JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 15, No. 2, March-April 1992

Fuel Consumption in Optimal Control

Jim Redmond* and Larry Silverberg†
North Carolina State University, Raleigh, North Carolina 27695

This paper treats the fuel optimal control of dynamic systems based on three cost functions. The three cost functions are a saturation condition, pesudofuel, and absolute fuel. They lead to bang-bang control, continuous control, and impulse control, respectively. A comparison of fuel consumption based on absolute fuel reveals that impulse control consumes 35% less fuel than continuous control in the case of controlling a linear damped harmonic oscillator. Impulse control consumes 33% less fuel than continuous control in the case of a rest-to-rest rigid-body maneuver. This result suggests the potential for fuel savings in larger order systems through the implementation of impulsive control strategies.

I. Introduction

THE minimization of absolute fuel as well as other fuel functions was first introduced by Krasovskii in Ref. 1, in which the optimal control problem was formulated as an n-variable function minimization problem. This process was later detailed by Neustadt, who first solved the minimum time problem[2] and then extended his results to include various optimal control problems. However, his objective was to develop numerical methods, and so Neustadt avoided the development of optimal control based on absolute fuel because of numerical difficulties in the formulation.[3] The absolute fuel function was later considered by Hajek but only for bounded control inputs.[4] This paper treats absolute fuel minimization with unbounded control inputs as demonstrated by Chukwu[5] and compares the solution with two other fuel functions. Note that interest in absolute fuel minimization has been rekindled due in part to the recent development of close approximations of the fuel optimal solution associated with controlling the motion of flexible structures.[6]

Optimal control is reviewed in Sec. II along with two fundamental theorems associated with the properties of convex sets. A general procedure is developed for the minimization of any cost function satisfying the third property of a norm.[7] Three optimizing fuel functions are considered in Sec. III. They yield bang-bang, continuous, and impulse control strategies. These control strategies are applied to the vibration suppression of a lightly damped oscillator in Sec. IV and to the rest-to-rest maneuver of a rigid body in Sec. V. A comparison of fuel consumption among the three control strategies for the two examples considered is given in Sec. VI. Concluding remarks are given in Sec. VII.

II. Optimal Control Formulation

Consider the linear time-invariant system

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{1}$$

where $x(t)$ denotes the $n \times 1$ state vector at time t, $u(t)$ denotes the $m \times 1$ control vector, and A and B are $n \times n$ and $n \times m$ coefficient matrices, respectively. Assume that the system of Eq. (1) is controllable. That is, there exists a control $u(t)$ that transfers $x(0) = x_0$ to $x(T) = x_1$ in some time T. Furthermore, assume that $u(t)$ is contained in a convex set U of admissible controls. That is, for any $u_1(t)$ and $u_2(t)$ contained in U, and for any λ_1 and λ_2 greater than zero such that $\lambda_1 + \lambda_2 = 1$, the control $u(t) = \lambda_1 u_1(t) + \lambda_2 u_2(t)$ is also contained in U. The solution to Eq. (1) is[8]

$$x(t) = e^{At}\left(x_0 + \int_0^t e^{-As}Bu(s)\,\mathrm{d}s\right) \tag{2}$$

in which e^{At} is the state transition matrix. At time T, Eq. (2) leads to

$$y(T) = \int_0^T e^{-At}Bu(t)\,\mathrm{d}t \tag{3}$$

where

$$y(T) = e^{-AT}x_1 - x_0 \tag{4}$$

is referred to as the reachable state. The reachable set is defined as the collection of reachable states obtained from all of the admissible controls, that is,

$$R = \{y(T):y(T) = \int_0^T e^{-At}Bu(t)\,\mathrm{d}t, \quad \forall\ u(t)\epsilon U\}$$

Note that an admissible control is capable of transferring x_0 to x_1 in time T only if $y(T)$ in Eq. (4) is contained in R. The set of reachable states is also convex as a result of the convexity of U. This is verified by considering two points contained in the reachable set $y_1(T)$ and $y_2(T)$ obtained from the control inputs $u_1(t)$ and $u_2(t)$, respectively. Since $u_1(t)$ and $u_2(t)$ are elements of the convex set U, the control $u_3(t) = \lambda_1 u_1(t) + \lambda_2 u_2(t)$ is also in U. By definition, the reachable state $y_3(T)$ obtained from $u_3(t)$ is in R, and

$$y_3(T) = \int_0^T e^{-At}Bu_3(t)\,\mathrm{d}t = \int_0^T e^{-At}B[\lambda_1 u_1(t) + \lambda_2 u_2(t)]\,\mathrm{d}t$$
$$= \lambda_1 y_1(T) + \lambda_2 y_2(T)$$

so R is convex. We now state two fundamental convexity theorems to be used later.[3,9]

Convexity Theorem 1. Consider any two vectors z and ζ contained in the convex set M and any vector η contained in the linear space V generated by M. If

$$\eta^T z \geq \eta^T \zeta, \quad \forall \zeta \tag{5}$$

then 1) z lies on the boundary of M, and 2) the vector η is an outward normal to the boundary of M at z.

Proof: Rather than a rigorous analytical proof, we construct a geometric proof. A convex set M and some vector η contained in the linear space generated by M are shown in Fig.

Received July 11, 1990; revision received Oct. 19, 1990; accepted for publication Dec. 26, 1990.

*Graduate Research Assistant, Mars Mission Research Center. Student Member AIAA.

†Associate Professor, Mars Mission Research Center. Member AIAA.

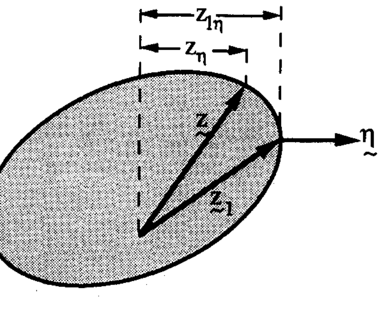

Fig. 1 Convexity theorem 1.

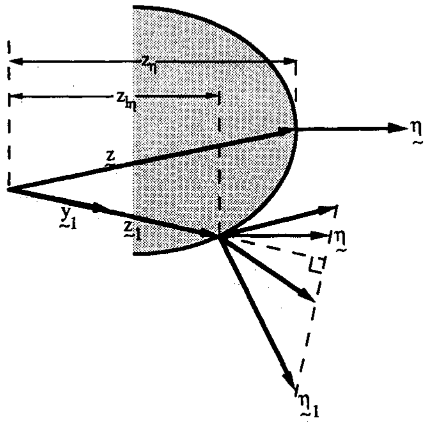

Fig. 2 Convexity theorem 2.

1; z clearly lies on the boundary of M. Otherwise z could be projected to the boundary to yield $\bar{\alpha}z$ for some $\bar{\alpha}>1$. Then, letting $\zeta = \bar{\alpha}z$ would yield the contradiction $\eta^T z < \eta^T \zeta$. Furthermore, η is the outward normal to M at z. To show this, let us assume η is not an outward normal to M at z. Moving η around the boundary, we locate the point on the boundary where η is an outward normal denoted by z_1. Since $\eta^T z = |\eta| z_\eta$, $\eta^T z_1 = |\eta| z_{1\eta}$, and $z_{1\eta} > z_\eta$, we have the contradiction $\eta^T z_1 > \eta^T z$.

End of Proof

Convexity Theorem 2. Consider any vector z contained on the boundary of the convex set M and the vector η contained in the linear space generated by M. The outward normal to the boundary of M located at z is denoted by $\eta(z)$, and we define the non-negative function $F(\eta) = \eta^T(z)z$. Let y_1 be contained in M and denote the extension of y_1 to the boundary of M by z_1, that is, $z_1 = \bar{\alpha} y_1$ in which $\bar{\alpha}$ denotes the largest β for which βy_1 is contained in M. If we constrain η to the hyperplane H expressed as

$$H = \{\eta : \eta^T y_1 = 1\}$$

then

$$F(\eta_1) = \bar{\alpha} \text{ and } F(\eta_1) \le F(\eta) \ \forall \eta \text{ in } H \tag{6}$$

where $\eta_1 = \eta(z_1)$ is the outward normal to M at z_1.

Proof: Figure 2 shows a portion of a convex set containing y_1. A few of the elements of H are shown at z_1. These include η_1, the outward normal to the boundary at z_1, and η, the outward normal to the boundary at z. First note that $F(\eta_1) = \eta_1^T z_1 = \eta_1^T \bar{\alpha} y_1 = \bar{\alpha}$. Next note that the projection of z in the $\eta(z)$ direction is greater than or equal to the projection of z_1 in the $\eta(z)$ direction, that is, $z_\eta \ge z_{1\eta}$. It follows that $\eta^T z \ge \eta^T z_1 = \eta^T \bar{\alpha} y_1 = \bar{\alpha}$, that is, $F(\eta) \ge F(\eta_1)$.

End of Proof

Multiplying Eq. (3) by any η in V yields

$$\eta^T y(T) = \int_0^T g^T(\eta,t) u(t) \, dt \tag{7}$$

in which

$$g(\eta,t) = (\eta^T e^{-At} B)^T \tag{8}$$

By convexity theorem 1, if a control maximizes Eq. (7) over U, then it produces a reachable state on the boundary of R. Upon maximizing Eq. (7), the controls that produce reachable states on the boundary of R are expressed as a function of η. The control that produces z_1 in convexity theorem 2 can then be expressed as $u_1 = u(\eta_1)$. From Eq. (3), the control that produces y_1 is $u_1/\bar{\alpha}$ since $\bar{\alpha} y_1 = z_1$. Indeed, any reachable state in the interior of R can as well be expressed as a function of η by its extension to the boundary of R.

Now define the cost function $C(u)$ that in turn is used to define the admissible set as follows:

$$U = \{u : C(u) \le C_{\max}\} \tag{9}$$

From Eq. (9), the admissible set bounds the cost by $C_{\max}$. Let us assume that the cost function satisfies the multiplicative property

$$C(\beta u) = |\beta| C(u) \tag{10}$$

for any β and for any control u. This property is recognized as the third property associated with the definition of a norm.[7] We now show that the control that produced y_1, given previously by

$$u^* = u_1/\bar{\alpha} \tag{11}$$

is the optimal control, that is, minimizes $C(u)$ over all u in U that produce y_1. Toward that end, we note from Eq. (11) that $C(u^*) = C(u_1)/\bar{\alpha} = C_{\max}/\bar{\alpha}$ since u_1 is on the boundary of U from which it follows that $C(u_1) = C_{\max}$. Let us now consider any other control u_2 in U that produces y_1 and show that $C(u_2) \ge C_{\max}/\bar{\alpha}$. To this end, we define $u_3 = u_2[C_{\max}/C(u_2)]$. Although it does not produce y_1, u_3 is admissible since $C(u_3) = C(u_2)(C_{\max}/C(u_2) = C_{\max}$. Thus, the state attained by u_3 is contained in the reachable set. From Eq. (3),

$$\int_0^T e^{-At} B u_3(t) \, dt = C_{\max}/C(u_2) \int_0^T e^{-At} B u_2(t) \, dt$$
$$= y_1 C_{\max}/C(u_2)$$

But $y_1 C_{\max}/C(u_2) \le \bar{\alpha} y_1$ since $\bar{\alpha} y_1$ is on the boundary of the reachable set. It follows that $C(u_2) \ge C_{\max}/\bar{\alpha}$, and so $C(u_2) \ge C(u^*)$.

III. Some Classical Cost Functions

The following describes three optimal control problems. In each problem, a control that produces a reachable state on the boundary is constructed and then expressed as a function of η. The construction of the optimal control follows directly.

Bang-Bang Control

Consider the case of minimizing the largest control input. This can be regarded as a saturation condition. The cost function is

$$C_1(u) = \max_{1 \le j \le m} \sup_{0 \le t \le T} |u_j(t)| \tag{12}$$

in which $u = [u_1, u_2, \ldots, u_m]^T$. From Eq. (9), the set of admissible controls is

$$U_1 = \{u(t) : \max_{1 \le j \le m} \sup_{0 \le t \le T} |u_j(t)| \le C_{\max}\} \tag{13}$$

which bounds each input by C_{max}. Without loss of generality, we let $C_{max} = 1$ since it turns out that the optimal control u^* is independent of C_{max}. Note that this further implies that the optimal control u^* exists for large enough C_{max} and for large enough T. From Eq. (7),

$$\eta^T y(T) = \int_0^T g^T(\eta,t)u(t)\,dt = \int_0^T \sum_{j=1}^m g_j(\eta,t)u_j(t)\,dt \quad (14)$$

where $g = [g_1, g_2, \ldots, g_m]^T$. Equation (14) is largest when the control is of the form $u_j = \text{sgn}(g_j)$, $j = 1, 2, \ldots, m$, that is, when u is expressed as a function η as

$$u = \text{sgn}(g) \quad (15)$$

By convexity theorem 1, the control in Eq. (15) produces reachable states on the boundary of the reachable set R_1 associated with Eq. (12). Substituting Eq. (15) into Eq. (14) yields the function

$$F(\eta) = \int_0^T \sum_{j=1}^m |g_j|\,dt \quad (16)$$

By convexity theorem 2,

$$\bar{\alpha} = \min_{\eta \in H} F(\eta) = F(\eta_1) \quad (17)$$

The optimal control is given by Eq. (11), in which $\bar{\alpha}$ and η_1 are determined from Eq. (17).

Continuous Control

Consider the cost function representing a pseudofuel

$$C_2(u) = \left(\int_0^T \sum_{j=1}^m |u_j(t)|^p\,dt\right)^{1/p} \qquad p > 1 \quad (18)$$

From Eq. (9), the admissible set of controls is

$$U_2 = \left\{u(t): \left(\int_0^T \sum_{j=1}^m |u_j(t)|^p\,dt\right)^{1/p} \leq 1\right\} \quad (19)$$

where we let $C_{max} = 1$ without loss of generality. Using Holder's inequalities for sums and integrals in conjunction with Eq. (7),[10]

$$\begin{aligned}\eta^T y(T) &= \int_0^T \sum_{j=1}^m g_j u_j\,dt \leq \int_0^T \sum_{j=1}^m |g_j u_j|\,dt \\ &\leq \int_0^T \left(\sum_{j=1}^m |g_j|^q\right)^{1/q} \left(\sum_{j=1}^m |u_j|^p\right)^{1/p} dt \\ &\leq \left(\int_0^T \sum_{j=1}^m |g_j|^q\,dt\right)^{1/q} \left(\int_0^T \sum_{j=1}^m |u_j|^p\,dt\right)^{1/p} \\ &\leq \left(\int_0^T \sum_{j=1}^m |g_j|^q\,dt\right)^{1/q} \end{aligned} \quad (20)$$

where $(1/p) + (1/q) = 1$. We now verify that u expressed as a function of η in the form

$$u_j = k|g_j|^{q/p}\,\text{sgn}(g_j), \qquad j = 1,2,\ldots,m \quad (21)$$

maximizes Eq. (20). Substituting Eq. (21) into the left side of Eq. (20) yields the right side of Eq. (20)

$$\begin{aligned}\eta^T y(T) &= \int_0^T \sum_{j=1}^m g_j k|g_j|^{q/p}\,\text{sgn}(g_j)\,dt \\ &= k\int_0^T \sum_{j=1}^m |g_j|^q\,dt \leq \left(\int_0^T \sum_{j=1}^m |g_j|^q\,dt\right)^{1/q}\end{aligned} \quad (22)$$

by letting

$$k = \left(\int_0^T \sum_{j=1}^m |g_j|^q\,dt\right)^{-1/p} \quad (23)$$

Thus, substituting Eqs. (21) and (23) into the left side of Eq. (20) we obtain by convexity theorem 2,

$$F(\eta) = \left(\int_0^T \sum_{j=1}^m |g_j|^q\,dt\right)^{1/q}, \qquad \bar{\alpha} = \min_{\eta \in H} F(\eta) = F(\eta_1) \quad (24)$$

The optimal control is again given by Eq. (11) in which $\bar{\alpha}$ and η_1 are determined from Eq. (24). From Eqs. (23) and (24), note that $k = F(\eta)^{-q/p}$.

Impulse Control

The final cost function considered is a measure of absolute fuel,

$$C_3(u) = \int_0^T \sum_{j=1}^m |u_j(t)|\,dt \quad (25)$$

From Eq. (9) the set of admissible controls is

$$U_{3a} = \left\{u(t): \int_0^T \sum_{j=1}^m |u_j(t)|\,dt \leq 1\right\} \quad (26)$$

where $C_{max} = 1$. Another possibility that we will consider separately for this cost function is the set of admissible controls defined by placing a bound on each input

$$U_{3b} = \left\{u(t): \int_0^T |u_j(t)|\,dt \leq 1, \qquad j = 1, 2, \ldots, m\right\} \quad (27)$$

For now, return to the admissible set of Eq. (26). From Eq. (7)

$$\begin{aligned}\eta^T y(T) &= \int_0^T \sum_{j=1}^m g_j u_j\,dt \leq \int_0^T \sum_{j=1}^m |g_j||u_j|\,dt \\ &\leq \max_{1 \leq j \leq m} \sup_{0 \leq t \leq T} |g_j|\end{aligned} \quad (28)$$

By inspecting Eq. (28), we conclude that impulse functions applied at the points where g_j is largest in turn maximizes Eq. (28). The supremum in Eq. (28) may occur at multiple j and at multiple instances of time τ_{ij} $i = 1, 2, \ldots, N_j$, where N_j equals zero if g_j does not contain the supremum. Thus, there may be an infinite number of impulsive control strategies that maximize Eq. (28). The controls that produce reachable states on the boundary of R_{3a} are of the form

$$u_j(t) = \sum_{i=1}^{N_j} c_{ij}\,\text{sgn}[g_j(\eta,\tau_{ji})]\delta(t - \tau_{ji}), \qquad 1 = \sum_{j=1}^m \sum_{i=1}^{N_j} c_{ij} \quad (29)$$

where c_{ij} are positive coefficients. Substituting Eq. (29) into the left side of Eq. (28) yields

$$F(\eta) = \max_{1 \leq j \leq m} \sup_{0 \leq t \leq T} |g_j|, \qquad \bar{\alpha} = \min_{\eta \in H} F(\eta) = F(\eta_1) \quad (30)$$

The optimal control is given by Eq. (11) in which η_1 and $\bar{\alpha}$ are determined from Eq. (30). Note that $u(\eta_1)$ produces a family of reachable states on the boundary of R depending on c_{ij}. The coefficients c_{ij} are selected so as to transfer the system from x_0 to x_1 in time T. Furthermore, note that multiple selections of c_{ij} can exist. We now turn to the set of admissible controls U_{3b}. From Eqs. (7) and (27)

$$\eta^T y(T) = \int_0^T \sum_{j=1}^m g_j u_j\,dt \leq \sum_{j=1}^m \sup_{0 \leq t \leq T} |g_j| \quad (31)$$

The supremum in Eq. (31) may occur at multiple instances of time τ_{ji}, $i = 1, 2, \ldots, M_j$, where $M_j \geq 1$. Again there may be a multitude of impulsive control strategies that maximize Eq.

(31). The control that produces reachable states on the boundary of R_{3b} is given by

$$u_j(t) = \sum_{i=1}^{M_j} c_{ij} \operatorname{sgn}[g_j(\tau_{ji},\eta)]\delta(t-\tau_{ji}), \qquad 1 = \sum_{i=1}^{M_j} c_{ij} \tag{32}$$

As stated previously, although any control of the form of Eq. (32) will map the system to the boundary of the reachable set, only certain choices of $c_{ij}(i = 1, 2, \ldots, M_j,\ j = 1, 2, \ldots, m)$ yield $\eta_2 = \eta_1$. Thus, care must be taken to choose a set of c_{ij} that results in a controller that drives the system from x_0 to x_1. We obtain by convexity theorem 2,

$$F(\eta) = \sum_{j=1}^{m} \sup_{0\le t\le T} |g_j|, \qquad \bar{\alpha} = \min_{\eta\in H} F(\eta) = F(\eta_1) \tag{33}$$

The optimal control is given by Eq. (11) in which η_1 and $\bar{\alpha}$ are determined by Eq. (33).

IV. Control of a Damped Harmonic Oscillator

Differences among the solutions given in the previous section are brought out by considering the damped harmonic oscillator. The motion of the oscillator is described by the differential equation

$$m\ddot{x}(t) + c\dot{x}(t) + kx(t) = u(t) \tag{34}$$

where m denotes mass, c denotes damping, and k denotes stiffness. Equation (34) is recast in the form of Eq. (1) in which

$$\frac{c}{m} = 2\alpha, \qquad \frac{k}{m} = \omega^2, \qquad \boldsymbol{x}(t) = \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix}$$

leading to the coefficient matrices

$$A = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -2\alpha \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 1/m \end{bmatrix}$$

This system is controllable.[8] The state transition matrix of the damped harmonic oscillator and the inverse at time T are

$$e^{At} = e^{-\alpha t}\begin{bmatrix} \cos\beta t + \dfrac{\alpha}{\beta}\sin\beta t & \dfrac{1}{\beta}\sin\beta t \\ -\left(\dfrac{\alpha^2+\beta^2}{\beta}\right)\sin\beta t & \cos\beta t - \dfrac{\alpha}{\beta}\sin\beta t \end{bmatrix}$$

$$e^{-AT} = e^{\alpha T}\begin{bmatrix} \cos\beta T - \dfrac{\alpha}{\beta}\sin\beta T & -\dfrac{1}{\beta}\sin\beta T \\ \left(\dfrac{\alpha^2+\beta^2}{\beta}\right)\sin\beta T & \cos\beta T + \dfrac{\alpha}{\beta}\sin\beta T \end{bmatrix} \tag{35}$$

where $\beta^2 = \alpha^2 - \omega^2$.

It is desired to drive the system from some initial displacement, $\boldsymbol{x}(0) = [x_0\ 0]^T$, to the origin, $\boldsymbol{x}(T) = [0\ 0]^T$, in some time T. From Eq. (3), we compute the desired reachable state

$$y(T) = e^{-AT}\boldsymbol{x}(T) - \boldsymbol{x}(0) = \begin{bmatrix} -x_0 \\ 0 \end{bmatrix} \tag{36}$$

Since $n = 2$, we write $\eta = [\eta_1\ \eta_2]^T$. From Eq. (8), the function $g(\eta,t)$ is reduced to

$$g(\eta,t) = De^{\alpha t}\sin(\beta t + \Phi) \tag{37}$$

where

$$D = [(-\eta_1/\beta + \eta_2\alpha/\beta)^2 + \eta_2^2]^{1/2}/m$$

$$\Phi = \tan^{-1}\left[\frac{\eta_2\beta}{-\eta_1 + \eta_2\alpha}\right] \tag{38}$$

For further reference, notice that $g(\eta,t)$ is an exponentially growing oscillatory function of time. Consistent with convexity theorem 2, we define the hyperplane

$$H = \left\{\eta = [\eta_1\ \eta_2]^T : \eta_1 = -\frac{1}{x_0}\right\} \tag{39}$$

Bang-Bang Control

Consider the cost function and corresponding boundary control given in Eqs. (11) and (15), respectively. Substituting $g(\eta,t)$ from Eq. (38) into Eq. (16) yields

$$F(\eta) = D\int_0^T e^{\alpha t}|\sin(\beta t + \Phi)|\, dt \tag{40}$$

for $\alpha \ll \beta$, the minimum value of Eq. (40) over the hyperplane occurs when η_2 is zero. Thus, from Eq. (40)

$$\bar{\alpha} = \frac{1}{x_0\beta m}\int_0^T e^{\alpha t}|\sin\beta t|\, dt \tag{41}$$

From Eq. (11), the optimal control is of the form

$$u^*(t) = \frac{1}{\bar{\alpha}}\operatorname{sgn}(\sin\beta t) \tag{42}$$

For example, let $m = k = 1$, $x_0 = 1$, and let the decay rate $\alpha = 0.03$. The selection of the final time T depends on the admissible set of controls. Let $T = 6\pi/\beta$, representing the period of time of three oscillations, and assume $C_{\max}$ is large enough that the optimal control is within U. Figure 3a shows the control input, revealing the bang-bang nature of the control. This is also characteristic of the minimum time solution for the same problem with bounded control inputs. Indeed, if the final time T is chosen so as to place the bang-bang control on the boundary of the admissible set, that is, $C(u) = C_{\max}$, then the minimum time solution is obtained.[2] The phase diagram shown in Fig. 3b confirms that the system is driven to the origin in three oscillations.

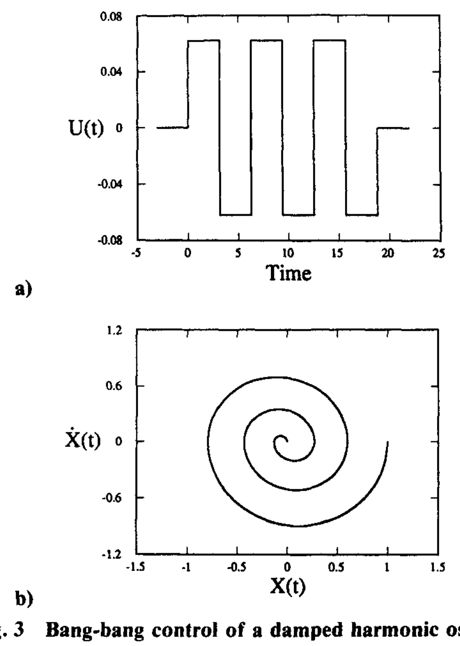

Fig. 3 Bang-bang control of a damped harmonic oscillator.

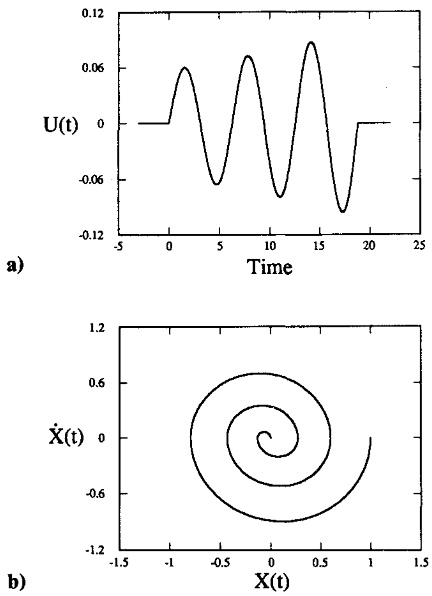

Fig. 4 Continuous control of a damped harmonic oscillator.

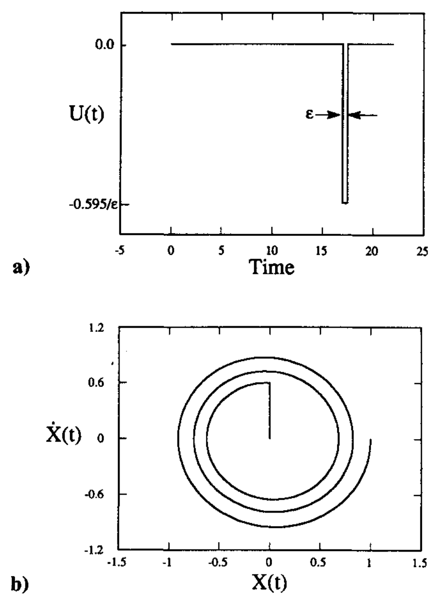

Fig. 5 Impulse control of a damped harmonic oscillator.

Continuous Control

Now consider the cost function in Eq. (18). Letting $p = q = 2$, the cost function takes on a more popular quadratic form. Substituting Eq. (37) into Eq. (24) yields

$$F(\eta) = \left[D^2 \int_0^T e^{2\alpha t} \sin(\beta t + \Phi)\, dt \right]^{1/2} \tag{43}$$

As in the previous example, for $\alpha \ll \beta$ the minimum value of $F(\eta)$ occurs when η_2 is equal to zero. From Eqs. (23) and (24) the optimal control is

$$u^*(t) = \frac{1}{\bar{\alpha}^2} \frac{1}{x_0 \beta m} e^{\alpha t} \sin(\beta t) \tag{44}$$

where

$$\bar{\alpha}^2 = \left[\frac{1}{\beta m x_0} \right]^2 \int_0^T e^{2\alpha t} \sin^2 \beta t\, dt \tag{45}$$

As in the previous example, we let $m = k = 1$ and $x_0 = 1$. Figure 4a shows that the control input is an exponentially growing oscillatory function. The control drives the system to the origin as shown in Fig. 4b.

Impulse Control

The cost function of Eq. (25) is now considered. Since the dimension of the control vector $m = 1$, the two sets of admissible controls U_{3a} and U_{3b} are identical. A unique maximum is determined to occur during the last half-period of the last oscillation by examining $g(\eta,t)$ in Eq. (37). As in the previous examples, we let $m = k = 1$, $x_0 = 1$, and $\alpha = 0.03$. Since $\alpha \ll \beta$, the maximum occurs at approximately $\tau = [(11\pi/2) - \Phi]/\beta$. Substituting $g(\eta,t)$ into Eq. (30) or (33) yields

$$F(\eta) = De^{\alpha\tau} |\sin(\beta\tau + \Phi)| \tag{46}$$

where $F(\eta)$ is minimized over H by letting $\eta_2 = 0$. We obtain

$$\bar{\alpha} = \frac{1}{\beta m x_0} e^{\alpha\tau} \tag{47}$$

in which $\tau = 11\pi/2\beta$. The final form of the optimal control is

$$u^*(t) = \beta m x_0 e^{-\alpha\tau} \operatorname{sgn}(\sin\beta\tau)\delta(t - \tau) \tag{48}$$

Impulse control is shown in Fig. 5a. The impulse control allows for a period of free oscillation, during which time the natural damping present in the system removes energy from the system. Then at the last instant the system's displacement is identically zero, an impulse of magnitude equal to the system's momentum is applied. The instantaneous change in velocity abruptly transfers the system to the origin at time τ. The timing of the impulse is critical. Figure 5b shows that the impulse is applied when the system potential energy is at a minimum and when the system kinetic energy is at a local maximum. This observation was useful in the development of near optimal control laws for large scale systems.[6]

V. Rigid-Body Maneuver

We now consider rest-to-rest rigid-body maneuvers. The equation of motion is

$$m\ddot{x}(t) = u(t) \tag{49}$$

Equation (49) is transformed into the state space by letting

$$\boldsymbol{x}(t) = \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1/m \end{bmatrix}$$

Once again, it is a simple matter to verify the controllability of this system. The state transition matrix and its inverse at time T become

$$e^{At} = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}, \quad e^{-At} = \begin{bmatrix} 1 & -T \\ 0 & 1 \end{bmatrix} \tag{50}$$

It is desired to transfer the system from the origin $\boldsymbol{x}(0) = [0\ 0]^T$ to some final point $\boldsymbol{x}(T) = [x_T\ 0]^T$. From Eq. (3), the reachable state is

$$y(T) = e^{-AT}\boldsymbol{x}(T) - \boldsymbol{x}(0) = \begin{bmatrix} x_T \\ 0 \end{bmatrix} \tag{51}$$

Since $n = 2$, $\eta = [\eta_1\ \eta_2]^T$. From Eq. (51) the hyperplane is

$$H = \left\{ \eta = [\eta_1\ \eta_2]^T : \eta_1 = \frac{1}{x_T} \right\} \tag{52}$$

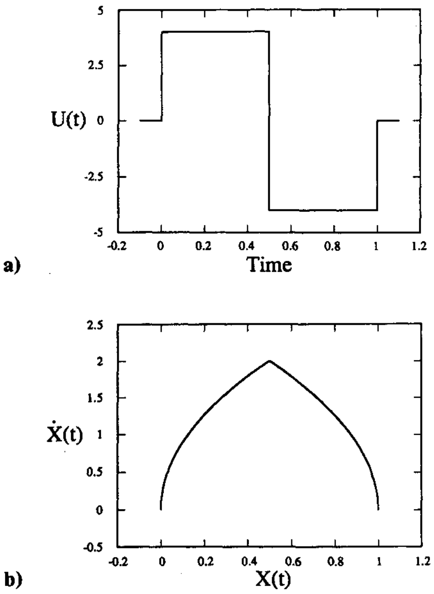

Fig. 6 Bang-bang control of a rigid body.

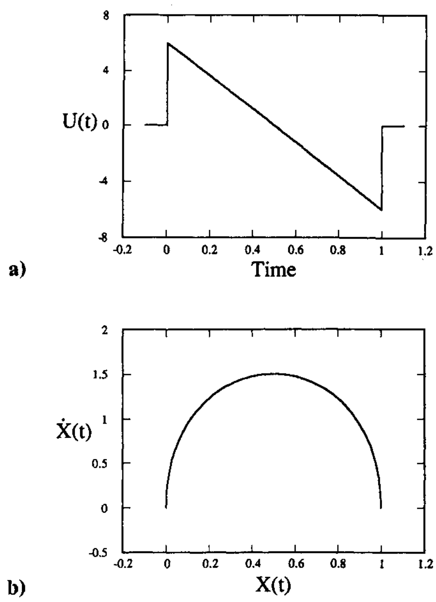

Fig. 7 Continuous control of a rigid body.

From Eq. (8)

$$g(\eta,t) = -\frac{\eta_1 t}{m} + \frac{\eta_2}{m} = \frac{-t}{mx_T} + \frac{\eta_2}{m} \tag{53}$$

Bang-Bang Control

Considering the cost function of Eq. (12), we obtain from Eqs. (53) and (16)

$$F(\eta) = \int_0^T \left| \frac{-t}{mx_T} + \frac{\eta_2}{m} \right| \mathrm{d}t \tag{54}$$

The minimization of Eq. (54) over the hyperplane is accomplished by distinguishing between three cases depending on the sign of the integrand (positive, negative, and sign change). We obtain $\eta_2 = T/2x_T$ and

$$\bar{\alpha} = \frac{T^2}{4mx_T} \tag{55}$$

The optimal control is

$$u^*(t) = \frac{4mx_T}{T^2}\mathrm{sgn}\left(\frac{T}{2} - t\right) \tag{56}$$

For example, let $m = 1$, $T = 1$, and $x_T = 1$. The associated control and the system response are shown in Fig. 6. Note that choosing T to place the control on the boundary of U_1 corresponds to the minimum time solution, as was the case with the damped oscillator.

Continuous Control

Consider the cost function of Eq. (18) for $p = q = 2$. Repeating the previous steps we obtain

$$F(\eta) = \frac{1}{m}\left[\frac{T^3}{3x_T^2} - \frac{\eta_2 T^2}{x_T} + \eta_2^2 T\right]^{1/2} \tag{57}$$

The minimum of Eq. (57) occurs when $\eta_2 = T/2x_T$, yielding

$$\bar{\alpha} = \frac{1}{m}\left[\frac{T^3}{12x_T^2}\right]^{1/2} \tag{58}$$

The optimal control becomes

$$u^*(t) = \frac{12x_T m}{T^3}\left[\frac{T}{2} - t\right] \tag{59}$$

Figure 7 shows the optimal control forces and the system response. The linearity of the control yields a smooth transition of the system from the initial state to the final state.

Impulse Control

Consider the cost function given in Eq. (25). From Eq. (30)

$$F(\eta) = \sup_{0 \le t \le T} \left| -\frac{t}{x_T m} + \frac{\eta_2}{m} \right| \tag{60}$$

From Eq. (60) and convexity theorem 2

$$\bar{\alpha} = \min_{\eta_2} \sup_{0 \le t \le T} \left| -\frac{t}{x_T m} + \frac{\eta_2}{m} \right| \tag{61}$$

Considering three possible ranges of the argument (positive, negative, sign change), the minimum of Eq. (60) occurs when $\eta_2 = T/2x_T$, yielding

$$\bar{\alpha} = \frac{T}{2x_T m} \tag{62}$$

This minimum occurs both at $\tau_1 = 0$ and $\tau_2 = T$. From Eq. (29) or (32) with N or $M = 2$, it is determined that a uniform distribution of the impulsive forces will drive the system from x_0 to x_1. Thus, letting $c_1 = c_2 = 1/2$ in Eq. (32), the optimal control is

$$u^*(t) = \frac{x_T m}{T}\mathrm{sgn}\left(\frac{T}{2} - t\right)[\delta(t) + \delta(t - T)] \tag{63}$$

This solution consists of one initial impulse to impart a velocity to the system as shown in Fig. 8. Then the system drifts to the desired position at time T, at which time a final impulse terminates the motion, leaving the system at rest.

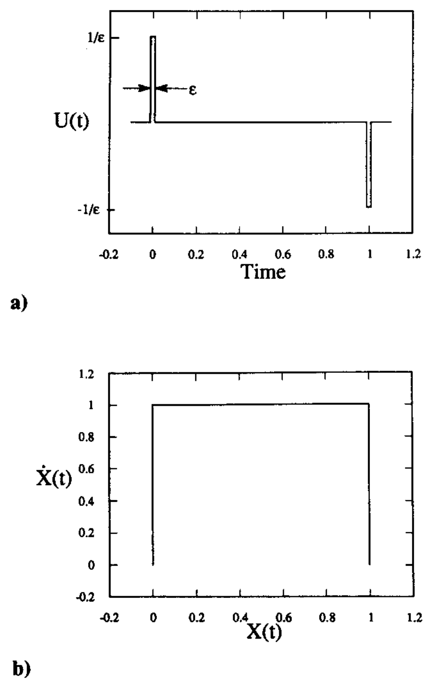

Fig. 8 Impulse control of a rigid body.

Table 1 Fuel consumption

	Bang-bang	Continuous	Impulse
Oscillator	1.168	0.921	0.595
Rigid body	4.0	3.0	2.0

VI. Fuel Consumption

The previous two sections have significant implications regarding the use of continuous control in propulsive systems. The fuel consumption in propulsive systems is[11]

$$Fu = \frac{1}{I_{sp}} \int_0^T |u(t)| \, dt \tag{64}$$

where $u(t)$ represents thrust and I_{sp} represents specific impulse. Letting $I_{sp} = 1$, the fuel consumed in the previously considered cases is given in Table 1.

Table 1 indicates that impulse control offers a significant reduction in fuel for both the damped oscillator and the rigid body. Of course, this result is expected since the third cost function truly represents absolute fuel. Compared with the continuous control, the impulse control offers a 35% savings of fuel for the damped oscillator and 33% for the rigid-body maneuver. The natural question arises whether a savings of this order of magnitude can be expected in larger order systems as well. A partial answer to this question can be found in Ref. 6, where a simple impulsive control approximation posted a savings on the order of 20% over continuous control for vibration suppression of a cantilevered beam.

VII. Final Comments

This paper developed a means for comparing three optimal control strategies on the basis of fuel consumption. First, a general cost function minimization procedure was developed via the application of two theorems associated with convex sets. This procedure is applicable to optimal control problems employing a cost function that satisfies a norm property. Three cost functions associated with control saturation, pseudofuel, and absolute fuel were introduced and minimized in accordance with the previously mentioned technique. The first two cost functions led to the previously well-documented bang-bang and continuous control strategies. The minimization of absolute fuel led to an impulsive strategy that was first suggested by L. W. Neustadt but not developed. The three control strategies were then implemented on two elementary systems and a comparison of fuel consumption was made. The impulsive control strategy consumed significantly less fuel than the continuous and bang-bang control strategies. This comparison suggests a potential for fuel savings in higher order systems through the use of impulsive control strategies. However, exact solutions to fuel optimal control for large-order systems are difficult if not impossible to achieve. The alternative is to develop near-optimal control strategies. This can be accomplished in one of two ways. The first way is to approximate the mathematical conditions of optimality and to construct approximating manifolds used in feedback control algorithms. The second option is to approximate qualitative dynamic characteristics exhibited by the exact solution to the problem. The second approach is in certain cases more amenable to the development of feedback control algorithms.

Acknowledgments

This research is supported in part by NASA Grant NAGW-1331 to the Mars Mission Research Center. Critical discussions with E. N. Chukwu of North Carolina State University provided the basis for the developments contained in Secs. II and III.

References

[1]Krasovskii, N. N., "On the Theory of Optimal Control," *Journal of Applied Mathematics and Mechanics,* Vol. 23, No. 4, 1959, pp. 899–919.

[2]Neustadt, L. W., "Synthesizing Time Optimal Control Systems," *Journal of Mathematical Analysis and Applications,* Vol. 1, 1960, pp. 484–493.

[3]Neustadt, L. W., "Minimum Effort Control Systems," *SIAM Journal on Control,* Vol. 1, No. 1, 1962, pp. 16–31.

[4]Hajek, O., "L₁-Optimization in Linear Systems with Bounded Controls," *Journal of Optimization Theory and Applications,* Vol. 29, No. 3, Nov. 1979, pp. 409–436.

[5]Chukwu, E. N., *Control of Hereditary Systems—Class Notes,* Dept. of Mathematics, North Carolina State Univ., Raleigh, NC, 1990.

[6]Foster, L. A., III, and Silverberg, L., "On-Off Decentralized Control of Flexible Structures," *Journal of Dynamic Systems, Measurement and Control,* Vol. 113, No. 1, 1991, pp. 41–47.

[7]Kreyszig, E., *Introductory Functional Analysis With Applications,* Wiley, New York, 1978, p. 59.

[8]Luenberger, D. G., *Introduction to Dynamic Systems—Theory, Models & Applications,* Wiley, New York, 1979.

[9]Hermes, H., and Lasalle, J. P., *Functional Analysis and Time Optimal Control,* Academic, New York, 1969, pp. 32–39.

[10]Taylor, A. E., and Mann, W. R., *Advanced Calculus,* 3rd ed., Wiley, New York, 1983, p. 188.

[11]Foster, L. A., III, *Reboost Attitude Control/Structure Interaction for MB-2,* Space Station Engineering and Integration Contractor, PSH-341-RP90-008, Reston, VA, March 1990.