# Time Varying Terminally Restrained Optimal Control Assessment

J. A. Gibson* and W. Sew Hoy†
*University of Canterbury, Christchurch, New Zealand*

**Terminal state restraints are represented by square law penalty functions with operating states progressively offset to reduce terminal errors by an iterative updating procedure based on biased terminal error data. Terminal restraints and transversality conditions, bias values, and the free terminal time parameter combine in unified boundary condition and parameter updating sets. The optimal control seeking algorithm is based on the Min-*H* method and in the comprehensive form presented contains substrategies for predictive control updating and compensation for large departures from optimal conditions. Algorithm effectiveness and insensitivity to spurious parameter variations are tested using fixed and time varying parameter versions of the nonlinear Van der Pol system with both stationary and moving terminal restraints and with time varying Liapunov stability boundaries. The flexible algorithm structure allows deletion of substrategies to provide simpler algorithms for less complex applications.**

## I. Introduction

COMPUTATIONAL techniques using a gradient approach for solution of the free time terminal restrained problem have usually required the satisfaction of the one of the terminal restraints as an iteration stopping condition.[1-3] The particular penalty function technique[4] forming the gradient of the cost with respect to the final time stops each iteration when this gradient is nulled. A requirement in common is that the a priori estimates of the initial optimal control must be sufficiently well chosen to insure satisfaction of the arbitrary stopping conditions. A different normalized scaling approach[5] transforms the free final time problem into a fixed time problem[6,7] without the necessity for arbitrary stopping conditions.

In this paper normalized time scaling is used with a biased square law penalty function representation of the terminal restraints in a Min-*H* search for optimal control.[8] The free time parameter is included in the bias parameter updating set and the transversality conditions with the terminal restraints. Thereby, some unification of mathematical formulation is achieved and, in computation, simultaneous rather than consecutive nulling of boundary errors is allowed. Substrategies provide Hamiltonian compensation for the effect of penalty function bias updating[9] through the costate system, and predictive control updating.[10]

The effectiveness of this comprehensive version of the algorithm is fully demonstrated using the nonlinear Van der Pol (VDP) system often chosen as the illustrative example by contributors in this field. However, this treatment goes further by considering both fixed and time varying parameter versions of the VDP system so that both stable and/or unstable system characteristics are encountered in the search trajectories tracking both stationary and moving terminal target restraints. An improved tolerance to spurious parameter variation is also demonstrated in a full experimental sensitivity analysis. Biased methods, requiring only modest penalty function weighting values even for high solution accuracy, avoid the computation sensitivity problems[9] of unbiased terminal restraint[11] techniques.

Received Feb. 2, 1977; revision received July 25, 1977. 
Index category: Analytical and Numerical Methods.
*Reader in Electrical Engineering.
†Postgraduate Scholar.

In contributing an improved and flexible method for assessing optimal control for terminally restrained systems, it is envisaged that appropriate versions of the biased penalty function algorithm will be used in off-line system design exercises establishing ideal performance data as a guide to specifying equipment performance and in on-line scheduling problems such as optimal control steering in aeronautical navigation systems.

## II. Algorithm Synthesis

### A. Problem Statement and Necessary Conditions

In the penalty function approach, the free time terminally restrained problem is to find the set of $m$ control variables $u(t)$ and the final time $T$ minimizing the cost functional:

$$J=\theta(x,T)+\int_0^T \phi(x,u,t)\,\mathrm{d}t \tag{1}$$

subject to $n$ first-order state equations

$$\dot{x}=f(x,u,t), \quad x(0)=x_0 \tag{2}$$

and with the Bolza cost $\theta$ suitably chosen to penalize restraint violation; e.g., in the biased penalty function approach to be used,

$$\theta=\left[h(T)+b\right]'\frac{M}{2}\left[h(T)+b\right] \tag{3}$$

where $h(T)$ is the set of errors in satisfying the set of $q$ terminal restraints

$$h(x,t)=0 \quad \text{at} \quad t=T \tag{4}$$

and $b$ is the set of $q$ restraint error bias values; diagonal $M$ is a $q\times q$ penalty weighting matrix; $T$, the final time, is a problem variable. The scalar Hamiltonian function is

$$H=\phi+p'f \tag{5}$$

providing a) optimal control via Pontryagin's minimum principle

$$\partial H/\partial u=0 \tag{6}$$

and b) $n$ first-order costate equations,

$$\dot{p}=-\partial H/\partial x=-\partial\phi/\partial x-(\partial f/\partial x)'p \tag{7}$$

There are $n+1$ transversality conditions at the terminal time:

$$p=\partial\theta/\partial x=(\partial h/\partial x)'M(h+b);\quad t=T \tag{8a}$$

$$H=-\partial\theta/\partial T=(\partial h/\partial T)'M(h+b);\quad t=T \tag{8b}$$

Adopting the control iteration method, reverse time ($\tau=-t+T$) computation of the costate system yields

$$p(\tau)=\int_0^{\tau}\left\{\partial\phi/\partial x+(\partial f/\partial x)'p(\tau)\right\}\mathrm{d}\tau+p(t=T) \tag{9}$$

Using the terminal costate condition in Eqs. (8), the costate system in Eq. (9) can be separated into terminal error ($h$) and terminal bias ($b$) dependent sections.

$$p_h(\tau)=\int_0^{\tau}\left\{\frac{\partial\phi}{\partial x}+\left(\frac{\partial f}{\partial x}\right)'p_h(\tau)\right\}\mathrm{d}\tau+\left\{\left(\frac{\partial h}{\partial x}\right)'Mh\right\}_T$$

$$p_b(\tau)=\int_0^{\tau}\left\{\left(\frac{\partial f}{\partial x}\right)'p_b(\tau)\right\}\mathrm{d}\tau+\left\{\left(\frac{\partial h}{\partial x}\right)'Mb\right\}_T$$

with

$$p=p_h+p_b \tag{10}$$

The purpose of this separation is to identify how bias values effect the Min-$H$ assessment through the costate system, so that in Sec. II E a compensating control correction [Eq. (31)] can be incorporated as a substrategy in the comprehensive algorithm.

### B. Transformation to a Normalized Time Scale

Introduce a normalized time scale variable

$$s=t/T;\quad 0\le s\le 1 \tag{11}$$

whereby in $s$-time the duration of the free $t$-time terminal restraint problem is fixed.

In $s$-time the transformed problem is to minimize

$$J=\theta(x,T)+T\int_0^1\phi(x,u,sT)\,\mathrm{d}s \tag{12}$$

with state and costate restraint equations:

$$\overset{\Delta}{x}=\partial H/\partial p=Tf(x,u,sT) \tag{13a}$$

$$\overset{\Delta}{p}_h=-(\partial H/\partial x)_h=-T\{(\partial\phi/\partial x)+(\partial f/\partial x)'p_h\} \tag{13b}$$

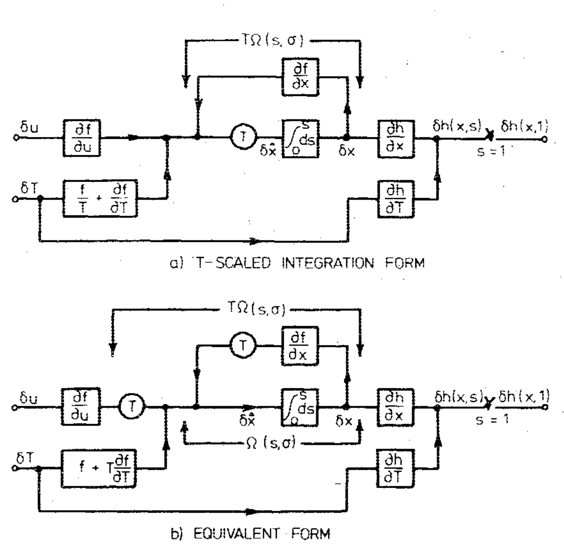



**Fig. 1 Schematic: perturbed $s$-time state system.**

$$\overset{\Delta}{p}_b=-(\partial H/\partial x)_b=-T(\partial f/\partial x)'p_b \tag{13c}$$

where $\Delta$ denotes the $s$-time differential operator d/d$s$.

The problem boundary and transversality conditions are

1) $s=0$

$$x(o)=x_o \tag{14}$$

2) $s=1$

$$h=h(x,T)=0 \tag{15}$$

$$H+\partial\theta/\partial T=\phi+p'f+\partial\theta/\partial T=0 \tag{16}$$

$$p=\partial\theta/\partial x \tag{17}$$

Combine Eqs. (16) and (17) into a single condition,

$$E=\phi+(\partial\theta/\partial x)'f+(\partial\theta/\partial T)=0 \tag{18}$$

and note that for the chosen case of a square law Bolza terminal cost function,

$$E=\phi+\{M(h+b)\}'\{(\partial h/\partial x)f+(\partial h/\partial T)\}=0$$

$$=E(x,u,b,T) \tag{19}$$

### C. Perturbed System Equations in $s$-time

The response to control, terminal time, and terminal error bias updating increments $\delta_u$, $\delta_T$, $\delta_b$ in the iterative solution seeking procedure following will be determined from the corresponding perturbed problem equations.

1) Perturbed state equations

$$\delta\overset{\Delta}{x}=\{f+T(\partial f/\partial T)\}\delta T+T\{(\partial f/\partial x)\delta x+(\partial f/\partial u)\delta u\} \tag{20}$$

with

$$\delta x(1)=\int_0^1(\delta\overset{\Delta}{x})\,\mathrm{d}s+\delta x(0)$$

i.e.,

$$\delta x(1)=T\Omega(1,\sigma)*\{(\partial f/\partial u)\delta u+(f/T+\partial f/\partial T)\delta T\} \tag{21}$$

where $\{T\Omega(1,\sigma)\}$ is the delayed ($\sigma$) transition matrix of the $s$-time perturbed state system in Fig. 1a evaluated at $s=1$.

2) Perturbed boundary conditions ($s=1$)

$$\delta h(1)=(\partial h/\partial x)\delta x+(\partial h/\partial T)\delta T \tag{22a}$$

$$\delta E(1)=(\partial E/\partial x)'\delta x+(\partial E/\partial u)'\delta u+(\partial E/\partial b)'\delta b$$

$$+(\partial E/\partial T)'\delta T \tag{22b}$$

For a square law Bolza cost in Eq. (3)

$$(\partial E/\partial x)=\partial\phi/\partial x+\{M(\partial h/\partial x)\}'\{(\partial h/\partial x)f+(\partial h/\partial T)\}$$

$$+\{(\partial^2 h/\partial x^2)f+(\partial h/\partial x)(\partial f/\partial x)+(\partial^2 h/\partial T\partial x)\}'M(h+b)$$

where

$$(\partial^2 h/\partial x^2)f=\left[\left\{\frac{\partial}{\partial x_1}\left(\frac{\partial h}{\partial x}\right)\right\}f:\left\{\frac{\partial}{\partial x_2}\left(\frac{\partial h}{\partial x}\right)\right\}f:$$

$$\cdots:\left\{\frac{\partial}{\partial x_n}\left(\frac{\partial h}{\partial x}\right)\right\}f\right] \tag{23a}$$

$$(\partial E/\partial u)=\partial\phi/\partial u+\{(\partial h/\partial x)(\partial f/\partial u)\}'M(h+b) \tag{23b}$$

$$(\partial E/\partial b)=M'\{(\partial h/\partial x)f+(\partial h/\partial T)\} \tag{23c}$$

$$(\partial E/\partial T) = \partial\phi/\partial T + \{M(\partial h/\partial T)\}'\{(\partial h/\partial x)f + (\partial h/\partial T)\}$$
$$+\{(\partial^2 h/\partial x\partial T)f + (\partial h/\partial x)(\partial f/\partial T)$$
$$+(\partial^2 h/\partial T^2)\}'M(h+b) \tag{23d}$$

It is noted in the alternative schematic, Fig. 1b, that since the perturbed state (20) and the costate (13) systems are adjoint, the transpose of the unit initial condition response of the $p_b$ section can be used to generate the transition matrix influence coefficients $\Omega(1,s)$ entering into schematic Fig. 1b and the perturbed state and terminal restraint Eqs. (21) and (22).

### D. Synthesis of an Augmented Terminal Restraint Set

Define an augmented bias updating set as

$$\delta b^+ = [\delta b' : \delta T]' \tag{24}$$

whereby the free time updating parameter $\delta T$ is combined with the bias updating set $\delta b$. Treat the combined transversality condition [Eq. (18)] as a terminal restraint and postulate an augmented set of terminal restraints

$$h^+ = [h' : E]' = 0 \tag{25}$$

and, corresponding to Eq. (19), write

$$h^+ = h^+(x,u,b^+) = 0 \tag{26}$$

The perturbed augmented terminal restraint set is

$$\delta h^+ = (\partial h^+/\partial x)\delta x + (\partial h^+/\partial u)\delta u + (\partial h^+/\partial b^+)\delta b^+ \tag{27}$$

with

$$\partial h^+/\partial b^+ = [\partial h^+/\partial b : \partial h^+/\partial T] \tag{28a}$$

$$\partial h^+/\partial b = [\partial h/\partial b : \partial E/\partial b]' = [0 : \partial E/\partial b]' \tag{28b}$$

$$\partial h^+/\partial T = [(\partial h/\partial T)' : \partial E/\partial T]' \tag{28c}$$

$$\partial h^+/\partial u = [(\partial h/\partial u)' : \partial E/\partial u]' = [0' : \partial E/\partial u]' \tag{28d}$$

$$\partial h^+/\partial x = [(\partial h/\partial x)' : \partial E/\partial x]' \tag{28e}$$

Note that writing $\delta T = 0$ implies $b^+ = b$, $h^+ = h$, and the treatment reduces to one with fixed terminal times.

### E. Predictive and Compensatory Min-*H* Control Updating

Incorporating optional predictive and Hamiltonian input data compensating substrategies in the Min-*H* control updating

$$\delta u = I(\delta u^* + \delta u_b) \tag{29}$$

where 1) $I$ is a scale factor based on a linear interpolative prediction[11] of the control increment needed to satisfy the Min-*H* requirement at the next iteration; 2) $\delta u^*$ is obtained directly from Eq. (6) or by a Min-$H$[8] hill descending technique,

$$\delta u^* = u^* - u \tag{30}$$

where $u$ is the control used at the current iteration, and 3) $\delta u_b$ compensating for costate increments from bias updating avoids unnecessary departures of Min-*H* input data from estimated optimal sets,

$$\delta u_b = (\partial u^*/\partial b^+)\delta b^+ \tag{31}$$

for which the sensitivity matrix $(\partial u^*/\partial b^+)$ is determined analytically or by sensitivity test perturbations $(\delta b^+)_s$ in complex problems.

Thus from Eqs. (29-31) the overall control updating increment is

$$\delta u = I\{\delta u^* + (\partial u^*/\partial b^+)\delta b^+\} \tag{32}$$

Predictive and/or compensating strategies may be omitted by writing $I = 1$ and $(\partial u^*/\partial b^+) = 0$ throughout this treatment.

### F. Incremental Terminal State Response to Control Updating

Substituting from Eq. (32) into Eq. (21),

$$\delta x(1) = T\Omega(1,\sigma)*[(\partial f/\partial u)I\{\delta u^* + (\partial u^*/\partial b^+)\delta b^+\}$$
$$+ (f/T + \partial f/\partial T)\delta T] \tag{33}$$

Express now the final time updating increment $\delta T$ as an element of the augmented bias updating set $\delta b^+$,

$$\delta x(1) = T\Omega(1,\sigma)*[\{(\partial f/\partial u)I\}\delta u^* + \{(\partial f/\partial u)I(\partial u^*/\partial b^+)$$
$$+ (\partial f/\partial b^+)\}\delta b^+]$$

in which

$$(\partial f/\partial b^+) = [0 : \partial f/\partial T]' = [0 : f/T + (\partial f/\partial T)]' \tag{34}$$

Thus, the response of the state system is expressed in terms of increments $\delta u^*$, $\delta b^+$ from the Min-*H* and augmented bias updating strategies.

### G. Augmented Bias Updating Strategy

Substituting $\delta u$, $\delta x$ in Eqs. (32) and (34) into Eq. (27):

$$\delta h^+(1) = \left(\frac{\partial h^+}{\partial x}\right) T\Omega(1,\sigma)*\left[\left\{\frac{\partial f}{\partial u}I\right\}\delta u^* + \left\{\frac{\partial f}{\partial u}I\frac{\partial u^*}{\partial b^+}\right.\right.$$
$$\left.\left. + \frac{\partial f}{\partial b^+}\right\}\delta b^+\right] + \left(\frac{\partial h^+}{\partial u}\right)\left[I\left\{\delta u^* + \left(\frac{\partial u^*}{\partial b^+}\right)\delta b^+\right\}\right]$$
$$+ \left(\frac{\partial h^+}{\partial b^+}\right)\delta b^+ = Q^+ + R^+\delta b^+ \tag{35}$$

in which

$$Q^+ = I\left[\frac{\partial h^+}{\partial u} + \frac{\partial h^+}{\partial x} T\Omega(s,\sigma)*\left\{\frac{\partial f}{\partial u}\right\}\right]\delta u^* \tag{36}$$

$$= I\left[\left(\frac{\partial h^+}{\partial u}\right)\delta u^* + \left(\frac{\partial h^+}{\partial x}\right) T\int_0^1 \left\{\Omega(1,\sigma)\times\left(\frac{\partial f}{\partial u}\right)\delta u^*\right\}d\rho\right] \tag{37}$$

and

$$R^+ = I\left[\frac{\partial h^+}{\partial u}\frac{\partial u^*}{\partial b^+} + \frac{\partial h^+}{\partial b^+} + \frac{\partial h^+}{\partial x} T\Omega(1,\sigma)*\left\{\frac{\partial f}{\partial u}\frac{\partial u^*}{\partial b^+}\right.\right.$$
$$\left.\left. + \left(\frac{1}{I}\right)\frac{\partial f}{\partial b^+}\right\}\right] \tag{38}$$

$$= I\left[\frac{\partial h^+}{\partial u}\frac{\partial u^*}{\partial b^+} + \frac{\partial h^+}{\partial b^+} + \frac{\partial h^+}{\partial x} T\int_0^1 \Omega(1,\sigma)\right.$$
$$\left.\times\left\{\frac{\partial f}{\partial u}\frac{\partial u^*}{\partial b^+} + \left(\frac{1}{I}\right)\left(\frac{\partial f}{\partial b^+}\right)\right\}d\rho\right] \tag{39}$$

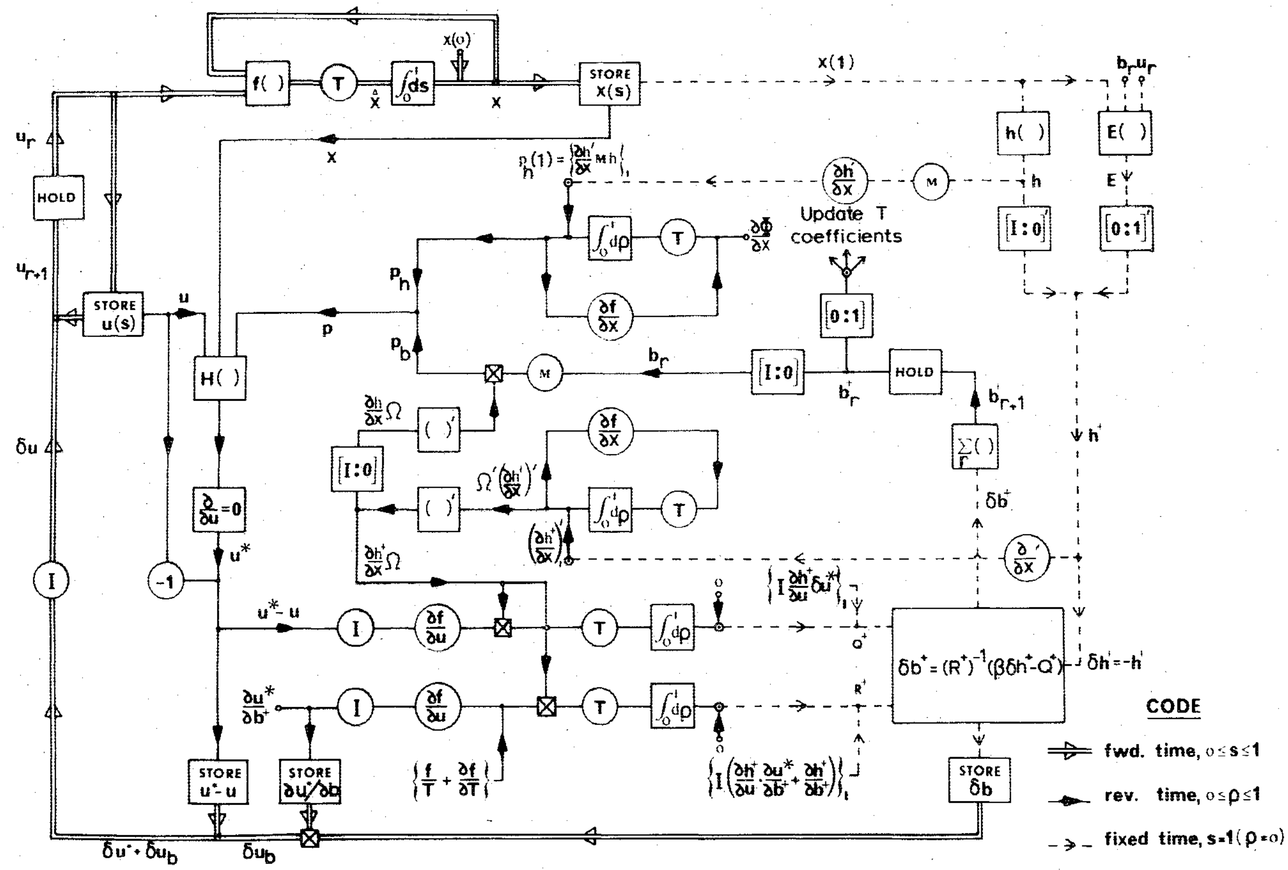



Fig. 2 Schematic: the comprehensive algorithm.

with normalized reverse time

$$\rho = 1 - s \tag{40}$$

Providing the state system is $\delta h^+$ controllable it follows from Eq. (35) that

$$\delta b^+ = (R^+)^{-1}(\delta h^+ - Q^+) \tag{41}$$

The augmented bias updating strategy requires that the updating set $(\delta b^+)_{r+1}$ for iteration $r+1$ be chosen so that to a first-order final time updating $(\delta T)_{r+1}$ and the overall control updating $(\delta u)_{r+1}$ generate augmented terminal restraint increments $(\delta h^+)_{r+1}$, cancelling current boundary condition errors $(h^+)_r$, viz,

$$(\delta h^+)_{r+1} = -\beta(h^+)_r \quad \text{at } s=1 \tag{42}$$

where $\beta(=1$ ideally) is an attenuating factor from an auxilliary algorithm desensitizing strategy[9] arranged to operate automatically in early iterations when the estimates of optimal control may be poor. From Eqs. (41) and (42) the incremental bias set is determined:

$$(\delta b^+)_{r+1} = -\{(R^+)^{-1}(\beta h^+ + Q^+)\}_r \tag{43}$$

**H. Overall Structure of the Comprehensive Algorithm**

The problem equations and solution-seeking strategies involved are consolidated for the case of a square law Bolza terminal cost function in the flow diagram schematic, Fig. 2, from which the elements in the overall structure of the biased penalty function algorithm are identified in Appendix B.

## III. Demonstrations of Algorithm Effectiveness

The following demonstrations use the comprehensive version of the algorithm to assess the optimal control characteristics of the terminally restrained VDP system described in Appendix A. An EAI 640 serial digital computer with a 32-bit word length was used and all integrations were performed with a fourth-order Runge-Kutta routine employing 100 steps. For all of the demonstration examples the following stopping criteria for accuracy were adopted:

$$(\delta u)_{rms} = \left[\frac{\sum_{i=1}^{100}(u_i^* - u_i)^2}{100}\right]^{1/2} < 10^{-2} \tag{44a}$$

$$(h^+)'h^+ < 10^{-8} \tag{44b}$$

Small penalty weights $\mu_1 = 1$, $\mu_2 = 1$ were chosen for all cases.

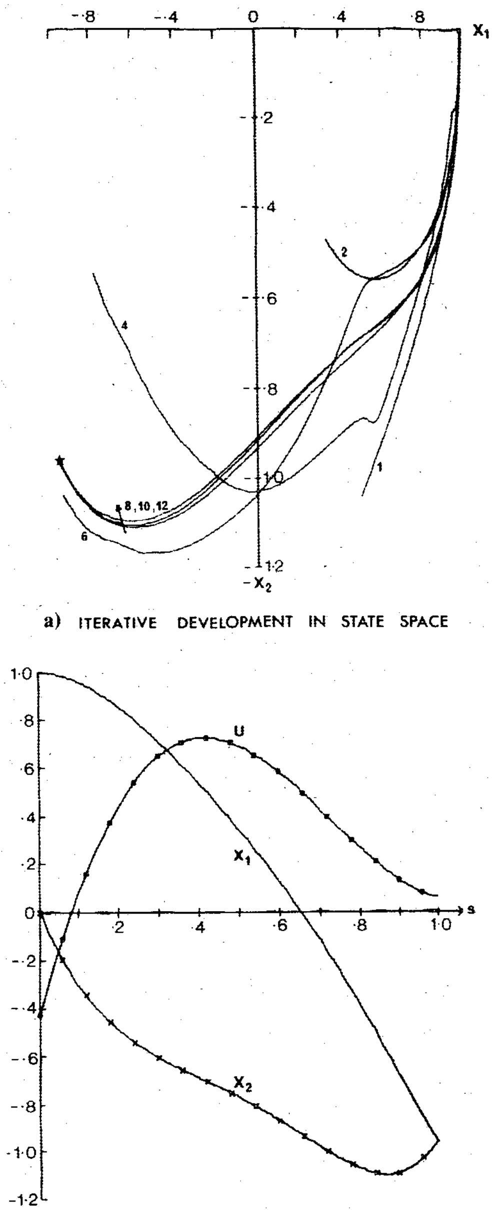
**A. Time Invariant System, Stationary Terminal Restraint**

For this demonstration case, the boundary restraints on the VDP system in Appendix A are

$$X(0) = \begin{bmatrix} 1 \\ o \end{bmatrix}, \quad h(x,T) = \begin{bmatrix} x_1(1) + 0.97 \\ x_2(1) + 0.96 \end{bmatrix} \tag{45}$$

The progression of 12 iterative solutions, Fig. 3a, in the development of optimal control achieved a very satisfactory terminal restraint error index of $(h^+)'h^+ = 0.43E-9$, with components $h'h = 0.17$, $E-9$ and $E'E = 0.26$ $E-9$. The greater part of the development was, as shown, attained in only 8 iterations with $(h^+)'h^+ = 0.17$ $E-4$. The augmented bias parameter set $(b^+)' = (b_1, b_2, T)$ developed from initial values (0, 0, 1) rapidly to ($-0.1846E1$, $0.9565E-1$, $0.2640E1$) at iteration 8 and then trimmed to final values ($-0.1850E1$, $0.9218E-1$, $0.2632E1$) at iteration 12, with a cost of $J = 3.38$.

The final optimal control and state trajectories shown in Fig. 3b are interesting in that maximum control effort is



a) ITERATIVE DEVELOPMENT IN STATE SPACE

b) THE OPTIMAL TRAJECTORIES

Fig. 3 Time invariant system, stationary terminal restraints.

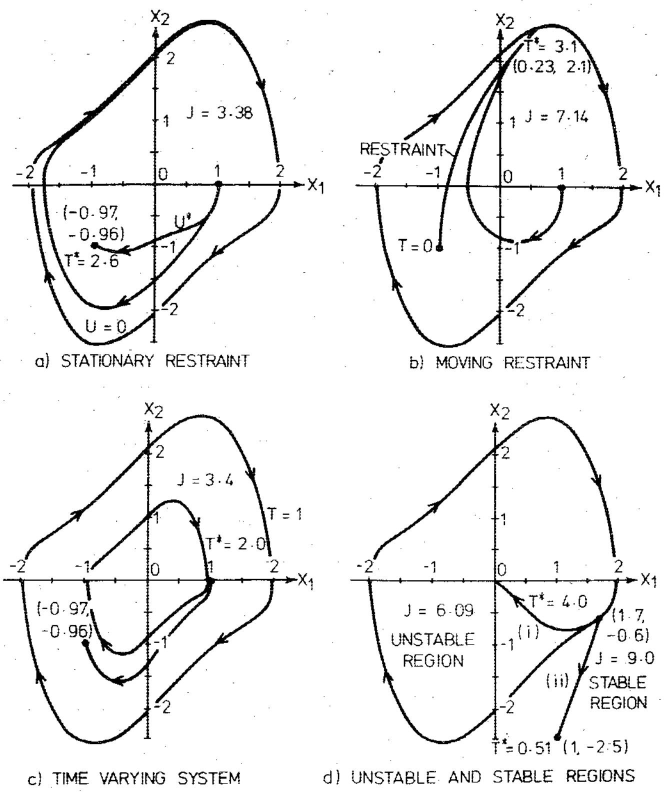



Fig. 4 Summary of cases.

called for in midperiod $s \doteq 4$, rather than toward the end of the period of optimal control ($s = 1$) when a greater control effort could have been expected in order to satisfy the restraints. Instead, with declining control effort, the controlled state trajectories merge with the trajectories (Fig. 4a) of the uncontrolled VDP system passing through the terminal restraint coordinate. The initial negative control ($0 < s < 0.08$) is reasoned as being necessary to reduce the large values of $x_1$ incurring initially the greatest cost in the quadratic cost functional in the appendix.

Thus, the performance of the algorithm is demonstrated first in the solution of a problem previously solved in earlier work[6] using a modified quasilinearization approach. In the demonstrations, rough initial estimates of optimal control ($u = 0$, $0 \leq s \leq 1$), final time ($T = 1$), control predictive factor ($I = 0.5$), and bias parameters ($b_1 = 0$, $b_2 = 0$) were found to be satisfactory as the starting parameter set. The quasilinearization method is reported to have failed unless starting final time $T$ estimates were very close to $T = 2.5$. The state portrait of the optimal solution is entered into the consolidation diagram, Fig. 4a, for comparison with the remaining demonstrations.

## B. Time Invariant System, Moving Terminal Restraint

In this demonstration the terminal restraints are time varying:

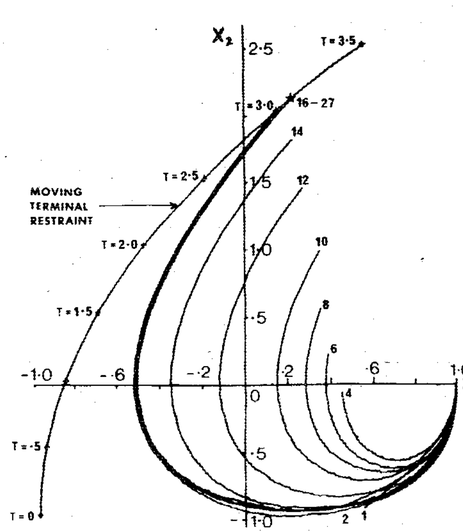
$$h(x,T) = \begin{bmatrix} x_1(1) + 0.97 - T^2/8 \\ x_2(1) + 0.96 - T \end{bmatrix} \tag{46}$$

The iterative development to the optimal portrait in state space is evident in Fig. 5a. A total of 27 iterations were required to catch the moving terminal restraint with a terminal restraint error index $(h^+)'h^+ = 0.90E-7$ comprising $h'h = 0.23E-8$ and $E'E = 0.84E-7$. Large changes in the state trajectories generated in the first four iterations are indicative of the action of the algorithm to revise the poor initial estimate ($u = 0$) of optimal control. Thereafter, from iterations 4 to 16 a more uniform progression, dominated by the final time $T$ updating strategy, to null the moving terminal restraint errors is evident. A final trimming phase involving a reduction of error index $(h^+)'h^+ = 0.26E-3$ at iteration 16 to the accepted value $E-8$ at iteration 27 is clear. Augmented bias values accumulated from $(b_1, b_2, T) = (0, 0, 1)$ initially through $(-0.6951E1, 0.4123E1, 0.1403E1)$ at iteration 4 and $(-0.1322E1, 0.3530E1, 0.3031E1)$ at iteration 16 to the final values $(-0.1051E1, 0.3573E1, 0.3095E1)$ at iteration 27 with cost $J = 7.14$.

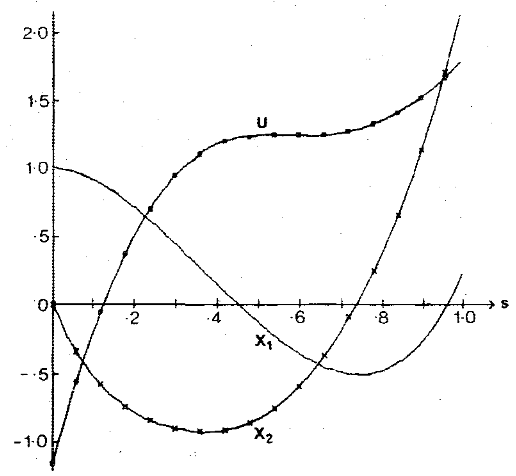
Optimal control and state trajectories are recorded in Fig. 5b. Compared with the previous fixed-restraint case the optimal control is interesting in showing that extra control effort, particularly toward the end period, is required to catch the moving target restraint. The optimal state portrait is entered into the consolidation diagram, Fig. 4b, and shows the result in relation to the uncontrolled VDP system.

## C. Time Vaying System, Fixed Terminal Restraints

For this demonstration the square law component of the amplitude dependent damping factor in the VDP system is modified so that the Liapunov boundary between unstable and stable regions of the state system varies with time. The time variance is chosen, Fig. 4c, so that during the period of

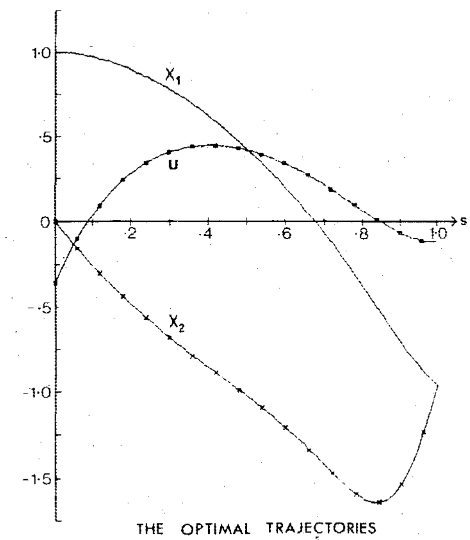



a) ITERATIVE DEVELOPMENT IN STATE SPACE



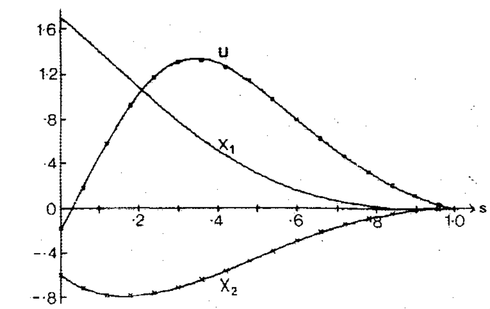

b) THE OPTIMAL TRAJECTORIES

**Fig. 5 Time invariant system, moving terminal restraints.**

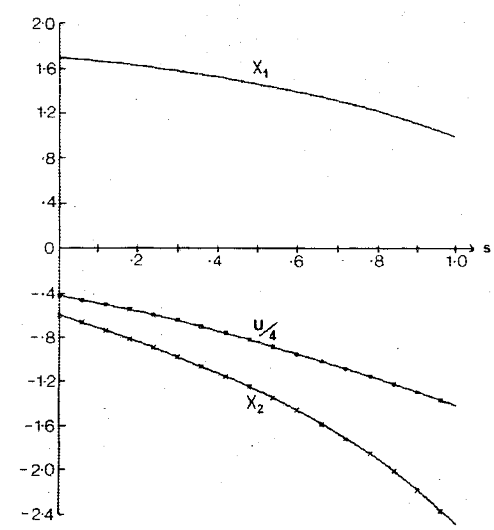



THE OPTIMAL TRAJECTORIES

**Fig. 6 Time varying system, stationary terminal restraints.**




a) UNSTABLE REGION : OPTIMAL TRAJECTORIES




b) STABLE REGION : OPTIMAL TRAJECTORIES

**Fig. 7 Unstable and stable system.**

optimal control the boundary contracts across the initial and terminal restraint locations. In this way it is believed that explicit system time variance is introduced significantly by causing the state system to change its stability characteristics radically in the course of the optimal control assessment. The state system is now

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = (1 - t^2 x_1^2)x_2 - x_1 + u \tag{47}$$

and the boundary restraints are those given in Sec. III-A.

For this more difficult case the following estimates of parameter values form a satisfactory initial parameter set: $u=0$, $I=0.1$, $T=1$, $b_1=\mu_1 h_1=0.1351E1$, $b_2=\mu_2 h_2=-0.3435E0$. The optimal trajectories, Fig. 6, for this case are similar to the corresponding time invariant case in Sec. III-A and Fig. 3b. In achieving an error index $(h^+)'h^+=0.37E-8$ comprising $h'h=0.21E-9$ and $E'E=0.34E-8$, a total of 25 iterations were required. The final augmented bias set was $(b_1, b_2, T)=(-0.2842E-1, -0.2342E0, 0.2036E1)$ with a cost of $J=0.3469E1$. The contracting of the stability boundaries actually helped the assessment since the family of stability boundaries approach a near correspondence with the optimal trajectory, Fig. 4c. Hence the control effort recorded in Fig. 6 is less than that shown in Fig. 3b. The result was surprising in that computation of an unstable state system was expected to yield larger errors than in the computation of a stable case and thus to present a more difficult task for the biased penalty function algorithm seeking the optimal control. This point is therefore examined further.

**D. Sensitivity to Stability of the State System**

The test cases, indicated i and ii in Fig. 4d, have a common initial state (1.7,-0.6) approximately on the stability boundary and fixed terminal restraints at (0, 0), (1,-2.5) roughly equidistant in state space from the initial state. Thus, starting on the boundary between unstable regions the two cases are chosen to involve a similar depth of penetration, respectively, into the regions so as to detect the relative sensitivity of the algorithm to the stability of the state system under optimal control. The state system here was the time invariant case in Appendix A.

The iterative development leading to the optimal trajectories in Fig. 7 was satisfactory in both cases starting from the initial parameter set: $u=0$, $I=0.5$, $T=1$, $b_1=0$, $b_2=0$.

1) Unstable region: Convergence to an error index $(h^+)'h^+=0.60E-9$ comprising $h'h=0.50E-9$ and $E'E=0.10E-9$ was achieved after 37 iterations. The final augmented bias set was $(b_1, b_2, T)=(0.4657E0, 0.145E\text{-}3, 0.4056E1)$ with a cost of $J=6.10$. In this unstable region the control effort in Fig. 7a is, as expected, similar to the previous fixed terminal restraint case in Fig. 3b.

2) Stable region: Convergence to an error index $(h^+)'h^+=0.68E-9$ comprising $h'h=0.75E-10$ and $E'E=0.60E-9$ was achieved after 14 iterations. The final augmented bias set was $(b_1,b_2,T)=(0.1466E2, 0.1140E2, -0.51150E0)$ with a cost $J=9.03$. The control effort, Fig. 7b, increases to a large value in order to steer the state system towards the terminal restraint while overcoming the opposing and increasing state feedback steering components of the uncontrolled VDP system.

It is interesting to note from the data in Fig. 4d that the duration of optimal control for case ii was nearly an order of 10 less than for case i, but the accumulated costs $J$ are similar in magnitude. These results are compatible with the square law costing on each of the state and control variable chosen as the cost function $\phi$ in the appendix. The cost of the large control effort in case ii to reduce the time spent in the expensive terminal restraint region of state space is, of course, justified by the cost function.

The results of these sensitivity tests indicate that, at least for the VDP illustrative example, the iteration sensitivity of the biased penalty function algorithm to system stability regions is not as great as expected in Sec. III-C. The reason why is believed to be a self-compensating characteristic of optimal control assessments in that, if the perturbed state system is stable, then the costate system is unstable and vice versa.

### E. Sensitivity to the Initial Estimate of Terminal Time

From the data presented in Table 1 the iteration sensitivity of the algorithm is satisfactorily small over a wide range $0.3 \le T \le 5$ of initial estimates of the duration of optimal control for the fixed restraint case in Sec. III-A with $T=2.63$. This asymmetrical result suggests that sensible underestimating is an advantage in choosing the initial terminal time value. Notwithstanding this aspect, the algorithm provides a definite improvement in sensitivity margin over the modified quasilinearization method[6] requiring an additional desensitizing strategy for convergence to the optimal solution from a starting final time of $T=2$.

### F. Sensitivity to the Initial Estimate of Optimal Control

Involving the properties of a complete control trajectory, an informative statement on control sensitivity is difficult to make. It should be noted, however, that all test demonstrations achieved convergence for a poor initial estimate of control, $u=0$. For practical problems a more intelligent estimate would normally be possible and thereby the convergence of the biased penalty function algorithm would be helped.

### G. Sensitivity to the Initial Estimate of Bias Values

A convenient estimate of initial bias values is the set $b=0$, equivalent to the unbiased penalty function approach, or alternatively the set $b=Mh$ thereby beginning with the basic bias approach.[9] Both cases were tried and found to be satisfactory, although in the demonstrations the particular bias set adopted was the one found to be the most suitable. Of course, if good initial estimates of bias values could be made from prior knowledge of the system characteristics, then improved iteration convergence would occur.

## IV. Discussion

The biased penalty function algorithm is unique compared with the alternatives referenced in this paper, in containing cumulative bias updating, predictive control updating, and Hamiltonian compensation strategies, each of which enhances convergence to the optimal control solution. Although it has been presented and tested in this comprehensive form, the algorithm design is flexible because the predictive and/or the compensation strategies are easily omitted for simple applications. The results presented are believed to be the only comprehensive and sensitivity-tested demonstrations in the optimal control literature of stable and unstable time varying systems with both stationary and moving terminal restraints.

Convergence limitations remain in the comprehensive algorithm. They originate from those of the first-order Min-$H$ method requiring at all times that the iterative estimates of optimal control are not far removed from the true optimal conditions. The Hamiltonian compensation for bias parameter updating greatly improves convergence in this respect, but even with the compensation the $\beta$-search [Eq. (42)] strategy desensitizing parameter updating is needed to achieve convergence with poor initial estimates of optimal control.

Although control and state restraints have not been included in this treatment, the basic penalty function relaxation method[8,12] is well suited to restraint implementation providing that an adequate restraint error index is applied to insure restraint violation is tolerable. The authors are actively researching the possibility of using automatic control and state restraint biasing with the same objective as in the case of terminal restraints, of achieving insensitive high-accuracy solutions while benefiting from the simplicity of the penalty function relaxation method.

## V. Conclusions

A comprehensive biased penalty function algorithm with predictive and compensatory substrategies has been carefully formulated stage by stage, performance and sensitivity tested in detail, and shown to be effective in determining optimal control of constant and time varying parameter versions of a nonlinear Van der Pol test example, subject to both stationary and moving terminal restraints. The composition of the algorithm is flexible with scope for adaptation to suit requirements of particular applications. These are believed to be principally in early system design exercises establishing

**Table 1 Sensitivity to initial estimates of final time[a]**

| $T$ | Initial estimates of augmented bias values $b_1$ | $b_2$ | Iteration at $\beta-$ search termination | Iterations requiring $\delta b^+$ attenuation | Iterations required for convergence |
|---|---|---|---|---|---|
| 0.25 | $0.1940E1$ | $0.7116E0$ | ... | ... | unsuccessful |
| 0.3 | $0.1925E1$ | $0.6625E0$ | 13 {14}[b] | 5 {6} | 17 {18} |
| 0.5 | $0.1846E1$ | $0.4660E0$ | 11 {13} | ... | 17 {19} |
| 1.0 | 0 | 0 | 5 { 5} | ... | 12 {12} |
| 2.0 | 0 | 0 | 16 {16} | ... | 21 {21} |
| 3.0 | $-0.7580E0$ | $0.1375E1$ | 17 {18} | 5 {6} | 24 {25} |
| 4.0 | 0 | 0 | 14 {14} | 3 {3} | 19 {19} |
| 5.0 | 0 | 0 | 23 {23} | 3 {3} | 27 {27} |
| 6.0 | 0 | 0 | ... | ... | unsuccessful |

[a] In all cases $u=0$, $I=0.5$. [b] { } indicates iterations with $b_1=b_2=0$ initially.

ideal optimal design data for equipment specification and in the scheduling of optimal control commands in operational systems.

## Appendix A: Van der Pol Illustrative Example

Performance index:

$$J=\int_0^T (x_1^2+x_2^2+u^2)\,dt$$

System state equations:

$$\dot{x}_1=x_2, \quad \dot{x}_2=(1-x_1^2)x_2-x_1+u$$

A specimen state space trajectory of the classical uncontrolled ($u=0$) Van der Pol system is given in Fig. 4a.

## Appendix B: Algorithm Structure

Figure 2 is a schematic of the composition of the overall algorithm, with readily identifiable principal sections relating to the state system $x$, unified terminal restraints $h^+$, the partitioned costate system $p_h$, $p_b$, Min-$H$ assessments $u^*$; augmented bias updating $\delta b^+$ through unified boundary condition error cancellation ($\delta h^+=-\beta h^+$); generation of the control bias compensation $\delta u_b$, and interpolative prediction ($I$) of the overall control updating $\delta u$. Particular details to note are 1) normalized $s$-time integration used; 2) terminal restraint unification $h^+$ is synthesized by the $[(q+1)\times 1]$ coupling operators $[I{:}0]'$, $[0{:}1]'$ partitioned respectively into $[q\times q{:}q\times 1]'$ and $[1\times q{:}1\times 1]'$ sections; 3) the augmented bias set $b^+$ is separated into bias $b$ and terminal time $T$ groups for parameter updating by $[(q+1)\times q]$ distribution operators $[I{:}0]$, $[0{:}1]$ partitioned respectively into $[q\times q{:}q\times 1]$ and $[1\times q{:}1\times 1]$ sections; 4) the augmented matrix of influence coefficients $(\partial h^+/\partial x)\Omega(1,\sigma)$ from the initial condition response of the $p_b$ costate section is similarly distributed to costate and to terminal error cancellation sections; 5) inputs to the $Q^+$, $R^+$ calculations cater for both time varying state systems and free terminal time specifications; 6) removal of the superscipts + reduces the representation to the fixed terminal time case; 7) writing $I=1$ removes the interpolative predictive control updating strategy; and 8) writing $(\partial u^*/\partial b)=[0]$ removes the Hamiltonian compensation strategy.

## Acknowledgment

The authors thank R. S. Long in the Department of Mathematics, University of Canterbury, for reading the manuscript.

## References

[1] Bryson, A. E. and Denham, W. F., "A Steepest-Ascent Method for Solving Optimum Programming Problems," *Journal of Applied Mechanics*, Vol. 29, June 1962, pp. 247-257.

[2] Gottlieb, R. G., "Rapid Convergence to Optimum Solutions Using a Min-*H* Strategy," *AIAA Journal*, Vol. 5, Feb. 1967, pp. 322-329.

[3] Chan, W. C. and Aly, G. M., "A Computational Method for Optimal Control Problems with a Free Final Time Using the Modified Quasilinearization and the Gradient Method," *International Journal of Control*, Vol. 18, Nov. 1973, pp. 1067-1075.

[4] Sage, A. P., *Optimal Systems Control*, 1st ed., Prentice Hall, Englewood Cliffs, N.J., 1968, p. 404.

[5] Long, R. S., "Newton-Raphson Operator: Problems with Undetermined End Points," *AIAA Journal*, Vol. 3, July 1965, pp. 1351-1352.

[6] Aly, G. M. and Chan, W. C., "Numerical Computation of Optimal Control Problems with Unknown Final Times," *Journal of Mathematical Analysis and Applications*, Vol. 45, Feb. 1974, pp. 274-284.

[7] Miele, A., Pritchard, R. E., and Damoulakis, J. N., "Sequential Gradient-Restoration Algorithm for Optimal Control Problems," *Journal of Optimization Theory and Applications*, Vol. 5, April 1970, pp. 253-282.

[8] Kelley, H. J., "Method of Gradients," *Optimization Techniques with Applications to Aerospace Systems*, edited by G. Leitman, Academic Press, New York, 1962, pp. 212-216.

[9] Gibson, J. A. and Sew Hoy, W., "Development of Biased Penalty Function Algorithms for Optimal Control with Terminal Restraints," *International Journal of Control*, Vol. 24, Nov. 1976, pp. 641-659.

[10] Gibson, J. A. and Sew Hoy, W., "An Extension of the Predictive Min. *H* Method to Multivariable Control," *International Journal of Control*, Vol. 21, Sept. 1975, pp. 795-799.

[11] Slater, G. L., "A Transformation Approach to the Terminal Control Problem," *AIAA Journal*, Vol. 14, Sept. 1976, pp. 1206-1209.

[12] Lowinger, J. F. and Gibson, J. A., "Desensitising Algorithms for State Restrained Optimal Control Assessments," *International Journal of Control*, Vol. 21, May 1975, pp. 353-373.