Suboptimal Control of Nonlinear Systems:

II. Constrained

A. P. J. Weber and L. Lapidus

Princeton University, Princeton, N. J.

The previous development in suboptimal control of nonlinear unconstrained systems is extended to constrained systems by the use of penalty functions in an augmented sense. This converts the constrained control problem into an unconstrained one with all of the advantages previously discussed. Numerical examples are presented.

In this second part we shall extend the development on suboptimal control of nonlinear unconstrained systems (*17*) to constrained systems by redefining the optimal L-Q (linear-quadratic) control problem to include constraints. By using penalty functions this converts a constrained control problem into an unconstrained one. The direct extension to suboptimal control of nonlinear systems with constraints then follows. To show the validity of the proposed penalty-function algorithm, numerical details will be presented for constrained linear and nonlinear problems.

The many techniques suggested for obtaining optimal control in the presence of inequality and equality constraints are discussed in the literature (*1, 9, 14*). The penalty function approach is analyzed in depth by Kelley (*7*), Lele and Jacobson (*10*), Okamura (*11*), Rothenberger (*12*), and Russell (*13*). In particular, Denn and Aris (*4*) and Chyung (*2*) applied this approach to discrete-time systems. It should be mentioned that Chyung, by means of the discrete maximum principle, treated the constrained L-Q problem in which penalty functions were used only for inequality state constraints. On the other hand, Fujisawa and Yasuda (*5*), without recourse to penalty functions, employed quadratic programming to solve an L-Q problem with a generalized quadratic performance index and inequality control constraints. Johnson and Wonham (*6*), Wonham and Johnson (*18*), Sirisena (*15*), and Deley (*3*) also considered the constrained L-Q problem, but their methods of solution did not proceed via the penalty function approach and, except in Deley's work, were restricted to a scalar control. We show that, by adapting the previous work on continuous-time penalty functions to the discrete-time case, the solution of the L-Q problem can be extended in a simple and computationally feasible fashion.

THE L-Q PROBLEM WITH CONSTRAINTS

Under discussion here is the discrete-time system

$$\mathbf{x}_{k+1} = \boldsymbol{\varphi}_k \mathbf{x}_k + \boldsymbol{\Delta}_k \mathbf{u}_k \qquad (k = 0, 1, \ldots, N-1) \quad (1)$$

with $\mathbf{x}_0 = \mathbf{x}(0)$ supplied, N fixed and the performance index

$$I\,[\mathbf{x}_0, N] = \tau \sum_{k=1}^{N} [\mathbf{x}^T{}_k \mathbf{Q}_k \mathbf{x}_k + \mathbf{u}^T{}_{k-1} \mathbf{R}_{k-1} \mathbf{u}_{k-1}] \quad (2)$$

to be minimized. First, we consider the case of general inequality constraints of the form

$$g_{i,k}(\mathbf{x}_k, \mathbf{u}_{k-1}) \leq 0 \qquad \begin{matrix} (i = 1, \ldots, m) \\ (k = 1, \ldots, N) \end{matrix} \quad (3)$$

where $g_{i,k}$ is the i^{th} component of the m-dimensional constraint vector $\mathbf{g}_k$. For simplicity, we take m to be a fixed integer for all k. Defined to have the same functional form as the vectors $\mathbf{g}_k$ is the m dimensional vector $\mathbf{g}_0(\mathbf{x}_0, \mathbf{u}_{-1})$, which does not serve as an actual constraint but which will be used later.

In g_0, we arbitrarily set the dummy fictitious control $\mathbf{u}_{-1} = \mathbf{0}$. Now we introduce a penalty function vector $\mathbf{c}_k$ with components

$$c_{i,k} \begin{cases} > 0 & \text{if} \quad g_{i,k} > 0 \\ = 0 & \phantom{\text{if}} \quad g_{i,k} \leq 0 \end{cases} \quad (4)$$

but with no $\mathbf{c}_0$ corresponding to $\mathbf{g}_0$. Corresponding to $\mathbf{c}_k$ is a penalty-weighting coefficient vector $\boldsymbol{\zeta}_k$ with components $\zeta_{i,k}$. The product of $\boldsymbol{\zeta}_k{}^T \mathbf{c}_k$ is termed the penalty and is zero only when all the constraints $\mathbf{g}_k$ are obeyed.

Calling the normal performance index I the unaugmented index, we now form a new index $\overline{I}$, called the augmented index, by attaching the sum penalties to I. Thus

$$\overline{I} = I + \tau \sum_{k=1}^{N} \boldsymbol{\zeta}_k{}^T \mathbf{c}_k \quad (5)$$

and we proceed to minimize $\overline{I}$ as if no constraints existed. This minimization is carried out via a series of minimizations in which the $\boldsymbol{\zeta}_k$ are increased resulting, at the same time in decreasing the components of $\mathbf{c}_k$. In the limit it is possible to show that as the $\boldsymbol{\zeta}_k$ are increased the two indices $\overline{I}$ and I will approach limiting values. At the same time the constraints will become progressively closer to being satisfied. In fact, as the vectors $\boldsymbol{\zeta}_k$ approach infinity, the constraints $\mathbf{g}_k \leq \mathbf{0}$ will be satisfied to whatever error is desired, and both $\overline{I}$ and I will approach the optimal performance index.

Paralleling the case of trajectory inequality constraints is the case of trajectory equality constraints with the form

$$h_{i,k}(\mathbf{x}_k, \mathbf{u}_{k-1}) = 0 \qquad \begin{matrix} (k = 1, \ldots, q) \\ (k = 1, \ldots, N) \end{matrix} \quad (6)$$

The constraint vector $\mathbf{h}_k$ of dimension q retains the same functional form at all time steps. Just as with $\mathbf{g}_0$, the q dimensional vector $\mathbf{h}_0(\mathbf{x}_0, \mathbf{u}_{-1})$, to be used later, has the same functional form as the other vectors $\mathbf{h}_k$, but does not serve as a constraint. For the penalty function vector, we now have $\mathbf{b}_k$ with components

$$b_{i,k} \begin{cases} > 0 & \text{if} \quad h_{i,k} \gtrless 0 \\ = 0 & \phantom{\text{if}} \quad h_{i,k} = 0 \end{cases} \quad (7)$$

but with no $\mathbf{b}_0$ corresponding to $\mathbf{h}_0$. Corresponding to each component $b_{i,k}$ are the components $\rho_{i,k}$, of the penalty-weighting coefficient vectors $\boldsymbol{\rho}_k$. The augmented index resulting from equality constraints is

$$\bar{I} = I + \tau \sum_{k=1}^{N} \boldsymbol{\rho}_k^T \mathbf{b}_k \tag{8}$$

and when inequality and equality constraints are considered together

$$\bar{I} = I + \tau \sum_{k=1}^{N} (\boldsymbol{\zeta}_k^T \mathbf{c}_k + \boldsymbol{\rho}_k^T \mathbf{b}_k) \tag{9}$$

To be more specific about the form of the penalty function, we first express the penalty function vectors in terms of the inequality constraints as

$$c_{i,k} = s^2_{i,k}(\mathbf{x}_k, \mathbf{u}_{k-1}) s_{i,k}(g_{i,k}) \tag{10}$$

where $\mathbf{s}_k$ is the Heaviside-unit step function vector at time step k, that is

$$s_{i,k}(\lambda_i) = \begin{cases} 1 & \lambda_i > 0 \\ 0 & \lambda_i \leq 0 \end{cases} \tag{11}$$

We now set up vectors $\boldsymbol{\eta}_k$ of additional state variables which will be equated to $\mathbf{g}_0$ and the constraint vectors $\mathbf{g}_k$. Thus

$$\eta_{i,k} = g_{i,k}(\mathbf{x}_k, \mathbf{u}_{k-1}) \quad (k = 0, \ldots, N) \tag{12}$$

or

$$\boldsymbol{\eta}_k = \mathbf{g}_k(\mathbf{x}_k, \mathbf{u}_{k-1})$$

$\boldsymbol{\eta}_0$ is the initial $\boldsymbol{\eta}_k$ vector corresponding to $\mathbf{g}_0$. This vector $\boldsymbol{\eta}_0$ does not appear in the performance index, it serves only as an initial condition from which the remaining $\boldsymbol{\eta}_k$ vectors are generated by means of recurrence relations about to be developed. From Equation (12),

$$\boldsymbol{\eta}_{k+1} = \boldsymbol{\eta}_k + \mathbf{g}_{k+1}(\mathbf{x}_{k+1}, \mathbf{u}_k) - \mathbf{g}_k(\mathbf{x}_k, \mathbf{u}_{k-1})$$

By means of the system Equation (1), $\mathbf{x}_{k+1}$ can be replaced to yield,

$$\boldsymbol{\eta}_{k+1} = \boldsymbol{\eta}_k + \mathbf{g}_{k+1}(\boldsymbol{\varphi}_k \mathbf{x}_k + \boldsymbol{\Delta}_k \mathbf{u}_k, \mathbf{u}_k) - \mathbf{g}_k(\mathbf{x}_k, \mathbf{u}_{k-1}) \quad (k = 0, \ldots, N-1) \tag{13}$$

Note in (13) that $\boldsymbol{\eta}_{k+1}$ is expressed entirely in terms of quantities evaluated at time $k\tau$ and earlier.

If $g_{i,k}$ in Equation (10) is replaced with $\eta_{i,k}$, then

$$c_{i,k} = \eta^2_{i,k} s_{i,k}(\eta_{i,k}) \quad (k = 1, \ldots, N) \tag{14}$$

Define an m^{th} order diagonal matrix $\mathbf{D}_k$ such that its diagonal elements are given in terms of $\boldsymbol{\zeta}_k$ and $\mathbf{s}_k$ as follows

$$D_{ii,k} = \zeta_{i,k} s_{i,k}(g_{i,k}) = \zeta_{i,k} s_{i,k}(\eta_{i,k}) \tag{15}$$

It is now possible to obtain the augmented performance index from relation (5). Into this equation is first substituted $\mathbf{c}_k$ from equality Equation (14). Then we have

$$\bar{I} = I + \tau \sum_{k=1}^{N} \left[\sum_{i=1}^{m} \zeta_{i,k} c_{i,k} \right]$$

$$\bar{I} = I + \tau \sum_{k=1}^{N} \left[\sum_{i=1}^{m} \zeta_{i,k} \eta^2_{i,k} s_{i,k}(\eta_{i,k}) \right]$$

Equation (15) may be employed to insert the matrix $\mathbf{D}_k$ into the above equation. When this is done, we obtain

$$\bar{I} = I + \tau \sum_{k=1}^{N} \boldsymbol{\eta}_k^T \mathbf{D}_k \boldsymbol{\eta}_k \tag{16}$$

If I is a quadratic index such as is treated in this work, then the augmented index $\bar{I}$ is also quadratic.

When it comes to trajectory equality constraints $h_{i,k}(\mathbf{x}_k, \mathbf{u}_{k-1}) = 0$, there need be little modification of the form of the above equations, except that the Heaviside step function is not used. Instead of Equation (10), we have

$$b_{i,k} = h^2_{i,k}(\mathbf{x}_k, \mathbf{u}_{k-1}) \tag{17}$$

for the penalty functions. Vectors $\boldsymbol{\beta}_k$ of extra discretized state variables which are equated to $\mathbf{h}_0$ and the N constraint vectors $\mathbf{h}_k$ are described by

$$\beta_{i,k} = h_{i,k}(\mathbf{x}_k, \mathbf{u}_{k-1}) \tag{18}$$

or

$$\boldsymbol{\beta}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{u}_{k-1})$$

$\boldsymbol{\beta}_0 = \mathbf{h}_0$ is included as an initial $\boldsymbol{\beta}_k$ vector. The vector $\boldsymbol{\beta}_0$ does not appear in the performance index, but is employed as an initial condition from which the other $\boldsymbol{\beta}_k$ vectors are calculated.

Just as in Equation (13), recurrence relations among the $\boldsymbol{\beta}_k$ vectors are

$$\boldsymbol{\beta}_{k+1} = \boldsymbol{\beta}_k + \mathbf{h}_{k+1}(\boldsymbol{\varphi}_k \mathbf{x}_k + \boldsymbol{\Delta}_k \mathbf{u}_k, \mathbf{u}_k) - \mathbf{h}_k(\mathbf{x}_k, \mathbf{u}_{k-1}) \tag{19}$$

When Equation (18) is used in (17), we come up with

$$b_{i,k} = \beta^2_{i,k} \tag{20}$$

Introduced at this point is a q^{th} order diagonal matrix $\mathbf{L}_k$, which is the analogue of matrix $\mathbf{D}_k$. Instead of the elements $D_{ii,k}$ of Equation (15), we supply the diagonal elements of $\mathbf{L}_k$

$$L_{ii,k} = \rho_{i,k} \tag{21}$$

where the $\rho_{i,k}$ are the penalty-weighting coefficients. Then

$$\boldsymbol{\rho}_k^T \mathbf{b}_k = \sum_{i=1}^{q} \rho_{i,k} \beta^2_{i,k} = \boldsymbol{\beta}_k^T \mathbf{L}_k \boldsymbol{\beta}_k$$

and the augmented index is

$$\bar{I} = I + \tau \sum_{k=1}^{N} \boldsymbol{\rho}_k^T \mathbf{b}_k = I + \tau \sum_{k=1}^{N} \boldsymbol{\beta}_k^T \mathbf{L}_k \boldsymbol{\beta}_k \tag{22}$$

If one is dealing with inequality and equality constraints, then

$$\bar{I} = I + \tau \sum_{k=1}^{N} (\boldsymbol{\eta}_k^T \mathbf{D}_k \boldsymbol{\eta}_k + \boldsymbol{\beta}_k^T \mathbf{L}_k \boldsymbol{\beta}_k) \tag{23}$$

THE AUGMENTED SYSTEM

We wish to show how the original constrained system is imbedded into an equivalent unconstrained system which is of larger dimension and is called the augmented system. Optimal control will be obtained for the original or unaugmented system by solving the augmented control problem in conjunction with the penalty function technique.

To start, the inequality constraints

$$\boldsymbol{\alpha}^- \leq \mathbf{u}_k \leq \boldsymbol{\alpha}^+ \quad (k = 0, 1, \ldots, N-1) \tag{24}$$

are considered. In the form of Equation (3), these constraints are really

$$\mathbf{g}_k(\mathbf{x}_k, \mathbf{u}_{k-1}) = \begin{cases} \mathbf{y}_k = \mathbf{u}_{k-1} - \boldsymbol{\alpha}^+ \leq \mathbf{0} \\ \mathbf{z}_k = -(\mathbf{u}_{k-1} - \boldsymbol{\alpha}^-) \leq \mathbf{0} \end{cases} \quad (k = 1, \ldots, N) \tag{25}$$

where the new vectors $\mathbf{y}_k$ and $\mathbf{z}_k$ are defined by this equation. $\mathbf{y}_0$ and $\mathbf{z}_0$ are also defined by their relation to $\mathbf{g}_0(\mathbf{x}_0, \mathbf{u}_{-1})$.

$$\mathbf{g}_0(\mathbf{x}_0, \mathbf{u}_{-1}) = \begin{cases} \mathbf{y}_0 = \mathbf{u}_{-1} - \boldsymbol{\alpha}^+ \\ \mathbf{z}_0 = -(\mathbf{u}_{-1} - \boldsymbol{\alpha}^-) \end{cases} \tag{26}$$

y_0, z_0, and the constraint vectors y_k and z_k are of the same dimension r as the control vectors. Thus m, the dimension of g_k, is $2r$.

Equation (12) for the additional state variables becomes

$$\eta_k = g_k = \begin{cases} y_k = u_{k-1} - \alpha^+ \\ z_k = -(u_{k-1} - \alpha^-) \end{cases} \quad (k = 0, \ldots, N) \tag{27}$$

so all y_k and z_k vectors are additional state variables. But for $k = 0$ and $u_{-1} = 0$,

$$\eta_0 = g_0 = \begin{cases} y_0 = -\alpha^+ \\ z_0 = \alpha^- \end{cases} \tag{28}$$

which supplies the initial conditions on the extra state variables.

For the difference Equations (13), the following is obtained

$$\eta_{k+1} = \begin{Bmatrix} y_{k+1} \\ z_{k+1} \end{Bmatrix} = \eta_k + g_{k+1} - g_k \quad (k = 0, \ldots, N-1)$$

$$= \begin{cases} y_k + u_k - \alpha^+ - (u_{k-1} - \alpha^+) \\ z_k - (u_k - \alpha^-) + (u_{k-1} - \alpha^-) \end{cases}$$

or

$$\begin{aligned} y_{k+1} &= y_k + u_k - u_{k-1} \\ z_{k+1} &= z_k - (u_k - u_{k-1}) \end{aligned} \tag{29}$$

In Equation (16) for the augmented cost appears the diagonal matrix D_k of order $m = 2r$ which contains the penalty-weighting coefficients $\zeta_{i,k}$. By Equation (15), the diagonal elements of D_k are

$$D_{ii,k} = \zeta_{i,k} s_{i,k}(\eta_{i,k}) \quad (i = 1, \ldots, m = 2r)$$

Consider the diagonal matrices E_k and H_k of order r. Let their diagonal elements be

$$\left.\begin{aligned} E_{ii,k} &= \zeta_{i,k} s_{i,k}(y_{i,k}) \\ H_{ii,k} &= \zeta_{r+i,k} s_{i,k}(z_{i,k}) \end{aligned}\right\} \quad (i = 1, \ldots, r) \tag{30}$$

Note that the penalty-weighting coefficient vector ζ_k is of dimension $2r$ and that its last r components are associated with z_k. Then

$$\begin{bmatrix} E_k & 0_{r\times r} \\ 0_{r\times r} & H_k \end{bmatrix} \quad (k = 1, \ldots, N) \tag{31}$$

We made use of the relationship $\eta_k = \begin{Bmatrix} y_k \\ z_k \end{Bmatrix}$ for $k = 1, \ldots, N$ in the arguments of the Heaviside step functions $s_{i,k}$ for Equation (30).

If η_k of Equation (27), D_k of Equation (31), and $I[x_0, N]$ of Equation (2) are substituted into $\bar{I}[x_0, N]$ of Equation (16), then

$$\bar{I}[x_0, N] = \tau \sum_{k=1}^{N} [x^T{}_k Q_k x_k + u^T{}_{k-1} R_{k-1} u_{k-1}] + \tau \sum_{k=1}^{N} (y_k{}^T, z_k{}^T) \begin{bmatrix} E_k & 0_{r\times r} \\ 0_{r\times r} & H_k \end{bmatrix} \begin{bmatrix} y_k \\ z_k \end{bmatrix} \tag{32}$$

Imbedding the Constrained System

The concept of imbedding a smaller system into a larger one is not new; it was used by Koepcke (8) to handle a system with a time delay in the control vector. For our inequality constrained control problem, the larger or augmented system equation is

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \\ z_{k+1} \\ u_k \end{bmatrix} = \begin{bmatrix} \varphi_k & 0_{n\times r} & 0_{n\times r} & 0_{n\times r} \\ 0_{r\times n} & I_r & 0_{r\times r} & -I_r \\ 0_{r\times n} & 0_{r\times r} & I_r & I_r \\ 0_{r\times n} & 0_{r\times r} & 0_{r\times r} & 0_{r\times r} \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ z_k \\ u_{k-1} \end{bmatrix} + \begin{bmatrix} \Delta_k \\ I_r \\ -I_r \\ I_r \end{bmatrix} u_k \quad (k = 0, \ldots, N-1)$$

or

$$w_{k+1} = \Psi_k w_k + \Omega_k u_k \tag{33}$$

Vector w_k and matrices Ψ_k and Ω_k are defined by the first part of Equation (33). The dimensions of the augmented quantities, and the unaugmented quantities to which they correspond, are listed in Table 1. (Also listed are augmented quantities to be introduced later.) Equation (33) incorporates the system Equation (1) and the Equations (29) for the extra state variables. In addition, u_k is entered into w_{k+1} so that it may be used in the calculations on the next step. Note that constant φ_k and Δ_k matrices imply constant Ψ_k and Ω_k arrays, respectively.

An initial condition w_0 for Equation (33) is supplied from Equation (28)

$$w_0 = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \\ u_{-1} \end{bmatrix} = \begin{bmatrix} x_0 \\ -\alpha^+ \\ \alpha^- \\ 0 \end{bmatrix} \tag{34}$$

In order to write the augmented performance index in terms of the augmented state w_k, a new matrix Γ_k is introduced of size $(n + 3r) \times (n + 3r)$.

$$\Gamma_k = \begin{bmatrix} Q_k & 0_{n\times 2r} & 0_{n\times r} \\ 0_{2r\times n} & \begin{bmatrix} E_k & 0_{r\times r} \\ 0_{r\times r} & H_k \end{bmatrix} & 0_{2r\times r} \\ 0_{r\times n} & 0_{r\times 2r} & 0_{r\times r} \end{bmatrix} (k = 1, \ldots, N) \tag{35}$$

Then Equation (32) for the augmented index becomes

$$\bar{I}[w_0, N] = \tau \sum_{k=1}^{N} [w^T{}_k \Gamma_k w_k + u^T{}_{k-1} R_{k-1} u_{k-1}] \tag{36}$$

Equations (33) and (36) define an unconstrained L-Q problem. The Γ_k matrices, which correspond to the arrays Q_k, are positive semidefinite because the elements of the diagonal arrays E_k and H_k are either positive or zero by Equation (30). Thus, the arrays of the augmented performance index, Equation (36), satisfy the same minimum requirements as do the matrices of the unaugmented index of equality Equation (2). It should be pointed out that the Γ_k matrix cannot be positive definite, because this array is singular although symmetric.

Feedback Control

No attempt will be made here to show that the augmented problem satisfies the conditions on uniqueness of optimal control or on controllability designed for the corresponding unconstrained unaugmented system. Calculation of the discrete-time feedback control law from the augmented system proceeds in the same manner as with an unconstrained system of a L-Q problem. It is necessary only to express optimal control equations in terms of the augmented system quantities listed in Table 1.

The feedback control relation for the unconstrained original system is

$$u_k = -K_k x_k \quad (k = 0, \ldots, N-1) \tag{37}$$

TABLE 1. CORRESPONDENCE BETWEEN AUGMENTED AND UNAUGMENTED QUANTITIES

Quantity	Dimensions*	Quantity	Dimensions*
$\mathbf{w}$	$(n+3r)\times 1$	$\mathbf{x}$	$n\times 1$
$\boldsymbol{\Psi}$	$(n+3r)\times(n+3r)$	$\boldsymbol{\varphi}$	$n\times n$
$\boldsymbol{\Omega}$	$(n+3r)\times r$	$\boldsymbol{\Delta}$	$n\times r$
$\boldsymbol{\Gamma}$	$(n+3r)\times(n+3r)$	$\mathbf{Q}$	$n\times n$
$\boldsymbol{\pi}$	$(n+3r)\times(n+3r)$	$\mathbf{P}$	$n\times n$
$\boldsymbol{\theta}$	$r\times(n+3r)$	$\mathbf{K}$	$r\times n$
$\bar{I}[\mathbf{w}_o, N]$	1×1	$I[\mathbf{x}_o, N]$	1×1

* Dimensions are for systems with inequality control constraints.

For the augmented system, the control becomes

$$\mathbf{u}_k = -\boldsymbol{\theta}_k\mathbf{w}_k \quad (k=0,\ldots,N-1) \tag{38}$$

Boundary condition

$$\mathbf{P}_N = \mathbf{0}$$

becomes

$$\boldsymbol{\pi}_N = \mathbf{0} \tag{39}$$

and the feedback arrays of the augmented system will be

$$\boldsymbol{\theta}_{k-1} = [\boldsymbol{\Omega}^T_{k-1}(\boldsymbol{\Gamma}_k + \boldsymbol{\pi}_k)\boldsymbol{\Omega}_{k-1} + \mathbf{R}_{k-1}]^{-1}\boldsymbol{\Omega}^T_{k-1}(\boldsymbol{\Gamma}_k + \boldsymbol{\pi}_k)\boldsymbol{\Psi}_{k-1} \tag{40}$$

$$(k=1,\ldots,N)$$

$$\boldsymbol{\pi}_{k-1} = (\boldsymbol{\Psi}_{k-1} - \boldsymbol{\Omega}_{k-1}\boldsymbol{\theta}_{k-1})^T(\boldsymbol{\pi}_k + \boldsymbol{\Gamma}_k)\boldsymbol{\Psi}_{k-1} \tag{41}$$

corresponding to $\mathbf{K}_{k-1}$ and $\mathbf{P}_{k-1}$, respectively.

Although $\mathbf{u}_k$ is calculated from $\mathbf{w}_k$ by Equation (38), $\mathbf{w}_{k+1}$ is not produced from $\mathbf{w}_k$ by Equation (33). $\mathbf{x}_{k+1}$ is obtained from $\mathbf{x}_k$ by integration of the unaugmented or original system equation with the value of $\mathbf{u}_k$ provided. Also, $\mathbf{y}_{k+1}$ and $\mathbf{z}_{k+1}$ are determined from the following modified form of Equation (27)

$$\left.\begin{array}{l}\mathbf{y}_{k+1} = \mathbf{u}_k - \boldsymbol{\alpha}^+ \\ \mathbf{z}_{k+1} = -(\mathbf{u}_k - \boldsymbol{\alpha}^-)\end{array}\right\} \quad (k=0,\ldots,N-1)$$

By analogy to the $\mathbf{P}_k$, the corresponding matrices $\boldsymbol{\pi}_k$ are positive semidefinite and symmetric.

Let us now examine the formation of the matrices $\boldsymbol{\Gamma}_k$. To begin, the unconstrained problem is worked out to see where the control violates the constraints. This serves to start the iterative integration procedure. On each pass, vectors $\mathbf{y}_k$ and $\mathbf{z}_k$ are determined by Equation (27). For the next pass, the elements of the diagonal arrays $\mathbf{E}_k$ and $\mathbf{H}_k$ in $\boldsymbol{\Gamma}_k$ are evaluated by means of Equation (30). What is required in this calculation of the elements are the Heaviside step function vectors $\mathbf{s}_k(\mathbf{y}_k)$ and $\mathbf{s}_k(\mathbf{z}_k)$ and the penalty-weighting coefficients $\zeta_{i,k}$.

Determination of the step functions is straightforward. However, there is a special way the weighting coefficients $\zeta_{i,k}$ are produced in the numerical examples of linear systems with constrained controls to appear later. In these examples, for one pass through the feedback matrix and system equations, all the coefficients $\zeta_{i,k}$ are set equal to a constant penalty-weighting coefficient ζ. Initially, ζ is made some small value. After that, at the beginning of each pass or iteration, ζ is increased or left unchanged. Usually ζ is always increased with each new pass, rather than sometimes being left the same for two or more consecutive passes. Increasing ζ stops when the control satisfies the constraints to within a given tolerance. In this work, meeting the inequality control constraints given by Equation (24) to within a tolerance μ means that

$$-\mu + \alpha_i^- \leq u_{i,k} \leq \alpha_i^+ + \mu \qquad \begin{array}{l}(i=1,\ldots,r)\\(k=0,\ldots,N-1)\end{array} \tag{42}$$

When this condition is met, it is said that convergence to the control constraints is attained.

PENALTY FUNCTION METHODS

Intuitively, the penalty function approach seems reasonable for the L-Q system. Reference to the augmented performance index of Equation (32) reveals the appearance of terms such as $\mathbf{x}_k^T\mathbf{Q}_k\mathbf{x}_k$ and $\mathbf{y}_k^T\mathbf{E}_k\mathbf{y}_k$. Suppose $\mathbf{Q}_k$ is diagonal with elements $Q_{ii,k} = \lambda_i$. $\mathbf{E}_k$ is diagonal, and by Equation (30), $E_{ii,k} = \zeta_{i,k}\, s_{i,k}(y_{i,k})$ where $\mathbf{y}_k = \mathbf{u}_{k-1} - \boldsymbol{\alpha}^+$. It has been shown that, as λ_i is made larger, the corresponding value of $|x_{i,k}|$ is diminished. So even apart from penalty function theory, for a given index i, if $y_{i,k} = u_{i,k-1} - \alpha_i^+ > 0$, so that $s_{i,k}(y_{i,k}) = 1$, it follows that as $\zeta_{i,k}$ is made large, then $|y_{i,k}|$ becomes small. But this means that $u_{i,k-1}$ is made to be close in value to α_i^+, or what is the same thing, $u_{i,k}$ comes closer to satisfying the constraint $u_{i,k-1} \leq \alpha_i^+$.

For inequality constraints, either the afore-mentioned technique or an alternate method of adjusting the penalty-weighting coefficients $\zeta_{i,k}$ can be used. The alternate method involves adjusting $\zeta_{i,k}$ separately for each component of the constraint vector at each time step. To demonstrate this technique, we will examine the behavior of $y_{i,k} = u_{i,k-1} - \alpha_i^+$ and its influence on $\zeta_{i,k}$ for one value of i and of k. The object is to make $u_{i,k-1} \leq \alpha_i^+$ or, what is the same thing, to make $y_{i,k} \leq 0$.

In Equation (30) we have $E_{ii,k} = \zeta_{i,k}$ and, to start, $\zeta_{i,k} = 0$. If on the first pass $u_{i,k-1}$ exceeds α_i^+, $\zeta_{i,k}$ is increased for the next pass. On every pass $\zeta_{i,k}$ will be gradually increased to an upper limit if, for all passes, $y_{i,k}$ does not become negative. When $y_{i,k}$ does become negative, $\zeta_{i,k}$ is reduced with zero as a lower bound.

In other words, $\zeta_{i,k}$ is increased or decreased gradually in steps between zero and a maximum depending on the number of passes for which $y_{i,k}$ is positive or negative. A switch of $y_{i,k}$ from negative to positive will start $\zeta_{i,k}$ increasing and conversely. If $y_{i,k}$ does not change sign but merely becomes zero, this situation is treated as if $y_{i,k}$ were still positive or negative, as the case may be. When $y_{i,k}$ is always zero from the start of the iterative procedure, $\zeta_{i,k}$ is always zero. A possible criterion for terminating such a penalty function procedure is to stop when all $\zeta_{i,k}$ are either zero or maximal and the constraints are met to within a given tolerance.

FURTHER POINTS

Other interesting developments in the current method of solving the L-Q problem with constraints (*16*) can be mentioned only briefly. The point is that for a unique and optimal control in the unconstrained problem to be guaranteed, it is necessary that the array $\boldsymbol{\Delta}^T_{k-1}(\mathbf{Q}_k + \mathbf{P}_k)\boldsymbol{\Delta}_{k-1} + \mathbf{R}_{k-1}$ be positive definite at all time steps k. In the constrained case, the equivalent statement holds for the array $\boldsymbol{\Omega}^T_{k-1}(\boldsymbol{\Gamma}_k + \boldsymbol{\pi}_k)\boldsymbol{\Omega}_{k-1} + \mathbf{R}_{k-1}$.

Secondly, our development has been restricted to the inequality control constraints of Equation (24) rather than the general constraints of Equation (3) or (6). The inclusion of different constraints applied either singly or together at all time steps or even at intermittent time steps can be easily developed. Finally, it is possible to include terminal equality state constraints in the formulations. When this general type of constraints is included, it may be necessary to linearize (apparent) the extra state variables $\boldsymbol{\eta}_k$. Even in the linear system case this will mean that an iterative procedure may be required for the updating of the linearized coefficients.

Suboptimal Control

To solve the nonlinear problem in a suboptimal fashion, we will combine the material of Part I (*17*) and the theory of the present paper.

Starting with a nonlinear state equation

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) \quad (43)$$

and a given initial condition $\mathbf{x}_0$, we transform this via apparent linearization to

$$\dot{\mathbf{x}}(t) = \mathbf{A}(\mathbf{x}(t), \mathbf{u}(t), t)\mathbf{x}(t) + \mathbf{B}(\mathbf{x}(t), \mathbf{u}(t), t)\mathbf{u}(t) \quad (44)$$

Then by using some control policy $\hat{\mathbf{u}}(t)$ and a resulting state $\hat{\mathbf{x}}(t)$, the arrays $\mathbf{A}$ and $\mathbf{B}$ are evaluated as functions of time according to

$$\mathbf{A}(t) = \mathbf{A}(\hat{\mathbf{x}}, \hat{\mathbf{u}}, t); \quad \mathbf{B}(t) = \mathbf{B}(\hat{\mathbf{x}}, \hat{\mathbf{u}}, t) \quad (45)$$

Now Equation (44) may be discretized to form

$$\mathbf{x}_{k+1} = \boldsymbol{\varphi}_k \mathbf{x}_k + \boldsymbol{\Delta}_k \mathbf{u}_k \quad (46)$$

where $\boldsymbol{\varphi}_k$ and $\boldsymbol{\Delta}_k$ can be considered functions of the trajectories $\hat{\mathbf{u}}(t)$ and $\hat{\mathbf{x}}(t)$. Equation (46), combined with the performance index of (2) and the constraints of (24), makes up the actual constrained L-Q problem.

To begin, a noniterative suboptimal control method (*17*) is used to evaluate $\hat{\mathbf{u}}(t)$ and $\hat{\mathbf{x}}(t)$ for the unconstrained nonlinear system; the $\boldsymbol{\varphi}_k$ and $\boldsymbol{\Delta}_k$ matrices are then calculated and frozen at their present values. Next, the augmented system is constructed, and for $\zeta_{i,k} = \zeta = 0$ (a penalty of zero) the corresponding L-Q problem is solved. This may be considered as the first pass of the penalty function procedure. We should recall that $\boldsymbol{\varphi}_k$ and $\boldsymbol{\Delta}_k$ are the only portions of the augmented arrays $\boldsymbol{\Psi}_k$ and $\boldsymbol{\Omega}_k$, respectively, which are variable, that is

$$\boldsymbol{\Psi}_k = \boldsymbol{\Psi}_k(\boldsymbol{\varphi}_k); \quad \boldsymbol{\Omega}_k = \boldsymbol{\Omega}_k(\boldsymbol{\Delta}_k) \quad (47)$$

After the first pass or iteration of the penalty function method, the remaining integrations are carried out with positive values of the coefficient ζ. For illustration, in the nonlinear numerical example to be discussed shortly, the sequence was $\zeta = 0$, $\zeta = 10$, $\zeta = 10^4$, and $\zeta = 10^8$ for four passes through the penalty function procedure. Convergence to the control constraints was achieved with μ in Equation (42) set to a small number, 10^{-7}.

The augmented system is employed only to produce the discrete-time control vectors $\mathbf{u}_k$. The control policy is the same for both the augmented state Equation (33) and the original state equation of the constrained system. Given a control vector $\mathbf{u}_k$, and the last state $\mathbf{x}_k$, an integration over one time step provides the next state $\mathbf{x}_{k+1}$. Then $\mathbf{x}_{k+1}$ is used to form the new augmented state $\mathbf{w}_{k+1}$. Further, the use of the nonlinear equation in the state vector integration applies every time a state trajectory is produced. Such a procedure is not limited to the time when the control constraints are being met. It is also employed during the re-evaluation of the matrices $\boldsymbol{\varphi}_k$ and $\boldsymbol{\Delta}_k$ about to be discussed.

Now let $\hat{\mathbf{u}}(t)$ be that control policy which satisfies Equation (42) and let $\hat{\mathbf{x}}(t)$ be the corresponding state trajectory. $\hat{\mathbf{u}}(t)$ and $\hat{\mathbf{x}}(t)$ then are the last trajectories produced by the penalty function method. As soon as the control constraints are met, the arrays $\boldsymbol{\varphi}_k$ and $\boldsymbol{\Delta}_k$ and, therefore, the matrices $\boldsymbol{\Psi}_k$ and $\boldsymbol{\Omega}_k$ are updated or recalculated from $\hat{\mathbf{u}}(t)$ and $\hat{\mathbf{x}}(t)$. At the same time, the arrays $\boldsymbol{\Gamma}_k$ are frozen at the values they had during the last integration or iteration of the penalty function procedure. These matrices $\boldsymbol{\Gamma}_k$ are the ones which contain the penalty-weighting coefficients. Holding them constant, then, is equivalent to making the assumption, even if it is incorrect, that while the matrices $\boldsymbol{\Psi}_k$ and $\boldsymbol{\Omega}_k$ (or $\boldsymbol{\varphi}_k$ and $\boldsymbol{\Delta}_k$) are being adjusted, the controls retain their same tendency to penetrate the constraints.

With the augmented cost matrices frozen, and arrays $\boldsymbol{\Psi}_k$ and $\boldsymbol{\Omega}_k$ newly evaluated, the control policy is recalculated from the augmented system and a new set of states is produced. Matrices $\boldsymbol{\Psi}_k$ and $\boldsymbol{\Omega}_k$ are again updated. Whether some of the constraints start being penetrated is of no concern. What matters is that the $\boldsymbol{\Psi}_k$ and $\boldsymbol{\Omega}_k$ arrays, and therefore the $\boldsymbol{\varphi}_k$ and $\boldsymbol{\Delta}_k$ arrays, continue to be updated until, for $\hat{\mathbf{x}}_k$ and $\hat{\mathbf{u}}_k$ denoting vectors of the previous iteration,

$$|x_{i,k} - \hat{x}_{i,k}| \leq 10^{-7} \quad (i = 1, .., n \text{ and } k = 1, .., N) \quad (48)$$

$$|u_{i,k} - \hat{u}_{i,k}| \leq 10^{-7} \quad (i = 1, .., r \text{ and } k = 0, .., N-1)$$

This procedure is none other than the iterative method (ITER) of suboptimal control discussed in Part I (*17*). However, it is the augmented system which is being used to yield the control vectors.

Convergence given by Equation (48) is described as convergence to the trajectories, those for the state and controls. Such convergence can also be viewed in terms of the coefficient matrices $\boldsymbol{\Psi}_k$ and $\boldsymbol{\Omega}_k$ or $\boldsymbol{\varphi}_k$ and $\boldsymbol{\Delta}_k$, which are evaluated as functions of the state and control trajectories. In turn, these matrices are employed in the calculation of a new set of trajectories. When the previous set of trajectories is sufficiently close to the new set, or when the old coefficient matrices differ little from the newly generated ones, convergence has occurred.

In the case of the nonlinear absorber problem, once this convergence was attained, it was found that the constraints were still obeyed to within the tolerance of 10^{-7}. There can be problems for which satisfying Equation (48) means that relation (42) for the control constraints is no longer obeyed. Whenever this is the case, the following is done: Let $\mathbf{x}^*(t)$ and $\mathbf{u}^*(t)$ be the final trajectories which satisfy Equation (48). Then the coefficient matrices $\boldsymbol{\Psi}_k$ and $\boldsymbol{\Omega}_k$ (or $\boldsymbol{\varphi}_k$ and $\boldsymbol{\Delta}_k$) are reevaluated with $\mathbf{x}^*(t)$ and $\mathbf{u}^*(t)$. Furthermore these arrays are frozen at these values during the process of again satisfying the control constraints. Where a new set of penalties has to be applied, the constant penalty-weighting coefficient ζ is brought back down to a low value such as 10. Application of the rest of the procedure for meeting the constraints proceeds as before.

In addition, if the conditions of Equation (48) are not satisfied after the control constraints are again met, the suboptimal operation is called upon. One can continue alternating the penalty function technique with that for suboptimal control until both Equation (42) and relation (48) are obeyed. Then the nonlinear constrained control problem will be completely solved for the suboptimal control. It should be emphasized that the entire manner of handling the nonlinear problem is suboptimal because the method of treating the nonlinearity is not optimal. Also note that it is assumed that convergence to the trajectories indicated by Equation (48) is possible. This convergence was attained for the nonlinear absorber example.

TABLE 2. SLIGHTLY AND MOST CONSTRAINED OPTIMAL CONTROL OF LINEAR ABSORBER*

(Indices I and $\bar{I}$ as a function of iteration number)

Iteration number	Penalty-weighting coefficient ζ	Slightly constrained control I	Slightly constrained control $\bar{I}$	Most constrained control I	Most constrained control $\bar{I}$
1	0	0.04654	0.04654	0.04654	0.04654
2	10	0.04697	0.04698	0.05512	0.05527
3	10^4	0.04699	0.04699	0.05611	0.05611
4	10^8	0.04699	0.04699	0.05612	0.05612

* Computing time of about 3 min.

Numerical Test Systems and Results

In order to test out the theoretical concepts, three different systems were analyzed computationally. The first system is a linear version of the nonlinear absorber presented in Part I (*17*). All the equations are in the literature (9) and we merely mention that Equation (1) is obtained directly with $\boldsymbol{\varphi}_k = \boldsymbol{\varphi}$ and $\boldsymbol{\Delta}_k = \boldsymbol{\Delta}$ as constants.

The second test problem is also a linear one and represents the behavior of an artificial earth satellite as detailed by Deley (3). In final form the equations are given by

$$\mathbf{x}_{k+1} = \boldsymbol{\varphi}\mathbf{x}_k + \boldsymbol{\Delta}\mathbf{u}_k \tag{49}$$

with

$$\boldsymbol{\varphi} = \begin{bmatrix} 0.54030 & 0.84147 \\ -0.84147 & 0.54030 \end{bmatrix}; \quad \boldsymbol{\Delta} = \begin{bmatrix} 0.45970 \\ 0.84147 \end{bmatrix} \tag{50}$$

The initial state and the constraints are

$$\mathbf{x}_0 = \mathbf{x}(0) = \begin{bmatrix} -5 \\ 3 \end{bmatrix} \tag{51}$$

and

$$-1 \leq \mathbf{u}_k = u_k \leq 1 \tag{52}$$

In this problem the control is scalar and physically represents the operation of stabilization gas jets.

The third test system is the nonlinear absorber system (*17*) covered in Part I.

THE CONSTRAINED LINEAR ABSORBER

For this completely linear system, we use $t_f = 25$, $\tau = 1$ and thus $N = 25$. In the quadratic performance index of Equation (2), we select $\mathbf{Q} = \mathbf{I}$ and $\mathbf{R} = \mathbf{0}$. Two cases of inequality constrained control are considered for the computations. These are

$$\boldsymbol{\alpha}^- = \begin{bmatrix} 0.0 \\ -0.025 \end{bmatrix} \leq \mathbf{u}_k \leq \begin{bmatrix} 0.2 \\ 0.3 \end{bmatrix} = \boldsymbol{\alpha}^+ \tag{53}$$

and

$$\boldsymbol{\alpha}^- = \begin{bmatrix} 0.0 \\ -0.025 \end{bmatrix} \leq \mathbf{u}_k \leq \begin{bmatrix} 0.1 \\ 0.15 \end{bmatrix} = \boldsymbol{\alpha}^+ \tag{54}$$

Equations (53) and (54) are referred to as the slightly constrained and the most constrained cases, respectively.

For each of the two constrained cases, four repetitive integrations were employed. The first was carried out with the constant scalar penalty-weighting coefficient ζ set equal to zero (the unconstrained L-Q problem is solved). For the second, ζ is increased to 10; for the third, ζ is 10^4, and finally ζ is raised to 10^8 on the fourth integration. On the fourth pass the constraints are considered satisfied, for then

$$-10^{-7} + \alpha_i^- \leq u_{i,k} \leq \alpha_i^+ + 10^{-7} \quad \begin{matrix}(i = 1, 2) \\ (k = 0, \ldots, N-1)\end{matrix} \tag{55}$$

where $\boldsymbol{\alpha}^+$ and $\boldsymbol{\alpha}^-$ for each example comes from Equations (53) and (54). That is, the constraints are met or convergence to the control constraints is attained to within a tolerance of 10^{-7}.

Table 2 shows the behavior of the unaugmented performance index I and the augmented performance index $\bar{I}$ as a function of iteration number for the two constrained cases and also for the unconstrained case listed as iteration 1. Note that, starting with the second iteration at which the difference $\bar{I} - I$ is first nonzero, this difference approaches zero as ζ is increased. Also, as ζ increases, the performance indices I and $\bar{I}$ approach an upper limit. This behavior of the indices was anticipated. Because the penalties for constraint violations are positive, the augmented cost $\bar{I}$ is larger than the unaugmented cost I when the two costs differ. The difference between the augmented and unaugmented index at the last iteration is actually zero to eight decimal places for the most constrained case. As expected, the (unaugmented) performance index increases as tighter constraints on the control are applied. Furthermore, the slightly constrained control policy yields a performance index only slightly larger than that for the unconstrained system.

Figures 1 and 2 show selected results for the unconstrained and two constrained control computational runs. To avoid cluttering Figure 1, we do not plot the slightly constrained control components. The control vector component u_1 (and u_2) is non-negative. Also, the unconstrained control is initially large and then diminishes to near zero rapidly with time. Constraining the control for the linear absorber system really means applying upper bounds on the control. Application of constraints to the control results in more control being employed after the given control component is no longer as large as the upper limit. That is, once a control component leaves its upper bound, it is never less than the corresponding unconstrained control component at any given time step. In addition, all control components are monotone nonincreasing.

In Figure 2, on the other hand, the trajectories of x_6 are not monotone. These trajectories possess an internal maximum which decreases as the control becomes more constrained. There is less control available initially to cause the x_6 trajectories to overshoot the time axis when the control is more constrained. However, the x_6 trajectories of the most constrained optimal control example end up above the corresponding curves for the slightly constrained

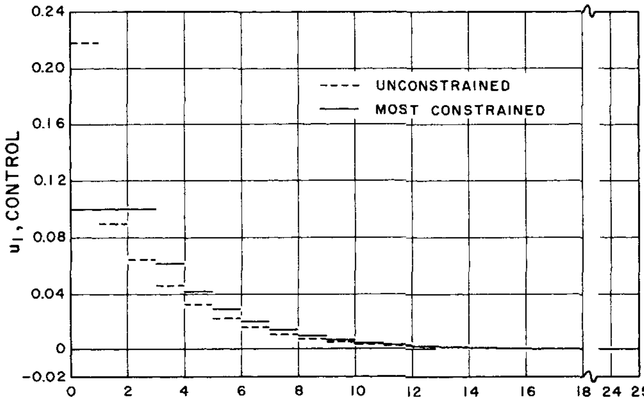


Fig. 1. Optimal control u_1 vs. time for the linear absorber.

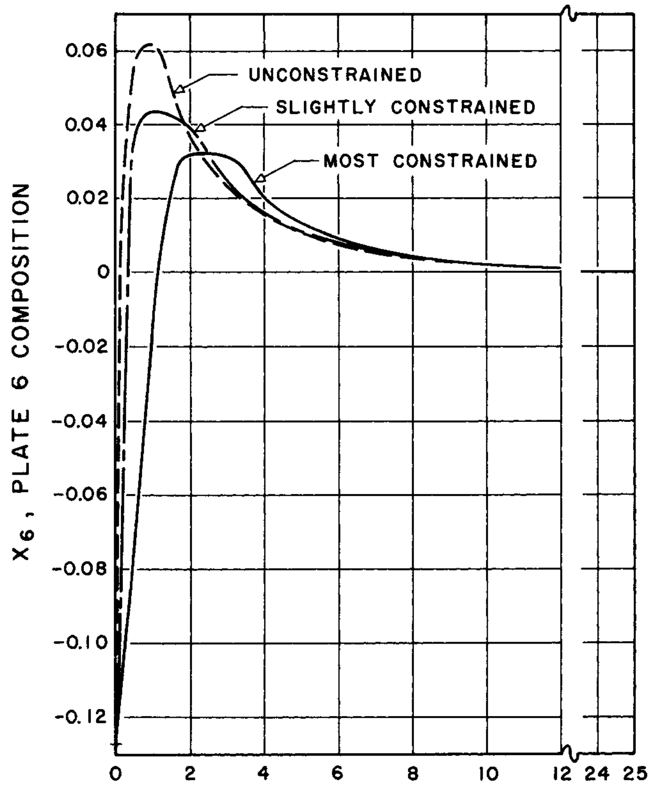


Fig. 2. x_6 vs. time for the linear absorber.

and unconstrained control cases.

This example shows that optimal control of a L-Q problem with constraints can be obtained via the penalty function method. The constraints can be met as exactly as desired, the dimensions of the state and control are not important, and feedback control is achieved. Moreover, the procedure appears to be computationally fast.

CONSTRAINED ARTIFICIAL EARTH SATELLITE

In the truly constrained problem, as specified by Equations (49) to (52), the final time t_f (and thus N) is not fixed; in fact, Deley specified an infinite final time. However, for the present computations a value of $t_f = 15$ was used; this is large enough to be considered infinite. Further, the matrices in the quadratic performance index were chosen to be

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}; \quad R = 0$$

to conform to Deley's approach. As neither of these matrices is positive definite, one cannot actually guarantee that, given enough time, the state or control will approach zero.

With the penalty function approach the procedure was as follows. For the first iteration, the unconstrained problem ($\zeta = 0$) is solved with $N = 15$, corresponding to $\tau = 1.0$. After the first iteration, eleven more passes are made with $\zeta = 10^4$. At this value of ζ, it is not possible to further minimize the performance index and still have the constraints met to within a reasonable tolerance. For $\zeta = 10^4$, this reasonable tolerance happens to be 10^{-3}. The next or thirteenth pass is made with ζ set at 10^7; this iteration is the last one neeeded because, at its completion, the constraints are met to within a tolerance of 10^{-6}.

Listed in Table 3 as a function of iteration number are the unaugmented index I and the augmented index $\bar{I}$. Also listed is the penalty-weighting coefficient ζ which varies

TABLE 3. CONSTRAINED OPTIMAL CONTROL OF ARTIFICIAL EARTH SATELLITE

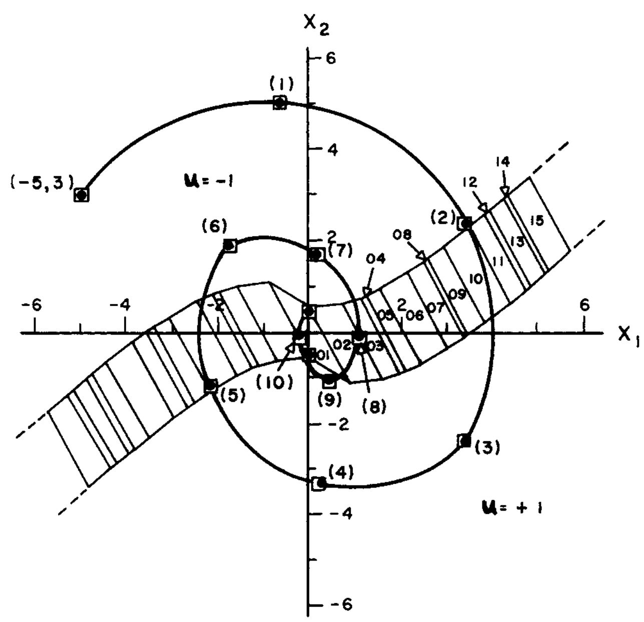
(Costs I and $\bar{I}$ as a function of iteration number)

Iteration number	Penalty-weighting coefficient ζ	I	$\bar{I}$
1	0	0	0
2	10^4	51.834	51.916
3	10^4	111.708	111.881
4	10^4	37.731	37.742
5	10^4	32.795	32.811
6	10^4	33.456	33.478
7	10^4	32.808	32.825
8	10^4	32.610	32.627
9	10^4	32.666	32.684
10	10^4	32.666	32.683
11	10^4	32.666	32.683
12	10^4	32.665	32.683
13	10^7	32.703	32.704

with iteration number. At the thirteenth pass appears the final result, an I of 32.703. In excellent agreement is Deley's value of $I = 32.702$ (3).

The general behavior of I and $\bar{I}$ as the iterations proceed (Table 3) is basically that expected from the theory. There are a few anomalies, such as I and $\bar{I}$ having lower values on iteration No. 8 than on iteration No. 12. However, this anomaly can be explained by the amount of violation of the constraints on iteration No. 8. The time to execute the entire program with its thirteen passes is little more than 20 sec.

Figure 3 shows a plot of x_2 versus x_1 for both Deley's method (3) and the penalty function method of finding constrained optimal control for the artificial earth satellite problem. This figure, with a few modifications, is Deley's. For time steps 1 through 10, there is a single number in parentheses next to each point indicating the corresponding time step. Round points indicate Deley's results and are connected by a smooth spiral curve. Results from the



—•— DELEY

□ PENALTY FUNCTION APPROACH

Fig. 3. x_2 vs. x_1 for the artificial earth satellite [adapted by permission from Stanford Electronics Laboratories and G. W. Deley (3)].

penalty function approach are shown by the squares superimposed upon Deley's graph. Each center of a square indicates the values of the coordinates x_1 and x_2. As each square contains most or all of the corresponding point of Deley, excellent agreement is obtained.

This problem is interesting because of the different regions of control which are possible. There are different areas blocked off by Deley in Figure 3; there is one region for which the control is $u = -1$ and another with $u = +1$. These are regions of saturation in which the control constraints hold exactly. A band separates the saturation regions from the corresponding unsaturated controls; that is, $-1 < u < 1$ is obeyed. When the state lies within the band, the optimal control is given by

$$u_k = -\mathbf{K}_k \mathbf{x}_k + d_k \tag{56}$$

where $\mathbf{K}_k$ is a feedback vector. Further, within the subregion in the band marked "01," $d_k = 0$; on either side of "01" are various subsections of the band for which Equation (56) still holds but $d_k \neq 0$. Throughout each subdivision of the band shown $\mathbf{K}_k$ and d_k are the same. At time step 5, the state vector $\mathbf{x} = \mathbf{x}_5$ falls in region "06" to the left and $\mathbf{x}_8$ falls into region "02" to the right. For this problem, these are the only two states which fall into regions of the band where $d_k \neq 0$.

Deley suggests that Equation (56) holds only when the control is unsaturated. The present penalty function method, however, generates this law irrespective of the specific region. Thus, if the augmented feedback matrix $\boldsymbol{\theta}_k$ is partitioned as follows

$$\underset{r \times (n+3r)}{\boldsymbol{\theta}_k} = [\underset{r \times n}{\hat{\boldsymbol{\theta}}_k}, \underset{r \times 3r}{\boldsymbol{\theta}_k^*}]$$

then the feedback control law, Equation (38), becomes

$$\mathbf{u}_k = -\boldsymbol{\theta}_k \mathbf{w}_k = -[\hat{\boldsymbol{\theta}}_k, \boldsymbol{\theta}_k^*] \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \\ \mathbf{z}_k \\ \mathbf{u}_{k-1} \end{bmatrix}$$

or

$$\mathbf{u}_k = -\hat{\boldsymbol{\theta}}_k \mathbf{x}_k - \boldsymbol{\theta}_k^* \begin{bmatrix} \mathbf{y}_k \\ \mathbf{z}_k \\ \mathbf{u}_{k-1} \end{bmatrix}$$

This equation is of the exact form of (56) for the scalar control of the present problem.

CONSTRAINED NONLINEAR ABSORBER

We shall analyze the computational features for the case $\mathbf{Q}_k = \mathbf{Q} = \mathbf{I}$ and $\mathbf{R}_k = \mathbf{R} = \mathbf{I}$ in the performance index and with $t_f = 25$, $\tau = 1$ and $N = 25$.

To serve as a comparison for the constrained control cases, unconstrained control is found for the nonlinear absorber by means of the ITER method of suboptimal control described in Part I. The control constraints will have the forms

$$\boldsymbol{\alpha}^- = \begin{bmatrix} 0 \\ -0.025 \end{bmatrix} \leq \mathbf{u}_k \leq \begin{bmatrix} 0.045 \\ 0.080 \end{bmatrix} = \boldsymbol{\alpha}^+ \tag{57}$$

and

$$\boldsymbol{\alpha}^- = \begin{bmatrix} 0 \\ -0.005 \end{bmatrix} \leq \mathbf{u}_k \leq \begin{bmatrix} 0.025 \\ 0.025 \end{bmatrix} = \boldsymbol{\alpha}^+$$

called the slightly constrained and the most constrained cases, respectively. Application of the constraints involves an iterative procedure. Also, as the system is nonlinear, iterative integration is necessary to update the coefficient matrices $\boldsymbol{\varphi}_k$ and $\boldsymbol{\Delta}_k$ of the system equation. Therefore, two convergence criteria are set up. The first deals with convergence to the control constraints. The second criterion, which is associated with the coefficient updating, deals with convergence of the state and control trajectories. In both cases [Equations (42) and (48)] the tolerance is 10^{-7}.

In order to initiate calculations for the ITER method of suboptimal control in the unconstrained case, the solution from noniterative suboptimal control of the nonlinear absorber (Part I) is provided. With the ITER method, four iterations were taken; at their completion, Equation (48) was obeyed by the state and control trajectories.

For the slightly constrained control problem, a noniterative method of suboptimal control is used on the nonlinear absorber system to begin the calculations. To this end, the control and state trajectories from the unconstrained suboptimal control of the nonlinear absorber (Part I) are used to evaluate the coefficient matrices $\boldsymbol{\psi}_k$ and $\boldsymbol{\Delta}_k$.

With this set of $\boldsymbol{\varphi}_k$ and $\boldsymbol{\Delta}_k$ arrays, iterative integration for the slightly constrained control case was initiated with convergence to the control constraints being attained in four iterations. Then convergence to the trajectories was attained in three more passes. At this point, the slightly constrained control problem was completely solved with only a total of seven iterations. Just as in the case of the constrained linear absorber, during convergence to the constraints, the penalty-weighting coefficient ζ was changed from one iteration to another. During the first iteration, ζ was zero corresponding to no penalties. For the second iteration, ζ was changed to 10. On the third and fourth passes, ζ was made 10^4 and 10^8, respectively. For the fifth, sixth, and seventh iterations—namely, those associated with convergence to the trajectories—ζ was left at 10^8.

The most constrained case was carried out in exactly the same manner as the slightly constrained case with two exceptions. First, to start the iterations, $\boldsymbol{\varphi}_k$ and $\boldsymbol{\Delta}_k$ were obtained from the solution to the slightly constrained problem just described. Second, and because of the starting matrices, only two iterations were necessary to converge to the trajectories once convergence to the constraints had been attained in four iterations. Now only six total iterations were required; for these, the previously used ζ were employed.

Table 4 presents selected results from the computations

TABLE 4. SLIGHTLY AND MOST CONSTRAINED SUBOPTIMAL CONTROL OF NONLINEAR ABSORBER

(Unaugmented and augmented indices as a function of iteration number)

	Iteration number	Penalty-weighting coefficient ζ	I	$\bar{I}$
	1	0	0.10822969	0.10822969
	2	10	0.10845426	0.10848159
Slightly	3	10^4	0.10851377	0.10851380
constrained	4	10^8	0.10851384	0.10851384
control	5	10^8	0.10851387	0.10851387
	6	10^8	0.10851387	0.10851387
	7	10^8	0.10851386	0.10851386
	1	0	0.10822996	0.10822996
	2	10	0.11764085	0.11925797
Most	3	10^4	0.12121198	0.12121428
constrained	4	10^8	0.12121742	0.12121742
	5	10^8	0.12121741	0.12121741
	6	10^8	0.12121741	0.12121741

IBM 7094 computer time is not excessive, requiring an average of 4 min. for each suboptimal control case.

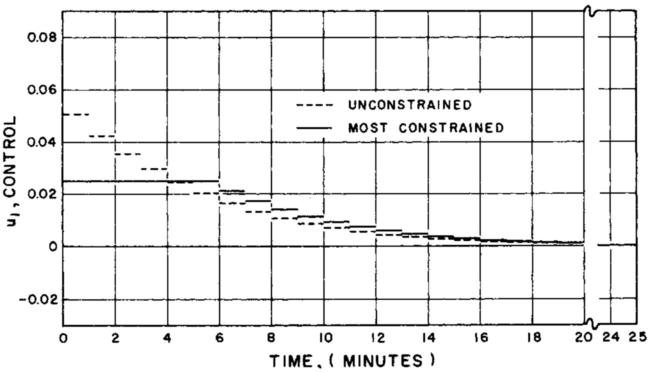


Fig. 4. Suboptimal control u_1 for the nonlinear absorber.

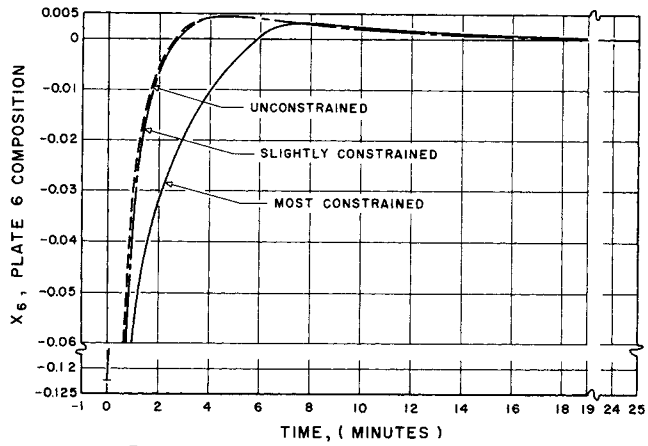


Fig. 6. x_6 vs. time for the nonlinear absorber.

in which the unaugmented and augmented indices are shown as a function of iteration number. We use eight decimal places to allow the changes to be more easily observed. The behavior is basically what would be expected from the theory; that is, after the first iteration $\bar{I}$ is greater than I, but when convergence occurs the two are at the same limiting value. Notice that the more the control is constrained, the greater the cost. For the ITER method of unconstrained suboptimal control, the index I is 0.10822984.

Figures 4 to 6 show some results of the computations. Once again these data are in agreement with what one would expect for this problem as the constraints change. Each constrained control vector component starts out at its upper bound and later gradually converges to the corresponding unconstrained control components. Also, initially, the constrained components of the state vector are prevented from approaching zero as rapidly as the corresponding unconstrained components.

CONCLUSIONS

We have presented a means of finding feedback control for multidimensional systems with quadratic indices and general constraints. A discrete-time penalty function method involving iterative integration was employed for satisfying the constraints. The constrained problem is reduced to an unconstrained L-Q problem by imbedding the constrained system into a larger unconstrianed one, called the augmented system. If the system is nonlinear, apparent linearization is used in the formation of the augmented system, and then iterative suboptimal control is combined with the penalty function technique.

Three numerical examples with inequality control constraints have been detailed. Optimal control was obtained for the first two examples, and suboptimal control was calculated for the third. For these examples, the constraints were met to within a small tolerance. Also calculations were performed with a minimum of computer time, so that real-time control is possible.

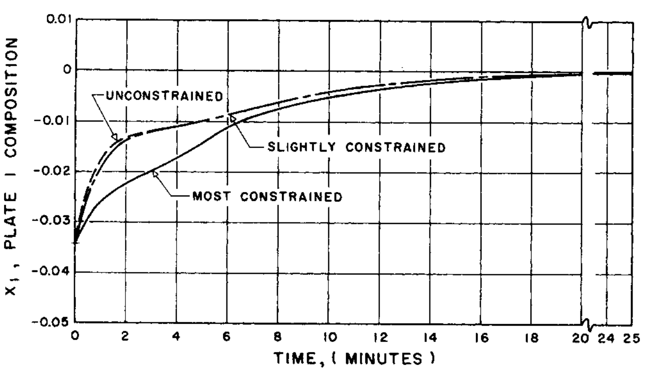


Fig. 5. x_1 vs. time for the nonlinear absorber.

ACKNOWLEDGMENTS

The authors wish to acknowledge the support of this work from National Science Foundation Grant NSF-GP 2858. Furthermore, this work made use of Princeton University computer facilities supported in part by National Science Foundation Grant NSF-GP-579.

NOTATION*

- $\mathbf{b}$ = vector of penalty functions, equality constraints
- $\mathbf{c}$ = vector of penalty functions, inequality constraints
- $\mathbf{d}$ = additive constant vector
- $\mathbf{D}$ = diagonal matrix
- $\mathbf{E}$ = diagonal matrix
- $\mathbf{g}$ = vector
- $\mathbf{h}$ = vector
- $\mathbf{H}$ = diagonal matrix
- $\bar{I}$ = augmented performance index
- $\mathbf{L}$ = diagonal matrix
- m = integer
- $\mathbf{s}$ = vector Heaviside-unit step function
- $\mathbf{w}$ = augmented state vector
- $\mathbf{y}$ = vector
- $\mathbf{z}$ = vector

Greek Letters

- $\boldsymbol{\alpha}$ = vector of inequality control constraints, $\boldsymbol{\alpha}^+$ upper and $\boldsymbol{\alpha}^-$ lower constraints
- $\boldsymbol{\beta}$ = vector
- $\boldsymbol{\Gamma}$ = augmented Q matrix
- $\boldsymbol{\xi}$ = vector of penalty weighting coefficients for inequality constraints
- $\boldsymbol{\eta}$ = vector
- $\boldsymbol{\theta}$ = augmented feedback matrix
- μ = tolerance
- $\boldsymbol{\pi}$ = augmented P matrix
- $\boldsymbol{\rho}$ = vector of penalty weighting coefficient for equality constraints
- $\boldsymbol{\Psi}$ = augmented $\boldsymbol{\varphi}$ matrix
- $\boldsymbol{\Omega}$ = augmented $\boldsymbol{\Delta}$ matrix

Subscripts

- i, j = vector or matrix elements

LITERATURE CITED

1. Bryson, A. E., Jr., and Y. Ho, "Applied Optimal Control: Optimization, Estimation, and Control," Blaisdell Publish-

* In addition to those in Part I.

ing Co., Waltham, Massachusetts (1969).
2. Chyung, D. H., *IEEE Trans. Autom. Control,* **12,** 37 (1967).
3. Deley, G. W., "Optimal Bounded Control of Linear Sampled-Data Systems Using Quadratic Performance Criteria," SUDAER No. 148, Department of Aeronautics and Astronautics, Stanford University, Stanford, California (1963).
4. Denn, M. M., and R. Aris, *Ind. Eng. Chem. Fundamentals,* **2,** 213 (1965).
5. Fujisawa, T., and Y. Yasuda, *J. Math. Anal. Applications,* **19,** 386 (1967).
6. Johnson, C. D., and W. M. Wonham, "On a Problem of Letov in Optimal Control," JACC Proc. (1964).
7. Kelley, H. J., "Method of Gradients" *in Optimization Techniques with Applications to Aerospace Systems,* G. Leitmann, Ed., Academic Press Inc., New York (1962).
8. Koepcke, R. W., *J. Basic Eng., Trans. Am. Soc. Mech. Engrs.,* Series D, 87, 74 (1965).
9. Lapidus, L., and R. Luus, "Optimal Control of Engineering Processes" Blaisdell Publishing Co., Waltham, Massachusetts (1967).
10. Lele, M. M., and D. H. Jacobson, *J. Math. Anal. Applications,* **26,** 163 (1969).
11. Okamura, K., *J. SIAM Control,* **2,** 317 (1965).
12. Rothenberger, B. F. "Quasilinearization as a Numerical Method," Ph.D. Dissert., Princeton University, Princeton, New Jersey (1966).
13. Russell, D. L., *J. SIAM Control,* Series A, **2,** 409 (1965).
14. Sage, A. P., "Optimum Systems Control," Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1968).
15. Sirisena, H. R., *Intern. J. Control,* **8,** 65 (1968). (1968).
16. Weber, A. P. J., "Suboptimal Control of Nonlinear and Inequality Constrained Control Systems," Ph.D. Dissert., Princeton University, Princeton, New Jersey (1969).
17. Weber, A. P. J., and L. Lapidus, *AIChE J.,* **17,** No. 4, 641 (1971).
18. Wonham, W. M., and C. D. Johnson, *J. Basic Eng., Trans. Am. Soc. Mech. Engrs.,* Series D, **86,** 107 (1964).

Manuscript received January 19, 1970, revision received April 27, 1970; paper accepted April 29, 1970.

Molecular Thermodynamics of Gases in Mixed Solvents

J. P. O'CONNELL

University of Florida, Gainesville, Florida

The general thermodynamic relation between Henry's constant for a solute in a mixed solvent and Henry's constants for the solute in the pure solvents is discussed in terms of limiting activity coefficients. The inadequacy of several common excess free energy expressions is shown, and results are presented for a form of Kirkwood-Buff solution theory based on pair distribution functions. Several unknown integrals in the theory are approximated by an empirical function of solvent composition and of solvent properties which generally predicts experimental data for gases in both simple and complex binary solvents with good accuracy. Prediction of fugacities of gaseous components in ternary and higher solvents and in nonideal solutions is considered.

Many chemical processes involve separations of solutes from multicomponent solvents. Distillation or stripping of gases, crystallization of solids, or extraction of slightly soluble substances all can involve this type of system. In these cases it is often convenient to use Henry's constant in the mixed solvent as the reference fugacity for the dilute component. This is based on experience with pure solvents which indicates that deviation from Henry's law often occurs only when significant amounts of solute are dissolved in the liquid. Although several treatments of solutes, particularly gases, in mixed solvents have appeared in recent years (*1* to *7*), most have dealt with particular equations for the excess free energy in the multicomponent mixture which are of limited application.

In this paper the general thermodynamic relationship for a solute in a mixed solvent is presented and the approximate equations which lead to previous results for gases are indicated. A general solution theory, based on the method of Kirkwood and Buff (*8*) but derived for a multicomponent solvent, is then used to obtain the mixed solvent Henry's constant in terms of pure solvent Henry's constants and integrals of the pair distribution functions of the system. Finally, since it is not possible to evaluate some of these integrals presently, particularly for complex mixtures, an empirical expression is developed for gases in mixed solvents which contains only readily accessible properties of the solvents and the solvent mixture. The results of this expression are compared with data for several systems of a gas in a binary solvent and the accuracy is better than with other methods not involving parameters from the multicomponent data. One multicomponent parameter is often sufficient to provide extremely good agreement for data over the entire range of solvent composition.

REFERENCE FUGACITY FOR A SOLUTE IN MIXED SOLVENTS

The primary aim in this work is to develop a method of predicting the fugacity of a solute in a mixed liquid solvent. This fugacity is equal to that in any other phase (solid, liquid, or gas) and is used to determine the distribution of the solute between the phases involved. The liquid fugacity can be expressed as

$$f_i^L = x_i\gamma_i(P^0)f_i^0(P^0)\exp\int_{P^0}^{P}\frac{\overline{v}_i}{RT}\,dP \qquad (1)$$

where the activity coefficient and reference state fugacity are both given at the reference pressure, and the system temperature.

The product $\gamma_i(P^0)f_i^0(P^0)$ can be obtained from two commonly used conventions. One is the symmetric convention for activity coefficients where all components follow

$$\lim_{x_i\to 1}(\gamma_i f_i^0) = f_i^{0L}(P^0) \qquad (2)$$

The reference fugacity is the pure component fugacity of i as a liquid at the reference pressure. For subcritical com-