12. Brown, G. M., and L. F. Stutzman, *Chem. Eng. Progr.*, **45**, 142 (1949).
13. Davis, P. C., A. F. Bertuzzi, R. L. Gore, and Frederick Kurata, *Am. Soc. Mech. Engrs. Trans.*, **201**, 245 (1954).
14. "Smithsonian Physical Tables," Publ. 4169, Smithsonian Inst., Washington, D. C. (1954).
15. Faulkner, R. C., Ph.D. thesis, Univ. Michigan, Ann Arbor (1959).
16. Jones, M. L., Jr., Ph.D. thesis, Univ. Michigan, Ann Arbor (1961).
17. Mage, D. T., Ph.D. thesis, Univ. Michigan, Ann Arbor (1964).
17a. ———, *J. Chem. Phys.*, **42**, 2977 (1965).
18. Jones, M. L., Jr., D. T. Mage, R. C. Faulkner, and D. L. Katz, *Chem. Eng. Progr. Symposium Ser. No. 44*, 59, 52 (1963).
19. Manker, E. A., D. T. Mage, A. E. Mather, J. E. Powers, and D. L. Katz, *Proc. Natl. Gas Proc. Assoc.*, 3 (1964).
20. Mage, D. T., M. L. Jones, Jr., D. L. Katz, and J. R. Roebuck, *Chem. Eng. Progr. Symposium Ser. No. 44*, **59**, 61 (1963).
21. Goff, J. A., and Serge Gratch, *Trans. Am. Soc. Mech. Engrs.*, 72, 741 (1950).
22. Friedman, A. S., and David White, *J. Am. Chem. Soc.*, **72**, 3931 (1952).
23. ———, and H. L. Johnston, unpublished data.
24. Schneider, W. G., *Can. J. Res.*, **27B**, 4, 339 (1949).
25. Pfefferle, W. C., Jr., J. A. Goff, and J. G. Miller, *J. Chem. Phys.*, **23**, 509 (1955).
26. Michels, A., and H. Wouters, *Physica*, **8**, 923 (1941).
27. Scheel, K., and W. Heuse, *Ann. Phys.*, **40**, 473 (1913).
28. *Ibid.*, **37**, 79 (1912).
29. Roebuck, J. R., and Harold Osterberg, *Phys. Rev.*, **43**, 60 (1933).
30. Roebuck, J. R., collection of experimental records donated to Univ. Michigan Library.
31. Mann, D. B., *U.S. Natl. Bur. Stds. Tech. Note 154A* (1962).
32. Pfenning, D. B., J. B. Canfield, and Riki Kobayashi, *J. Chem. Eng. Data*, **10**, 9 (1965).
33. Bloomer, O. T., and K. N. Rao, *Inst. Gas. Technol. Res. Bull.* 18, (1952).
34. Strobridge, T. R., *U.S. Natl. Bur. Stds. Tech. Note 129A* (1962).

Manuscript received March 9, 1965; revision received September 9, 1965; paper accepted September 9, 1965. Paper presented at A.I.Ch.E. San Francisco meeting.

The Time-Optimal Control of Discrete-Time Linear Systems with Bounded Controls

HERBERT A. LESSER and LEON LAPIDUS

Princeton University, Princeton, New Jersey

The discrete-time, time-optimal control of high-order, linear systems with bounded controls is formulated as a problem in linear programming. The solution obtained requires a certain minimum number of controls to be on their bounds. Numerical results are obtained for a system with controls bounded on one and both sides and extensions indicated for state variable constraints. The method proves to be extremely fast and simple to solve on a high-speed digital computer.

In recent years there has been considerable interest in the optimal control of dynamic systems with much of the work to date being based on dynamic programming (*1*) or the maximum principle (*12, 14*). Proposed methods for obtaining solutions often work well for simple systems, but numerical difficulties are frequently encountered in deriving the optimal control of systems which are either highly nonlinear or of high order.

This paper describes the first phase of a study (9) of the time-optimal control of high-order, linear systems with bounded controls. This particular phase of the study treated discrete-time (sampled data) systems, that is, systems whose control variables may be changed at the discrete times $k\Delta t$, $k = 1, 2, \ldots$ but must be held constant for $k\Delta t < t < (k + 1)\Delta t$, $k = 0, 1, 2, \ldots$. The time increment Δt is called the *sampling period.* It is desired to find that sequence of controls which will bring the system from its initial state to some desired state in a minimum number of sampling periods.

The control of high-order systems is especially important in the chemical process industries where most of the process systems have either a large number of state

Herbert A. Lesser is with Esso Production Research Company, Houston, Texas.

variables or are distributed-parameter systems which may be approximated by high-order, lumped-parameter systems. The control variables always have physical bounds on them. Although few such systems are strictly linear, many of them may be considered linear under certain operating conditions.

Interest in the control of discrete-time systems was initiated by a study of Kalman's in 1957 (*8*). In 1961 Desoer and Wing extended some of Kalman's preliminary results to obtain the time-optimal control of some specific systems (*2* to *4*). In reference *3*, a second-order system with a single control variable u, $|u(k\Delta t)| \leq 1$, was considered. In reference *2*, the n^{th} order case was treated with the restrictions of a single control variable u, $|u(k\Delta t)| \leq 1$, and a transfer function with real, distinct, nonpositive poles. Desoer and Wing (*4*) presented a time-optimal strategy for a general single-control, linear-discrete system, with the only restrictions being that the system be controllable in the Kalman sense (*7*) and that $|u(k\Delta t)| \leq 1$ for all k. It was shown how to determine $u(k\Delta t)$ as a function of the n dimensional state vector $\mathbf{x}$ of the system, such that $\mathbf{x}$ could be brought to $\mathbf{0}$ in minimum time. In 1961, Ho (*6*) formulated the general-discrete, optimal-control problem as a problem in nonlinear programming. He discussed some of the properties of the time-optimal case.

The following year Zadeh and Whalen (*16*) pointed out how the time-optimal problem could be converted to a problem in linear programming. This concept has been implemented and extended by Lesser (*9*) and Mangasarian (*11*) to obtain some numerical results. The use of linear programming is of special interest because of the general availability of well-written linear programs which require very little computational time on a high-speed digital computer. Torng (*15*) has also implemented the concept of Zadeh and Whalen in a study done independently and at about the same time as that of Lesser. He has demonstrated that the linear programming solution may not be unique, in which case a minimum-fuel criterion is invoked to choose between alternate solutions.

In the present paper a theorem is developed which describes the nature of the linear programming solutions suggested by Zadeh and Whalen. In addition, it is explained how the linear program can be modified to allow the controls to be bounded on one or both sides and to allow for constraints on the state vector. A single numerical example involving a lumped-parameter system is detailed to show the versatility of the method. A second example, presented in Appendix C,* is briefly outlined to indicate the possible direction for more complex systems (distributed-parameter systems).

THEORETICAL DEVELOPMENT

Statement of the Control Problem

The dynamic behavior of the system of interest is given by the set of linear, ordinary differential equations

$$\dot{x}_i = \sum_{j=1}^{n} a_{ij}x_j + \sum_{j=1}^{r} b_{ij}u_j; \quad i = 1, 2, \ldots, n \qquad (1)$$

where the x_i's are the state variables, the u_i's are the control variables, and $\dot{x}_i$ is the time derivative of x_i. The coefficients a_{ij} and b_{ij} may, in general, be time dependent. However, since so little computational work of any kind has been done on the time-optimal control of this system, numerical examples with constant coefficients were chosen for this study. The control variables are considered to be bounded above and/or below. The set of system equations may be written in a more compact form than Equation (1):

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu} \qquad (2)$$

where $\mathbf{x}$ is an n vector $(x_1, x_2, \ldots, x_n)$, $\mathbf{u}$ is an r vector $(u_1, u_2, \ldots, u_r)$, $\mathbf{A}$ is an $(n \times n)$ matrix of the coefficients a_{ij} and $\mathbf{B}$ is an $(n \times r)$ matrix of the coefficients b_{ij}. It is desired to find the control sequence $\mathbf{u}(k\Delta t)$, $k = 0, 1, 2, \ldots$, which will drive the state vector $\mathbf{x}$ from its initial state $\mathbf{x}(0)$ to a desired state $\mathbf{x}^d$ in minimum time.

It should be pointed out that if $\mathbf{x}^d \neq \mathbf{0}$, a linear transformation of Equation (2) can always be made so that the desired state of the new state vector will be $\mathbf{0}$. If each control is bounded on both sides and $\mathbf{x}^d$ is taken to be a steady state of the system, a new control vector arises under the transformation, each component of which is bounded by a negative number below and a positive number above. Therefore, no generality will be lost in subsequent discussion when $\mathbf{x}^d$ is set equal to $\mathbf{0}$ and u_i is assigned bounds $-\alpha_i$ and β_i, $i = 1, \ldots, r$.

Thus, the problem to be solved is one of finding the control sequence $\mathbf{u}(k\Delta t)$, $k = 0, 1, 2, \ldots, (N-1)$, which will bring the system from $\mathbf{x}_0$ to $\mathbf{0}$ in a minimum number of time (sampling) periods N, subject to $-\alpha_i \leq u_i \leq \beta_i$, $i = 1, \ldots, r$. The quantities α_i and β_i are positive. The case of either all α_i or all β_i being infinite is also of interest, since this yields controls bounded on one side only.

The performance index to be optimized is not always taken by investigators to be N. As a typical illustration, the performance index was taken by Lapidus et al. (*10*) to be P, where

$$P = \sum_{k=1}^{N} \mathbf{x}^T(k\Delta t)\mathbf{Q}\mathbf{x}(k\Delta t), \quad N \to \infty$$

and

$$\mathbf{x}^d = \mathbf{0}$$

The matrix $\mathbf{Q}$ is a diagonal matrix with positive elements. In this case the performance index is a weighted sum of squares of the deviation of the state variables from their desired values, summed over all control periods. The result of minimizing P is to bring $\mathbf{x}(k\Delta t)$ as near as possible to $\mathbf{0}$ at every sampling instant $k\Delta t$. Thus, the elements of $\mathbf{x}$ approach $\mathbf{0}$ very quickly at first, then approach it asymptotically with time (*10*). On the other hand, minimizing the number of sampling periods N required to reach $\mathbf{0}$ does not place any restriction on the distance of $\mathbf{x}$ from $\mathbf{0}$ for $k < N$. One might then anticipate that the elements of $\mathbf{x}$ will deviate more greatly from zero (for $k < N$) than they would for the minimum P problem.

SOLUTION IN THE ABSENCE OF CONTROL CONSTRAINTS

Since Equation (2), the dynamic equation of interest, is linear, it may be integrated directly and a corresponding discrete equation obtained. Thus

$$\mathbf{x}([k+1]\Delta t) = \exp(\mathbf{A}\Delta t)\mathbf{x}(k\Delta t) + \int_0^{\Delta t} \exp(\mathbf{A}s)\,ds \cdot \mathbf{Bu}(k\Delta t) \qquad (3)$$

The matrices $\mathbf{A}$ and $\mathbf{B}$ are constant and the control vector $\mathbf{u}(k\Delta t)$ is taken to be constant over the time interval $k\Delta t < t < (k+1)\Delta t$. For compactness of notation, the following substitutions are made:

$$\mathbf{G} = \exp(\mathbf{A}\Delta t) = \sum_{i=0}^{\infty} \frac{(\mathbf{A}\Delta t)^i}{i!}$$

* Deposited as document 8645 with the American Documentation Institute, Photoduplication Service, Library of Congress, Washington 25, D. C., and may be obtained for $1.25 for photoprints or 35-mm. microfilm.

$$J = \int_0^{\Delta t} \exp(\mathbf{A}s)\mathbf{B}ds$$
$$\mathbf{x}_k = \mathbf{x}(k\Delta t)$$
$$\mathbf{u}_k = \mathbf{u}(k\Delta t)$$

Then Equation (3) becomes

$$\mathbf{x}_{k+1} = \mathbf{G}\mathbf{x}_k + \mathbf{J}\mathbf{u}_k \quad (4)$$

Repeated use of this equation gives the relation between the initial and final conditions:

$$\begin{aligned}
\mathbf{x}_1 &= \mathbf{G}\mathbf{x}_o + \mathbf{J}\mathbf{u}_o \\
\mathbf{x}_2 &= \mathbf{G}\mathbf{x}_1 + \mathbf{J}\mathbf{u}_1 = \mathbf{G}^2\mathbf{x}_o + \mathbf{G}\mathbf{J}\mathbf{u}_o + \mathbf{J}\mathbf{u}_1 \\
\mathbf{x}_3 &= \mathbf{G}\mathbf{x}_2 + \mathbf{J}\mathbf{u}_2 = \mathbf{G}^3\mathbf{x}_o + \mathbf{G}^2\mathbf{J}\mathbf{u}_o + \mathbf{G}\mathbf{J}\mathbf{u}_1 + \mathbf{J}\mathbf{u}_2 \\
&\vdots \\
\mathbf{x}_N &= \mathbf{G}^N\mathbf{x}_o + \mathbf{G}^{N-1}\mathbf{J}\mathbf{u}_o + \mathbf{G}^{N-2}\mathbf{J}\mathbf{u}_1 + \ldots + \mathbf{G}\mathbf{J}\mathbf{u}_{N-2} + \mathbf{J}\mathbf{u}_{N-1}
\end{aligned}$$

or

$$\mathbf{x}_N = \mathbf{G}^N\mathbf{x}_o + \sum_{k=0}^{N-1} \mathbf{G}^{N-1-k}\mathbf{J}\mathbf{u}_k \quad (5)$$

Replace the matrix product $\mathbf{G}^{N-1-k}\mathbf{J}$ by $\mathbf{L}_{N-1-k}$. Then Equation (5) becomes

$$\mathbf{x}_N = \mathbf{G}^N\mathbf{x}_o + \sum_{k=0}^{N-1} \mathbf{L}_{N-1-k}\mathbf{u}_k \quad (6)$$

The sampled-data, time-optimal problem with no constraints may now be stated in the following fashion:

Problem 1: Given an n dimensional state vector $\mathbf{x}$ which has an initial condition $\mathbf{x}_o$ and whose final position in phase space is given by Equation (6), find that sequence of controls $\mathbf{u}_k$, $k = 0, 1, \ldots, (N-1)$, which makes $\mathbf{x}_N = \mathbf{x}^d$ and minimizes the number of sampling periods N.

The final condition may be written

$$\mathbf{L}\mathbf{U} = \mathbf{x}^d - \mathbf{G}^N\mathbf{x}_o \quad (7)$$

where

$$\mathbf{U} = \begin{bmatrix} \mathbf{u}_o \\ \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_{N-2} \\ \mathbf{u}_{N-1} \end{bmatrix} \quad (8)$$

and

$$\mathbf{L} = [\mathbf{L}_{N-1}\ \mathbf{L}_{N-2} \ldots \mathbf{L}_1\ \mathbf{J}] \quad (9)$$

The vector $\mathbf{U}$ has Nr components and the matrix $\mathbf{L}$ has dimensions $n \times Nr$. The array $\mathbf{LU}$ is thus $n \times 1$ and has the same dimensions as the right-hand side of Equation (7). If there are no constraints on the control variables, the minimum number of sampling periods will be given by

$$Nr = n$$

or

$$N = n/r \quad (10)$$

provided n/r is an integer and provided the matrix $\mathbf{L}$ has a nonzero determinant, that is, $\det(\mathbf{L}) \neq 0$. If n/r is not an integer, the minimum N will, in general, be the first integer greater than n/r. Equation (10) is another way of requiring matrix Equation (7) to be a set of n equations in n unknowns. It is well known that when the right-hand side of such a set of equations does not contain all zeros, a solution exists if, and only if, the determinant of the coefficients of the unknowns is not zero. This explains the requirement $\det(\mathbf{L}) \neq 0$. The solution to Problem 1 then becomes

$$\mathbf{U} = \mathbf{L}^{-1}(\mathbf{x}^d - \mathbf{G}^N\mathbf{x}_o),\ N = \frac{n}{r};\ \det(\mathbf{L}) \neq 0 \quad (11)$$

SOLUTION IN THE PRESENCE OF CONTROL CONSTRAINTS

Suppose now that the control variables are constrained, which is always the case in the real world. The solution given by Equation (11) may require negative input concentrations, unattainable temperatures or flow rates, and the like. Compliance with the constraints would therefore necessitate rejecting Equation (11) as giving the required control, which would, in turn, prevent the system from reaching the desired final state in n/r sampling periods.

It will now be assumed that the state vector has been defined in such a way that $\mathbf{x}^d$ of Equation (7) is the null vector $\mathbf{0}$. The bounds on the control vector $\mathbf{u}_k$ will be given by

$$-\alpha_i \leqq (\mathbf{u}_k)_i \leqq \beta_i;\ i = 1, 2, \ldots, r \quad (12)$$

where all the α_i and β_i are positive. The equation to be satisfied is then

$$\mathbf{L}\mathbf{U} = -\mathbf{G}^N\mathbf{x}_o \quad (13)$$

where the vector $\mathbf{U}$ consists, as before, of the N control vectors $\mathbf{u}_k$. The sampled-data, time-optimal control problem with bounded controls can now be stated in the following fashion:

Problem 2: Given an n dimensional state vector $\mathbf{x}$ which has an initial condition $\mathbf{x}_o$ and whose final position in phase space is given by Equation (6), find that $\mathbf{U}$ which satisfies Equation (13) and minimizes the number of sampling periods N, subject to the control constraints of Equation (12).

If the solution given by Equation (11) with $\mathbf{x}^d = \mathbf{0}$ does not give any control element at any instant of time which lies outside of its restricted region, then

$$\mathbf{U} = -\mathbf{L}^{-1}\mathbf{G}^N\mathbf{x}_o \quad (14)$$

is the solution to the problem, and the minimum number of sampling periods is $N = n/r$, provided n/r is an integer. In this case, the constraints on the controls do not affect the solution to the problem. However, if any control element lies outside its restricted region, a larger number of sampling periods will be needed to achieve the desired result, and Equation (14) must be rejected as the required solution.

To solve Problem 2 when Equation (14) produces inadmissible controls or when n/r is not an integer, use is made of linear programming, as suggested by Zadeh and Whalen (*16*). It is, of course, assumed that the origin can be reached in a finite number of sampling periods with constraints on the controls. A linear program solves for those Q unknowns which will simultaneously satisfy a set of S algebraic linear equations $Q > S$ and minimize a given linear combination of the unknowns. It will be shown how Problem 2 can be formulated as such a set of equations with one unknown to be minimized.

An error vector $\mathbf{E}$ is introduced where

$$\mathbf{E} = \mathbf{L}\mathbf{U} + \mathbf{G}^N\mathbf{x}_o \quad (15)$$

The problem is solved when the $\mathbf{U}$ is found which makes $\mathbf{E} = \mathbf{0}$ and at the same time minimizes N. The vector $\mathbf{E}$ may be forced to $\mathbf{0}$ by seeking to minimize $\max_i |E_i|$ over the control vector $\mathbf{U}$, subject to the control constraints. This is equivalent to minimizing g subject to the $2n$ constraints:

$$\begin{aligned}
g + E_i &\geqq 0;\ i = 1, \ldots, n \\
g - E_i &\geqq 0;\ i = 1, \ldots, n
\end{aligned} \quad (16)$$

Minimizing g will force $\mathbf{E}$ to be $\mathbf{0}$ for the minimum number or any larger number of sampling periods. The procedure consists then of fixing N and minimizing g. If $g = 0$, then all $E_i = 0$; the procedure is repeated after decreasing N until the smallest N is found such that g

= 0. This N and the corresponding $\mathbf{U}$ are the desired results. If the initial N yields a nonzero g, then a larger N obviously must be tried.

Linear programming requires the unknown variables to be non-negative (5). A new control vector $\mathbf{U}'$ is therefore introduced, whose components are all non-negative:

$$\mathbf{U}' = \mathbf{U} + \left.\begin{bmatrix} \alpha \\ \alpha \\ \cdot \\ \cdot \\ \cdot \\ \alpha \\ \alpha \end{bmatrix}\right\} N \ \alpha \text{ vectors}$$

$$\mathbf{U}' = \mathbf{U} + \underset{\sim}{\alpha} \tag{17}$$

$$\mathbf{H} = \begin{array}{c} n\left\{\begin{array}{c} \\ \\ \\ \end{array}\right. \\ n\left\{\begin{array}{c} \\ \\ \\ \end{array}\right. \\ Nr\left\{\begin{array}{c} \\ \\ \\ \\ \end{array}\right. \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & \overbrace{\hphantom{L_{N-1}\ L_{N-2}\ldots L_1\ J}}^{Nr} & \overbrace{\hphantom{-1\ldots-1}}^{n} & \overbrace{\hphantom{-1\ldots-1}}^{n} & \overbrace{\hphantom{1\ldots 1}}^{Nr} \end{array} \\ \begin{bmatrix} \begin{matrix}1\\ \vdots \\ 1\end{matrix} & \mathbf{L}_{N-1}\ \ \mathbf{L}_{N-2}\ldots\ \mathbf{L}_1\ \ \mathbf{J} & \begin{matrix}-1 & & \\ & \ddots & \\ & & -1\end{matrix} & & \\ \begin{matrix}1\\ \vdots \\ 1\end{matrix} & -\mathbf{L}_{N-1}\ -\mathbf{L}_{N-2}\ldots -\mathbf{L}_1\ -\mathbf{J} & & \begin{matrix}-1 & & \\ & \ddots & \\ & & -1\end{matrix} & 0 \\ 0 & \begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix} & 0 & & \begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix} \end{bmatrix} \end{array}$$

The r components of $\underset{\sim}{\alpha}$ are the α_i's of Equation (12). In addition $2n$ state slack variables s_i and Nr control slack variables c_i are introduced. The purpose of the state slack variables, which, like the control variables, must be non-negative, is to remove the inequality signs in Equation (16). In other words

$$\begin{gathered} g + E_i - s_i = 0;\ i = 1, \ldots, n \\ g - E_i - s_{i+n} = 0;\ i = 1, \ldots, n \end{gathered} \tag{18}$$

are equivalent to Equation (16) if s_i and s_{i+1} are non-negative for all i. Similarly, the control slack variables are also non-negative and insure that the control variables will not exceed their upper limits. Typically

$$(\mathbf{u}_k')_i - (\underset{\sim}{\alpha} + \underset{\sim}{\beta})_i + c_i = 0$$

insures that $(\mathbf{u}_k')_i$ will not exceed $(\underset{\sim}{\alpha} + \underset{\sim}{\beta})_i$.

Equations (12), (15), (16), and (17) may be combined with the slack variables to give, with $\underset{\sim}{\beta}'$, an Nr dimensional vector defined similarly to $\underset{\sim}{\alpha}'$:

$$g\begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ 1 \end{bmatrix} + \mathbf{L}\mathbf{U}' - \begin{bmatrix} s_1 \\ s_2 \\ \cdot \\ \cdot \\ \cdot \\ s_{n-1} \\ s_n \end{bmatrix} = \mathbf{L}\underset{\sim}{\alpha}' - \mathbf{G}^N\mathbf{x}_o \tag{19}$$

$$g\begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ 1 \end{bmatrix} - \mathbf{L}\ \mathbf{U}' - \begin{bmatrix} s_{n+1} \\ s_{n+2} \\ \cdot \\ \cdot \\ \cdot \\ s_{2n-1} \\ s_{2n} \end{bmatrix} = \mathbf{L}\underset{\sim}{\alpha}' + \mathbf{G}^N\mathbf{x}_o \tag{19}$$

$$\mathbf{U}' + \begin{bmatrix} c_1 \\ \cdot \\ \cdot \\ \cdot \\ c_{Nr} \end{bmatrix} = \underset{\sim}{\alpha}' + \underset{\sim}{\beta}' \tag{19}$$

These equations can be combined and written as a single vector equation:

$$\mathbf{H}\mathbf{z} = \mathbf{W} \tag{20}$$

where $\mathbf{z}$ is a $(1 + 2Nr + 2n)$ dimensional vector containing g, Nr dimensional control vectors $\mathbf{u}_k'$, $2n$ state slack variables, and Nr control slack variables, in that order. The $(2n + Nr)$ dimensional vector $\mathbf{W}$ contains $\mathbf{L}\underset{\sim}{\alpha}' - \mathbf{G}^N\mathbf{x}_o$, $-\mathbf{L}\underset{\sim}{\alpha}' + \mathbf{G}^N\mathbf{x}_o$, and $\underset{\sim}{\alpha}' + \underset{\sim}{\beta}'$, in that order. The matrix $\mathbf{H}$ has dimensions $(2n + Nr) \times (1 + 2Nr + 2n)$ and is set up as

Equation (20) represents a set of $(2n + Nr)$ scalar equations in $(1 + 2Nr + 2n)$ unknowns; that is, there are $(1 + Nr)$ more unknowns than equations. There are several linear programming algorithms for solving such a system (5). The basic technique is first, to find a set of $(1 + Nr)$ unknowns which, when set equal to zero, will leave a solvable set of $(2n + Nr)$ equations in the same number of unknowns. The solution of these equations must, in addition, be a feasible solution, that is, the solution must yield all non-negative quantities. The set of equations is solvable if the determinant of the coefficients of the $(2n + Nr)$ variables does not vanish. The value of the quantity being minimized, a linear sum of the unknown variables is recorded. In this case the quantity being minimized is g. A decrease in g is sought by setting one of the $(2n + Nr)$ variables equal to zero and by replacing it with one of the $(1 + Nr)$ variables which were originally set equal to zero. The new set of $(2n + Nr)$ equations produces a new feasible solution, the new value of g is recorded, and the procedure continues until no further improvement in g can be obtained. Systematic schemes have been developed for choosing successive sets of $(2n + Nr)$ equations which will produce successive feasible solutions and monotonic convergence of g to a minimum (5).

If N is sufficiently large, so that g can be reduced to zero, all the s_i will be zero. This follows from

$$\begin{gathered} g + E_i - s_i = 0 \\ g - E_i - s_{i+n} = 0 \end{gathered}$$

If $g = 0$ and both s_i and s_{i+n} are non-negative, this can be true if, and only if, $s_i = s_{i+n} = E_i = 0$. However, from the composition of $\mathbf{H}$, it follows that at most $(n + 1)$ of the s_i's can be included in the set of $(1 + Nr)$ variables set equal to zero. Furthermore, this maximum number of zero s_i's is possible if at most one case of $s_i = s_{i+n} = 0$

is included. (If any more of the s_i's are set equal to zero, the determinant of the coefficients of the remaining equations will vanish.) The zero values for the remaining s_i's arise in the solution of the $(2n + Nr)$ equations.

Of the $(1 + Nr)$ variables to be set equal to zero, it has been shown that at most $(n + 1)$ of them can be s_i's. The remaining $(Nr - n)$, or more, variables must be chosen from the Nr control variables and Nr control slack variables. If a variable set equal to zero is a control variable, then that control variable is on its lower bound. If a control slack variable is set equal to zero, its corresponding control variable is forced to be on its upper bound. Therefore, of the Nr scalar control variables, at least $(Nr - n)$ of them will be at their bounds, and the remaining control variables are free to be anywhere within the closed interval formed by their bounds.

These results can now be summarized in a theorem.

Theorem: If an n dimensional state vector whose transient behavior is given by Equation (6) can be brought from an initial state $\mathbf{x}_0$ to the origin in a finite number of sampling periods N, and, if the r dimensional control vectors, $\mathbf{u}_k$, $k = 0, 1, \ldots, (N - 1)$, are bounded in accordance with Equation (12), then the following exists.

1. The desired control sequence $\{\mathbf{u}_k\}$ arises as a solution to the linear programming problem associated with Equation (20), where the quantity being minimized g is the first element of $\mathbf{z}$.
2. Of the Nr control variables contained in the N control vectors $\mathbf{u}_k$, at least $(Nr - n)$ of them will be on their bounds if $\min(g) = 0$, and the remaining ones may appear anywhere within the closed interval formed by their respective bounds.
3. The minimum N, for which $g = 0$ is the solution of the linear program, is the minimum number of sampling periods sought in Problem 2, and the $\{\mathbf{u}_k\}$ associated with this N is the corresponding $\{\mathbf{u}_k\}$ sought in the problem.

Two further problems similar to Problem 2 should also be mentioned here. These are the problems where $\mathbf{U}$ is bounded above or below, but not both. If $\mathbf{U}$ is bounded only below, then the control-vector constraint equation is dropped from Equation set (19). In Equation (20), $\mathbf{H}$ becomes a $(2n) \times (1 + Nr + 2n)$ matrix, $\mathbf{z}$ becomes a $(1 + Nr + 2n)$ dimensional vector, and $\mathbf{W}$ becomes a $(2n)$ dimensional vector. The solution proceeds as before except that the final equation set has n equations in Nr unknown control variables (the components of $\mathbf{U}'$). This leads to $(Nr - n)$ controls being zero, which is equivalent to being on their lower bounds, and n controls are free to be anywhere above.

The second problem similar to Problem 2 would bound $\mathbf{U}$ from above only. This can be made equivalent to the lower bound problem.

$$\mathbf{U} \leq \underset{\sim}{\beta'}$$

implies

$$0 \leq \underset{\sim}{\beta'} - \mathbf{U} = \mathbf{U}''$$

Combine this with Equation (15) instead of Equation (17). The solution to the linear programming problem will be the same as for the lower bound (only) problem, except that $\mathbf{U}''$ will be obtained instead of $\mathbf{U}'$. There will be $(Nr - n)$ components of $\mathbf{U}''$ set equal to zero, which is equivalent to the corresponding components of $\mathbf{U}$ being on their upper bounds.

There are some additional points in connection with Problem 2 and its variations which are suitable for discussion. First, the linear programming solution is, in general, not unique. Corresponding to the minimum N, there will generally exist more than one control sequence $\{\mathbf{u}_k\}$ which will bring the system to the origin in that number of sampling periods. However, the linear program generates only one such sequence. The multiplicity of solutions is discussed more fully in Appendix B.* A second point of interest is that the linear programming method can be extended to handle constraints on the state variables and their time derivatives. Details for handling this extension are given in Appendix A (see page 152).

NUMERICAL EXAMPLE

To illustrate the computational procedure just described, a study was made of an absorption tower of the plate variety. Lapidus and his co-workers (*10*) have used this system to illustrate the application of dynamic programming to sampled-data control systems. The problem considered by these authors was not to minimize time, but the quantity P where

$$P = \sum_{j=1}^{N} (\mathbf{x}^d - \mathbf{x}_j)^T \mathbf{Q} (\mathbf{x}^d - \mathbf{x}_j), \; N \to \infty$$

The effect of minimizing P is to drive $\mathbf{x}$ to $\mathbf{x}^d$, but in a different manner than when N is being minimized. This difference was briefly discussed earlier.

A material balance around the m^{th} plate of the tower yields

$$L(x_{m-1} - x_m) + G(y_{m+1} - y_m) = H\frac{dy_m}{dt} + h\frac{dx_m}{dt} \tag{21}$$

A linear equilibrium relationship is assumed; that is

$$y_m = ax_m + b \tag{22}$$

Combination of Equations (21) and (22) yields

$$\frac{dx_m}{dt} = \frac{d}{e}x_{m-1} - \left(\frac{d+1}{e}\right)x_m + \frac{1}{e}x_{m+1}; \quad m = 1, 2, \ldots, n \tag{23}$$

where $d = L/Ga$ and $e = (Ha + h)/Ga$. It is assumed that the system is initially at a steady state given by

$$x_m(0) = x_{mo};\; m = 1, 2, \ldots, n \tag{24}$$

There are two boundary conditions to be met, namely, the inlet liquid and vapor compositions:

$$\begin{aligned} x_o(t) &= u_1(t) \\ y_{n+1}(t) &= \overline{u}_2(t) \end{aligned} \tag{25}$$

Equations (23) to (25) may be combined to yield a vector differential equation and associated boundary condition:

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ \mathbf{x}(0) &= \mathbf{x}_o \end{aligned} \tag{26}$$

where $\mathbf{A}$ is a tridiagonal matrix given by

$$\mathbf{A} = \begin{bmatrix} -\left(\frac{d+1}{e}\right) & \frac{1}{e} & & & \bigcirc \\ \frac{d}{e} & -\left(\frac{d+1}{e}\right) & \frac{1}{e} & & \\ & \frac{d}{e} & \ddots & \ddots & \\ & & \ddots & \ddots & \frac{1}{e} \\ \bigcirc & & & \frac{d}{e} & -\left(\frac{d+1}{e}\right) \end{bmatrix}$$

* See footnote on page 144.

Then $(n \times 2)$ matrix $\mathbf{B}$, the n dimensional vector $\mathbf{x}$, and the two-dimensional vector $\mathbf{u}$ are given, respectively, by

$$\mathbf{B} = \begin{bmatrix} \frac{d}{e} & 0 \\ 0 & 0 \\ \cdot & \cdot \\ \cdot & \cdot \\ 0 & 0 \\ 0 & \frac{1}{e} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_{n-1} \\ x_n \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

where

$$u_2 = (\overline{u_2} - b)/a$$

Equation (26) is simply the linear vector differential equation equivalent to Equation (2). In (*10*) it was assumed that the system was at a steady state $\mathbf{x}_o$, and it was desired to force the system to a new steady state $\mathbf{x}^d$. A sampled-data control sequence $\{\mathbf{u}_j\}$ was sought which would minimize the quantity P as N became infinite, where

$$P = \sum_{j=1}^{N} (\mathbf{x}^d - \mathbf{x}_j)^T \mathbf{Q} (\mathbf{x}^d - \mathbf{x}_j)$$

and $\mathbf{x}_j$ is given by Equation (5) if N is replaced by j in that equation. A sampling period of 1 min. was used, and $\mathbf{Q}$ was set equal to the identity matrix. The problem parameters, the initial condition $\mathbf{x}_o$, and the final condition $\mathbf{x}^d$ are all given in the reference mentioned and may be consulted by the reader if needed. As a point of interest, $\mathbf{x}_o$ and $\mathbf{x}^d$ were both chosen as equilibrium or steady state concentrations.

The dynamic optimization carried out succeeded in bringing $\mathbf{x}$ to $\mathbf{x}^d$ much faster than would have been the case, had a single control $\mathbf{u}^d$ been applied, where $\mathbf{u}^d$ is the control needed to maintain $\mathbf{x}$ at the steady state $\mathbf{x}^d$. The latter transient response was called by Lapidus et al. the normal transient response. The transient response obtained by minimizing P is shown for x_3 and x_6 in Figures 1 and 2, respectively. These components of $\mathbf{x}$ were chosen to illustrate the effect of the controls on plate concentrations far from and near the controls.

Use will now be made of the linear programming approach to demonstrate how the concentration vector $\mathbf{x}$ may be brought to its desired state in minimum time. When the variables in Equation (21) are defined in the manner of Lapidus et al., the compositions are bounded by zero below and have no bounds above. This assumes, of course, that solubility limits are neglected. If the variables in Equation (21) were defined in such a way that the compositions were weight fractions, then the compositions would be bounded by unity above. Computer programs were written for variables defined in both ways. This was done to study the use of linear programming in cases where the controls are bounded on either one or both sides. The programs were written, for the most part, in FORTRAN and run on a 32K IBM 7090/7094 computer.

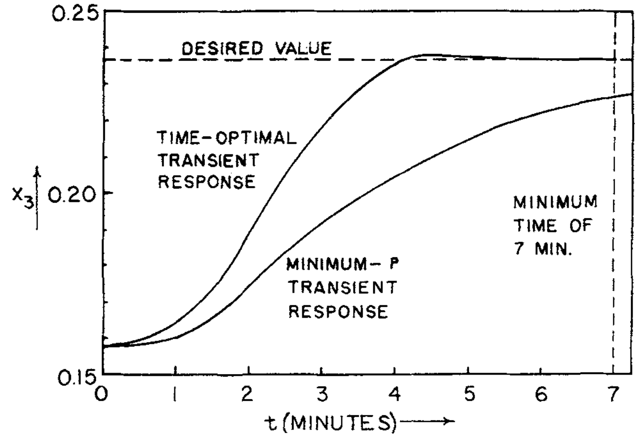

Fig. 1. Minimum P and time-optimal transient responses of x_3 in absorber. (Sampling period = 1 min., controls bounded below.)

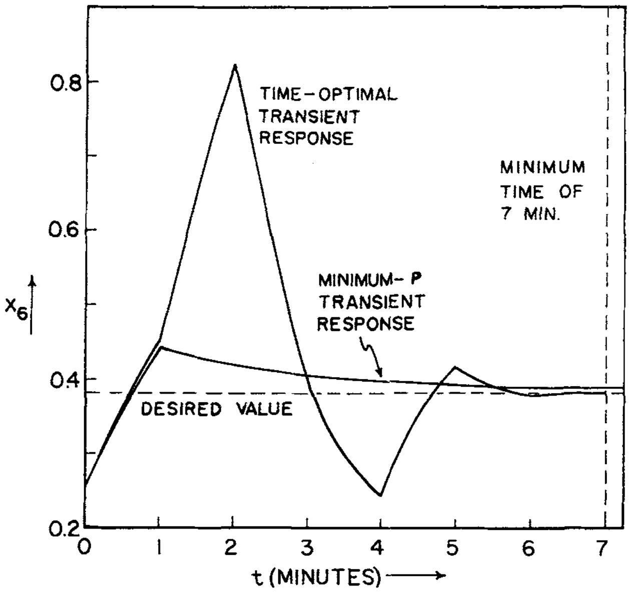

Fig. 2. Minimum P and time-optimal transient responses of x_6 in absorber. (Sampling period = 1 min., controls bounded below.)

Once the quantities required by Equation (20) are determined and put in their proper storage locations, the linear programming subroutine is entered. The subroutine, SHARE program RS MSUB, solves for that $\mathbf{z}$ which minimizes a linear sum of the elements of $\mathbf{z}$; in this case, the first element g is the only element in the summation. The elements of $\mathbf{z}$ which give the required control vectors $\mathbf{u}_k'$ are then used with Equation (4) to give the transient response of $\mathbf{x}$. (In this problem $\mathbf{u}_k \equiv \mathbf{u}_k'$, since $\mathbf{u}_k$ has a natural lower bound of 0.) It should be pointed out that $\alpha_1 = 0$ in this problem. In some problems an α_i or β_i equal to zero could make the origin of the state vector space unreachable for any N, however large. In the absorber problem such an effect was not present.

The first computer program was written with the variables of Equation (21) defined as Lapidus defined them; that is, the controls were bounded from below only. This required modifying the vectors and matrix of Equation (20), as mentioned earlier. Table 1 lists the main cases investigated, giving the sampling period Δt, the number of sampling periods N, and g, the maximum absolute composition error after N sampling periods. Seven-place accuracy was obtained in the calculations. A zero entry in Table 1 implies a value of g less than 10^{-7}. The running time per case on the IBM 7090/7094 computer was 10 to 12 sec. This included the time needed to calculate the transient response of the composition vector once the required control had been obtained.

It can be seen in Table 1 that for sampling periods of 1.0, 0.5, and 0.25 min., the minimum number of sampling periods required to reach the desired state was 7, 11, and 20, respectively. Not unexpected is the result that the minimum time required to reach the desired state decreases as the sampling period decreases. Another result is that for a fixed total time, g can be made smaller as

TABLE 1. MAXIMUM COMPOSITION DEVIATION FOR THE ABSORBER WITH ONLY LOWER BOUNDS ON CONTROLS

Δt, min.	N	g, lb. solute/ lb. absorbent
1.0	5	1.876500×10^{-4}
1.0	6	1.766656×10^{-5}
1.0	7	0
0.5	10	2.918065×10^{-5}
0.5	11	0
0.25	14	1.399807×10^{-3}
0.25	16	3.434860×10^{-4}
0.25	18	9.534832×10^{-5}
0.25	19	1.836266×10^{-5}
0.25	20	0

the sampling period decreases. This can be seen in Table 1 by observing g for the three cases of $N\Delta t = 5$ min. These trends are reasonable, since as N increases (but $N\Delta t$ fixed), there is more freedom allowed in switching the controls, and hence it should be possible to approach nearer to the desired state. For all cases in which g was forced to zero, all but six of the elements of $\mathbf{U}'$ were zero. This agrees with the result obtained earlier that at most n controls are not on their bounds when $g = 0$. In Tables 2, 3, and 4 the three time-optimal control sequences are given. Figures 1 and 2 compare the time-optimal transient responses of x_3 and x_6 with a sampling period of 1 min. to the corresponding minimum P transient response of Lapidus. It can be seen that the time-optimal response of each composition reaches its desired value well in advance of the minimum P response. Lapidus reports a deviation in x_6 of 4.0×10^{-5} after twenty sampling periods. This compares with no deviation after seven sampling periods with the time-optimal response.

Figures 3 to 6 give the time-optimal responses associated with sampling periods of 0.5 and 0.25 min. The compositions x_1 and x_6 experience larger fluctuations about their final values than do the intermediate compositions. This can be attributed to the proximity of x_1 and x_6 to the controls u_1 and u_2, respectively; in other words, x_1 and x_6 damp the response of $x_2 - x_5$ to u_1 and u_2, but are not themselves damped.

The second computer program written for the absorber utilized Equation (21) also but considered the x_i's and y_i's to be weight fractions. This implied that L and G were the total liquid and vapor flow rates, respectively. Similarly, H and h were interpreted as total holdups. The same numerical values of the parameters were used as in the first program, although the parameters had different physical units than before. The inlet concentrations were bounded in accordance with

TABLE 2. THE TIME-OPTIMAL CONTROL AND TRANSIENT RESPONSE OF THE ABSORBER WITH ONLY LOWER BOUNDS ON CONTROLS
Sampling period: 1 min.

t, min.	$u_1, = x_o$	$u_2, = y_7/a$
0	0.5155963	0.7865756
1	0	1.5306913
2	0	0
3	0	0
4	0	0.6403382
5	0	0.3818456
6	0	0.4183832

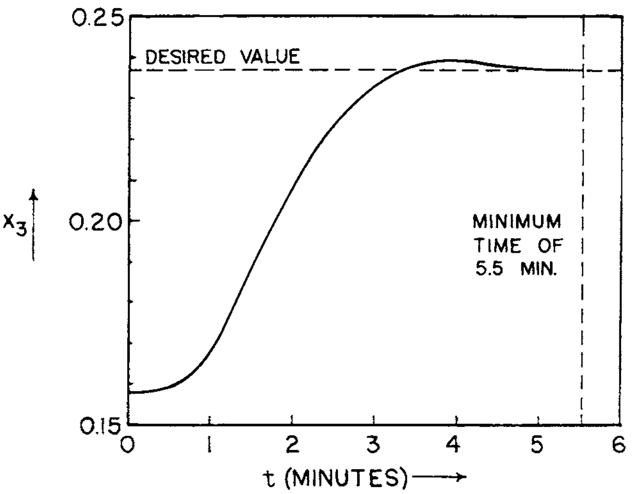

Fig. 3. Time-optimal transient response of x_3 in absorber. (Sampling period = 0.5 min., controls bounded below.)

$$0 \leq x_o \leq 1.0$$

$$0 \leq y_7 \leq 1.0$$

This meant that the controls were, in turn, bounded in accordance with

$$0 \leq u_1 \leq 1.0$$

$$0 \leq u_2 \leq 1.3888889$$

since $u_1 = x_o$, $u_2 = y_7/a$, and $a = 0.72$. Because both u_1 and u_2 again have lower bounds of zero, $\mathbf{U} = \mathbf{U}'$. Thus $\mathbf{U}$ is again obtained directly from the linear program solution for $\mathbf{z}$ in Equation (20). Computer running time per case was again 10 to 12 sec.

Table 5 indicates the principal cases considered. Sampling periods of 1 and 0.5 min. were used. Computer storage limitations prevented a 0.25-min. sampling period from being considered, since the minimum N for this case was so large. A maximum of sixteen sampling periods could be handled without the matrices and vectors involved in the solution overflowing the 32K storage capacity of the computer. This corresponded to a (44×77) dimensional $\mathbf{H}$ matrix. The minimum times required to reach the desired state were 7 and 6 min. for the two different sampling periods. This compares with 7 and 5.5 min. for the problem with only lower bounds on the controls. The time-optimal control sequence is given in Table 6. Both Tables 2 and 6 represent minimum times of 7 min. The reason for this can be seen by observing the control sequence of Table 2, where the controls were bounded below only. All but one control was within the bounds placed on the controls of Table 6; the one control which exceeded the bound of the other problem did so by a small amount. Thus when bounds were placed on both sides of the controls, small adjustments could be made in all the controls of Table 2, bringing the one inadmissible control down to its upper bound and still transferring the composition vector to its desired state in 7 min. (See Appendix B for a discussion on the occurrence of still other 7-min. solutions.)

For a sampling period of 0.5 min. and only a lower bound on the control, the minimum time was 5.5 min. But the required control sequence (Table 3) contained a control which was considerably above the upper bound of the other problem. In meeting the constraint a radically different control policy evolved (Table 7), and the desired state could not be reached in less than a minimum

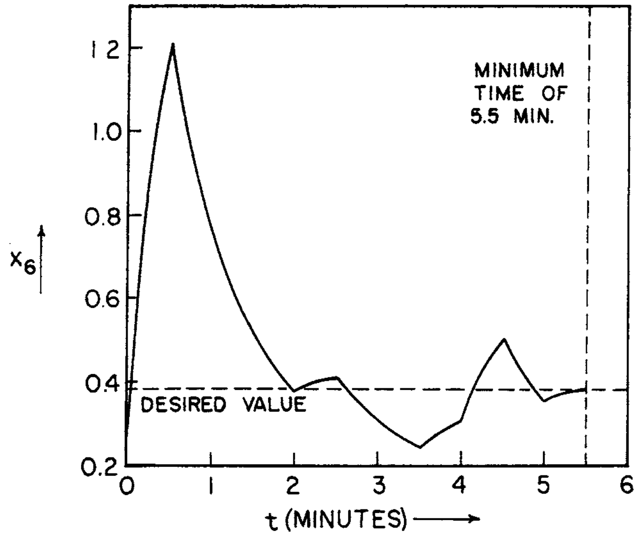

Fig. 4. Time-optimal transient response of x_6 in absorber (Sampling period = 0.5 min., controls bounded below.)

time of 6 min. The transient response of x_6 for this case is shown in Figure 7. The response of x_3 is similar to Figure 3 and is, therefore, not shown. The control sequence of Table 4 (with a sampling period of 0.25 min.) had several large controls, such that, if the problem could have been solved with the controls constrained above, that is, if more storage capacity had been available, one would expect a larger minimum time, as was the case for a 0.5-min. sampling period.

For each case listed in Table 5 in which $g = 0$ was the result of the linear program, six elements of U' were not on their bounds. This agrees with the theorem stated earlier. It is noted again that for $N\Delta t$ fixed and $g \neq 0$, g decreases as Δt decreases. By comparing corresponding cases in Tables 1 and 5, it is also seen that the additional constraints have increased g (unless $g = 0$). This is to be expected, since additional constraints allow for less freedom in picking a control sequence to drive the compositions to their desired state. The maximum composition deviation g is therefore greater.

From a practical point of view, it may be desirable to limit the extreme oscillations of x_6 by placing constraints

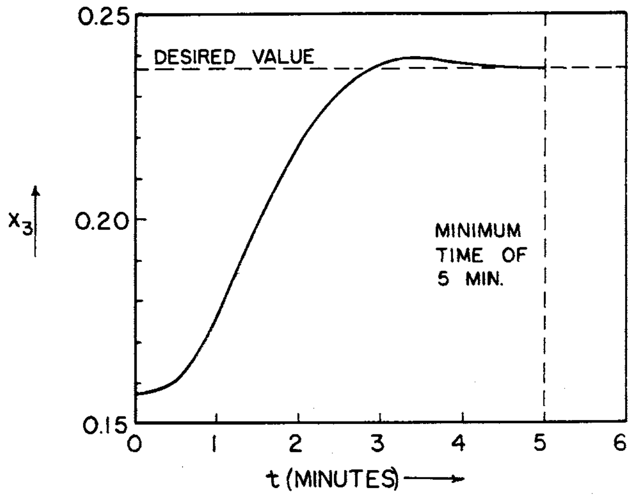

Fig. 5. Time-optimal transient response of x_3 in absorber. (Sampling period = 0.25 min., controls bounded below.)

TABLE 3. THE TIME-OPTIMAL CONTROL AND TRANSIENT RESPONSE OF THE ABSORBER WITH ONLY LOWER BOUNDS ON CONTROLS
Sampling period: 0.5 min.

t, min.	$u_1, = x_o$	$u_2, = y_7/a$
0	0	4.1781770
0.5	0.6248582	0
1.0	0	0
1.5	0	0
2.0	0	0.4774132
2.5	0	0
3.0	0	0
3.5	0	0.4296610
4.0	0	1.1514743
4.5	0	0
5.0	0	0.4803640

TABLE 4. THE TIME-OPTIMAL CONTROL AND TRANSIENT RESPONSE OF THE ABSORBER WITH ONLY LOWER BOUNDS ON CONTROLS
Sampling period: 0.25 min.

t, min.	$u_1, = x_o$	$u_2, = y_7/a$
0	0	8.5543973
0.25	1.3366968	0
0.50	0	0
	zero controls	
3.00	0	0
3.25	0	1.4506953
3.50	0	0
3.75	0	1.2451392
4.00	0	0.7741332
4.25	0	0
4.50	0	0
4.75	0	0.6444009

TABLE 5. MAXIMUM COMPOSITION DEVIATION FOR THE ABSORBER WITH CONTROLS BOUNDED ABOVE AND BELOW

Δt, min.	N	g, lb. solute/ lb. liquid
1.0	5	5.248003×10^{-4}
1.0	6	3.644988×10^{-5}
1.0	7	0
0.5	10	1.503775×10^{-4}
0.5	11	2.551380×10^{-5}
0.5	12	0

TABLE 6. THE TIME-OPTIMAL CONTROL AND TRANSIENT RESPONSE OF THE ABSORBER WITH CONTROLS BOUNDED ABOVE AND BELOW
Sampling period: 1 min.

t, min.	$u_1, = x_o$	$u_2, = y_7/a$
0	0.4467076	0.9620660
1	0	1.3888889
2	0.0185291	0
3	0	0
4	0	0.6488019
5	0	0.3796857
6	0	0.4185208

TABLE 7. THE TIME-OPTIMAL CONTROL AND TRANSIENT RESPONSE OF THE ABSORBER WITH CONTROLS BOUNDED ABOVE AND BELOW
(Sampling period: 0.5 min.

t, min.	$u_1, = x_o$	$u_2, = y_7/a$
0	0	0.8898799
0.5	0.8624622	1.3888889
1.0	0	1.3888889
1.5	0	1.0570987
2.0	0	0
2.5	0	0
3.0	0	0
3.5	0	0
4.0	0	0.4576379
4.5	0	1.1507888
5.0	0	0
5.5	0	0.4798705

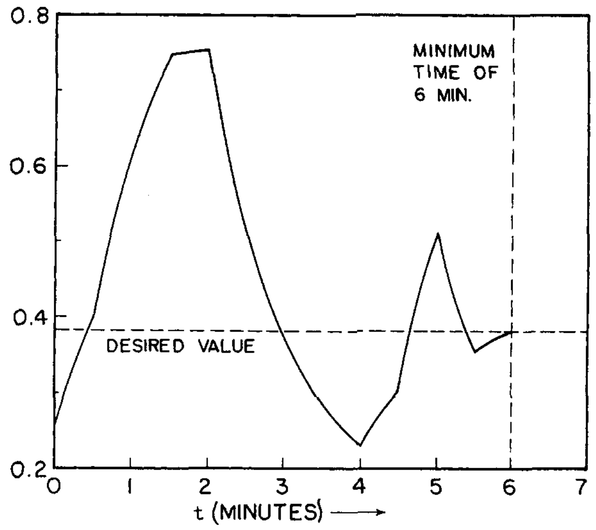

Fig. 7. Time-optimal transient response of x_6 in absorber. (Sampling period = 0.5 min., controls bounded above and below.)

on it and all the other x_i. This will necessarily increase the minimum time. It may also be advantageous to require that $\dot{x}$ be bounded at all sampling times, in order to smooth the transient response of x. This restriction would also increase the minimum time. Both of these specifications can, however, be built into the linear programming model. Appendix A explains how this can be carried out.

CONCLUSIONS

The time-optimal control of discrete-time linear systems has been investigated in the presence of constraints on the controls. A theorem has been established which provides for a solution to the problem with linear programming used. It allows for at most n controls not being on their bounds, where n is the order of the system. The theorem has been utilized to obtain the time-optimal control of an absorption tower with two control variables. Several effects were studied. Among these was the effect of varying the sampling period. The effect of bounding the controls from one or both sides was also illustrated. The matrices and vectors encountered in the method have large dimensions which are increasing functions of n, r, and N, r being the number of control variables and N being the number of sampling periods. It is, therefore, not surprising that these three quantities must not be too large if modern computer storage capacities are not to be exceeded. As storage capacities increase, the values of n, r, and N for which solutions are obtainable will increase. It will also be possible to obtain results in the presence of constraints on the state vector, the theory of which is given in Appendix A. The nonunique nature of the linear programming solutions is discussed in Appendix B, and the extension of the method to distributed-parameter system is presented in Appendix C.

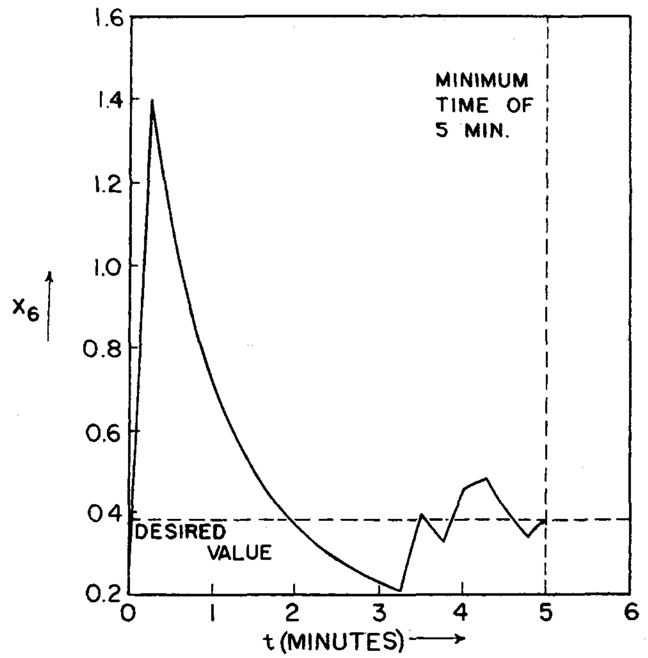

Fig. 6. Time-optimal transient response of x_6 in absorber. (Sampling period = 0.25 min., controls bounded below.)

ACKNOWLEDGMENT

The authors gratefully acknowledge the financial support in the form of fellowships by the National Science Foundation (Grant GP-509) and the Ford Foundation. The use of computer facilities at Princeton University was obtained with the support of National Science Foundation Grant GP-579. Appreciation is also expressed to O. L. Mangasarian of the Shell Development Company, who pointed out to the first author the work of Zadeh and Whalen (*16*).

NOTATION

A = matrix of a_{ij} coefficients
a = constant in linear equilibrium relationship
a_{ij} = state variable coefficient in set of linear, ordinary differential equations
B = matrix of b_{ij} coefficients
b = constant in linear equilibrium relationship
b_{ij} = control coefficient in set of linear, ordinary differential equations
c_i = control slack variable
d = L/Ga
E = error vector
e = $(Ha + h)/Ga$
G = $\exp(\mathbf{A}\,\Delta t)$
G = gas flow rate, lb./min.
g = $\max_i |E_i|$
H = linear programming coefficient matrix
H = vapor holdup per plate, lb.
h = liquid holdup per plate, lb.

$\mathbf{J} \quad = \int_0^{\Delta t} \exp(\mathbf{A}\, s)\, \mathbf{B}\, d\, s$

k = number of sampling periods
$\mathbf{L}$ = matrix defined by Equation (9)
$\mathbf{L}_{N-1-k} = \mathbf{G}^{N-1-k}\, \mathbf{J}$
L = liquid flow rate, lb./min.
N = total number of sampling periods
n = order of system
P = performance criterion of Lapidus et al.
Q = number of unknowns
$\mathbf{Q}$ = weighting matrix in expression for P
r = number of control variables
S = number of algebraic equations
s_i = state slack variable
$\mathbf{U}$ = control vector containing $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{N-1}$
$\mathbf{U}' = \mathbf{U} + \underset{\sim}{\alpha}'$
$\mathbf{U}'' = \underset{\sim}{\beta}' - \mathbf{U}$
$\mathbf{u}$ = r dimensional control vector
$\mathbf{W}$ = linear programming constant vector
$\mathbf{x}$ = n dimensional state vector
x_m = liquid composition on m^{th} tray, lb. solute/lb. absorbent
y_m = vapor composition on m^{th} tray, lb./solute/lb. inert
$\mathbf{z}$ = linear programming unknown vector

Greek Letters

$-\underset{\sim}{\alpha}$ = lower bound on $\mathbf{u}$
$\underset{\sim}{\alpha}'$ = vector containing N α's
$\underset{\sim}{\beta}$ = upper bound on $\mathbf{u}$
$\underset{\sim}{\beta}'$ = vector containing N β's
Δt = sampling period

Subscripts

k = number of sampling period
m = tray number
mo = initial condition on m^{th} tray

Superscripts

d = desired value
T = transpose

LITERATURE CITED

1. Bellman, R., "Dynamic Programming," Princeton Univ. Press, N. J. (1957).
2. Desoer, C. A., and J. Wing, *IRE Trans. Automatic Control*, **6**, 111-125 (1961).
3. *Ibid.*, 5-15.
4. Desoer, C. A., and J. Wing, *J. Franklin Inst.*, **272**, 208-228 (1961).
5. Gass, S. I., "Linear Programming," McGraw-Hill, New York (1958).
6. Ho, Y. C., *J. Basic Eng.*, **83**, 53-58 (1961).
7. Kalman, R. E., in "Proc. First International Congress of Automatic Control," Moscow (1960).
8. ———, *IRE Wescon Convention Record*, Pt. 4, 130-135 (1957).
9. Lesser, H. A., Ph.D. thesis, Princeton Univ., N. J. (1964).
10. Lapidus, Leon, Eugene Shapiro, Saul Shapiro, and R. E. Stillman, *A.I.Ch.E. J.*, **7**, 288-294 (1961).
11. Mangasarian, O. L., private communication.
12. Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze, and F. F. Mishenko, "The Mathematical Theory of Optimal Processes," (authorized translation from the Russian), Interscience, New York (1962).
13. Propoi, A. I., *Avtomat. Telemekh.*, **24**, 912-920 (1963); English translation, *Automation Remote Control*, **24**, 837-844 (1963).
14. Rozonoer, L. I., *Avtomat. Telemekh.*, **20** (1959); English translation, *Automation Remore Control*, **20**, 1288-1302, 1405-1421, 1517-1532 (1960).
15. Torng, H. C., *J. Franklin Inst.*, **277**, 28-44 (1964).
16. Zadeh, L. A., and B. H. Whalen, *IRE Trans. Automatic Control*, **AC-7**, 45-46 (1962).

APPENDIX A: CONSTRAINTS ON BOTH CONTROL AND STATE VECTORS

To constrain the state vector, Equation (5) is rewritten for ν sampling periods:

$$\mathbf{x}_\nu = \mathbf{G}^\nu \mathbf{x}_0 + \sum_{k=0}^{\nu-1} \mathbf{L}_{\nu-1-k}\, \mathbf{u}_k;\ \nu = 1, 2, \ldots, N-1 \qquad \text{(A1)}$$

The $2n(N-1)$ slack variables φ_i are introduced, along with the constraint vectors $-\underset{\sim}{\lambda}$ and $\underset{\sim}{\theta}$. The elements of $\underset{\sim}{\lambda}$ and $\underset{\sim}{\theta}$ are all positive and correspond to the lower and upper limits, respectively, of the elements of $\mathbf{x}$. Then the condition that $\dot{\mathbf{x}}_\nu$ be within its constraints (for $1 \leqq \nu \leqq N-1$) is equivalent to

$$\mathbf{G}^\nu \mathbf{x}_0 + \sum_{k=0}^{\nu-1} \mathbf{L}_{\nu-1-k}\, \mathbf{u}_k - \underset{\sim}{\theta} + \begin{bmatrix} \varphi_{2n\nu-2n+1} \\ \cdot \\ \cdot \\ \cdot \\ \varphi_{2n\nu-n} \end{bmatrix} = 0$$

$$-\mathbf{G}^\nu \mathbf{x}_0 + \sum_{k=0}^{\nu-1} \mathbf{L}_{\nu-1-k}\, \mathbf{u}_k - \underset{\sim}{\lambda} + \begin{bmatrix} \varphi_{2n\nu-n+1} \\ \cdot \\ \cdot \\ \cdot \\ \varphi_{2n\nu} \end{bmatrix} = 0$$

Once the vectors $\underset{\sim}{\lambda}$, $\underset{\sim}{\theta}$, and $\mathbf{G}^\nu \mathbf{x}_0$ are transferred to the other side and each $\mathbf{u}_k$ is converted to $\mathbf{u}_k'$, the equations may be combined with those of Equation (19). This, of course, will change the makeup of $\mathbf{H}$, $\mathbf{z}$, and $\mathbf{W}$ in Equation (20). The dimensions of $\mathbf{H}$ will become $[2n + Nr + 2n(N-1)] \times [1 + 2Nr + 2n + 2n(N-1)]$, which reduces to $[N(r + 2n)] \times [1 + 2N(r + n)]$.

Constraints can be placed on $\dot{\mathbf{x}}$ at the beginning and end of each sampling period. At the sampling time $\nu\Delta t$, $\dot{\mathbf{x}}$ has both a left- and right-hand limit. The left-hand limit is given by

$$\dot{\mathbf{x}}_{\nu-} = \mathbf{A}\, \mathbf{x}_\nu + \mathbf{B}\, \mathbf{u}_{\nu-1} \qquad \text{(A2)}$$

and the right-hand limit by

$$\dot{\mathbf{x}}_{\nu+} = \mathbf{A}\, \mathbf{x}_\nu + \mathbf{B}\, \mathbf{u}_\nu \qquad \text{(A3)}$$

Equations (A1) and (A2) may be combined to yield

$$\dot{\mathbf{x}}_{\nu-} = \mathbf{A}\, \mathbf{G}^\nu \mathbf{x}_0 + \mathbf{A} \sum_{k=0}^{\nu-1} \mathbf{L}_{\nu-1-k}\, \mathbf{u}_k + \mathbf{B}\, \mathbf{u}_{\nu-1} \qquad \text{(A4)}$$

If $-\underset{\sim}{\sigma}$ and $\underset{\sim}{\rho}$ are the constraint vectors for $\dot{\mathbf{x}}_{\nu-}$ (for $1 \leqq \nu \leqq N$), and if $2n(N-1)$ slack variables γ_i are introduced, then the condition that $\dot{\mathbf{x}}_{\nu-}$ be within its constraints (for $1 \leqq \nu \leqq N$) is equivalent to

$$\mathbf{A}\, \mathbf{G}^\nu \mathbf{x}_0 + \mathbf{A} \sum_{k=0}^{\nu-1} \mathbf{L}_{\nu-1-k}\, \mathbf{u}_k + \mathbf{B}\, \mathbf{u}_{\nu-1} - \underset{\sim}{\rho} + \begin{bmatrix} \gamma_{2n\nu-2n+1} \\ \cdot \\ \cdot \\ \cdot \\ \gamma_{2n\nu-n} \end{bmatrix} = 0$$

$$-\mathbf{A}\, \mathbf{G}^\nu \mathbf{x}_0 - \mathbf{A} \sum_{k=0}^{\nu-1} \mathbf{L}_{\nu-1-k}\, \mathbf{u}_k - \mathbf{B}\, \mathbf{u}_{\nu-1} - \underset{\sim}{\sigma} + \begin{bmatrix} \gamma_{2n\nu-n+1} \\ \cdot \\ \cdot \\ \cdot \\ \gamma_{2n\nu} \end{bmatrix} = 0$$

The vectors $\underset{\sim}{\rho}$, $\underset{\sim}{\sigma}$, and $\mathbf{A}\, \mathbf{G}^\nu \mathbf{x}_0$ are transferred to the other side, and each $\mathbf{u}_k$ is converted to $\mathbf{u}_k'$. The resulting equations may be combined with those of Equation (19). As before, the makeup of $\mathbf{H}$, $\mathbf{z}$, and $\mathbf{W}$ in Equation (20) will be altered accordingly. Constants on $\dot{\mathbf{x}}_{\nu+}$ may be handled in analogous fashion to those on $\dot{\mathbf{x}}_{\nu-}$. Since the dimensions of $\mathbf{H}$ are increasing functions of N, r, and n, modern computer storage capacities place stringent limits on the values of N, r, and n that can be treated when the additional constraints of this Appendix are included in the formulation of the problem.

The application of linear programming to optimal control problems with other constraints on the state vector has been discussed by Propoi (*13*).

Manuscript received March 26, 1965; revision received August 30, 1965; paper accepted September 1, 1965. Paper presented at A.I.Ch.E. San Francisco meeting.