$. By the definition of S_k, $\lambda_i a_{ij}$ is zero if j is not in S_k, hence, from (1),

$$\sum_{j \epsilon S_k} \lambda_i a_{ij} = \lambda_i , \qquad i \epsilon S_k . \tag{17}$$

Let B_k be the square submatrix of A_{kk} consisting of all elements whose row and column indices are both in S_k. Then (17) shows that B_k can be written in the form $L'(I - A')$, where A' is an averaging matrix satisfying (1). Hence, by Theorem I, B_k is singular, and the rows of B_k are linearly dependent. But since the ith row of A_{kk}, for all $i \in S_k$, is the ith row of B_k augmented with zeros, the same linear dependence holds among the rows of A_{kk}. Hence, A_{kk} is singular; hence its determinant, which is b_k, is zero, Q.E.D.

Theorem V: If b_k is zero, then the kth station does not transmit directly or indirectly to every other station.

Proof: Let A_{kk} be defined as in the proof of Theorem IV. By hypothesis, A_{kk} is singular; hence there exists a column matrix x, with elements x_i not all zero, such that $A_{kk}x = 0$, or equivalently

$$\sum_{j \neq k} \lambda_i a_{ij} x_j = \lambda_i x_i , \qquad i \neq k. \tag{18}$$

Let M be the magnitude of the x_i having the largest magnitude. Let S_k be defined now as the set of all indices i for which $| x_i | = M$; obviously S_k is nonempty. Now (18) implies

$$\left| \sum_{j \neq k} \lambda_i a_{ij} x_j \right| = \lambda_i M, \qquad i \in S_k . \tag{19}$$

Now (1) and $| x_j | \leqq M$ imply that this can be true only if $\lambda_i a_{ik} = 0$ and $| x_j | = M$ whenever $\lambda_i a_{ij} > 0$. Hence, for all $i \in S_k$, $\lambda_i a_{ij} = 0$ except when $j \in S_k$, and thus the ith station cannot receive directly or indirectly from the kth. Since S_k is nonempty, the kth station does not transmit directly or indirectly to all stations, Q.E.D.

It follows from these theorems that the formal solution (16) is valid for the set of equations (15) if and only if there is at least one station that transmits directly or indirectly to all other stations.

If there is no such station, the matrix of coefficients on the left side of (15) is singular, and the set of equations has either no solution or an infinity of solutions, depending on the values of the v_i. Since a solution defines an equilibrium state in which all oscillators run at the same frequency, this means that the oscillators will run at the same frequency only if their free-running frequencies are appropriately adjusted; that is, the system is not self-synchronizing.

If there is only one station that transmits directly or indirectly to all others, that station is the master, setting the frequency for the whole system. A single master receives from no other station, since any station that transmitted to it would thereby transmit indirectly to all other stations. Thus, a station can become a master simply by the loss of all inputs from other stations. However, if two stations lose all their inputs, the system fails to synchronize, since neither station sends directly or indirectly to every other.

If more than one station sends directly or indirectly to all others, these stations are mutually synchronized, and jointly establish the system frequency. Any station that does not send directly or indirectly to every other station is in effect a slave station.

V. SUMMARY AND CONCLUSIONS

The process for initiation of synchronous operation described in Section III is not necessarily recommended as the best possible. It is intended as a constructive existence proof, showing that there exists a method of setting up a synchronized system of geographically separated clocks that will lead to a final frequency that can be determined in advance. The second part of that section shows, in perhaps unnecessary detail, that the behavior of a system, once it has been set up, can be determined without considering how it was set up, so that it is not necessary to specify the set-up procedure before studying its steady-state or dynamic behavior.

Under these circumstances the equation for equilibrium frequency developed in Section IV plays no part in the process of setting the system in synchronism and adjusting it to run at the desired frequency. It serves to identify the factors that affect the final frequency and indicate the quantitative effect of each factor, and as such would appear to find its greatest usefulness in the design and control of the configuration of system interconnections.

The connectivity condition evolved in Section IV permits the inclusion of single-master systems in the same general class as completely mutually synchronized systems. It is suggested that these two types are in fact opposite extremes of a more general class in which the most useful configurations may have some intermediate hierarchical form.

REFERENCES

1. Beneš, V. E., unpublished memorandum, 1959.
2. Gersho, A. and Karafin, B. J., Mutual Synchronization of Geographically Separated Oscillators, B.S.T.J., this issue, pp. 1673–1704.
3. Runyon, J. P., unpublished memorandum, 1959.
4. Karnaugh, M., A Model for the Organic Synchronization of Communications Systems, B.S.T.J., this issue, pp. 1705–1735.
5. Guillemin, E. A., *The Mathematics of Circuit Analysis,* John Wiley & Sons, New York, 1949, pp. 17, 103.