Optimal Startup Control of a Jacketed Tubular Reactor

D. R. HAHN, L. T. FAN, and C. L. HWANG

Kansas State University, Manhattan, Kansas 66502

The optimal startup policy of a jacketed tubular reactor, in which a first-order, reversible, exothermic reaction takes place, is presented. A distributed maximum principle is presented for determining weak necessary conditions for optimality of a diffusional distributed parameter system. A numerical technique is developed for practical implementation of the distributed maximum principle. This involves the sequential solution of the state and adjoint equations, in conjunction with a functional gradient technique for iteratively improving the control function.

This paper presents an optimal policy for startup of a jacketed tubular reactor in which a first-order, reversible, exothermic reaction is taking place. The optimal control policy is determined by using a distributed maximum principle. The control or decision variable is the wall temperature of the reactor, which is manipulated to minimize a given performance index. Computational results are obtained for a case with and without a constraint on the maximum reaction temperature.

The mathematical model for the jacketed tubular reactor is a continuous distributed parameter flow system, which gives rise to a set of coupled nonlinear, one-dimensional, second-order, parabolic partial differential equations. A distributed maximum principle used by previous workers, for example Denn et al (*1*), is extended to a general system of nonlinear diffusion equations, with two-point boundary conditions consisting of linear relationships between the dependent variables and their axial gradients. A set of necessary conditions for optimality is obtained for a fairly general performance index.

In general, equations of the type treated cannot be solved by analytic methods and even numerical techniques for coupled, highly nonlinear, axial diffusion equations are not generally available. Therefore an iterative computational technique involving a gradient in functional space is presented, which enables the numerical implementation of the distributed maximum principle. It is shown that the technique is capable of accommodating inequality constraints on state variables by the addition of an appropriate penalty function to the performance index.

A DISTRIBUTED MAXIMUM PRINCIPLE

A distributed maximum principle is presented for determining weak necessary conditions for optimality for a class of distributed systems. Due to the complexity of partial differential equations, a completely general maximum principle, as exists for lumped-parameter systems (*2, 3*), has not been found. However, sufficient generality has been retained that the results apply to a wide variety of systems of interest in process control.

D. R. Hahn is at the Bendix Corporation, Kansas City, Missouri.

System Description

Attention will be focused on systems which may be described by a general nonlinear vector partial differential equation of the form

$$\mathbf{u}_t(x,t) = \mathbf{f}\{\mathbf{u}(x,t),\ \mathbf{u}_x(x,t),\ \mathbf{u}_{xx}(x,t),\ \boldsymbol{\theta}(t),\ x,\ t\} \tag{1}$$

where $\mathbf{u}$ is an s-dimensional state vector defined on a normalized one-dimensional spatial domain x from $x = 0$ to $x = 1$ and over a fixed time interval $t = 0$ to $t = t_f$. The control vector $\boldsymbol{\theta}$ is considered to be a function of time only and is r-dimensional. An independent variable appearing as a subscript denotes partial differentiation with respect to that variable.

Equation (1) is augmented by the following set of initial and boundary conditions:

$$\xi_i(\mathbf{u}) = 0 \qquad \text{at } t = 0,\ i = 1, \ldots, s \tag{2}$$

$$\phi_m(\mathbf{u}, \mathbf{u}_x) = 0 \text{ at } x = 0,\ m = 1, \ldots, p \tag{3}$$

$$\psi_n(\mathbf{u}, \mathbf{u}_x) = 0 \text{ at } x = 1,\ n = 1, \ldots, q = 2s - p \tag{4}$$

It is desired to determine the control function $\boldsymbol{\theta}$ which yields a minimum for the following generalized objective functional:

$$S = \int_0^1 F\{\mathbf{u}(x,t_f),\ x\}\,dx + \int_0^{t_f}\int_0^1 G\{\mathbf{u}(x,t),\ \boldsymbol{\theta}(t)\ x, t\}\,dx\,dt \tag{5}$$

In order to obtain necessary conditions for optimality, it is required to find a relationship which expresses variations of the objective functional δS in terms of control perturbations $\delta\boldsymbol{\theta}$.

Adjoining the variational system equation of Equation (1) as an equality constraint with the variational objective functional of Equation (5) yields

$$\delta S = \int_0^1 F_{\mathbf{u}}{}^T\,\delta\mathbf{u}\,dx\big|_{t=t_f} + \int_0^{t_f}\int_0^1 \{G_{\mathbf{u}}{}^T\,\delta\mathbf{u} + G_{\boldsymbol{\theta}}{}^T\,\delta\boldsymbol{\theta} - \mathbf{z}^T[\delta\mathbf{u}_t - \mathbf{f}_{\mathbf{u}}\,\delta\mathbf{u} - \mathbf{f}_{\mathbf{u}_x}\,\delta\mathbf{u}_x - \mathbf{f}_{\mathbf{u}_{xx}}\,\delta\mathbf{u}_{xx} - \mathbf{f}_{\boldsymbol{\theta}}\,\delta\boldsymbol{\theta}]\}\,dx\,dt \tag{6}$$

where $\mathbf{z}(x,t)$ is an s-dimensional adjoint vector. After some manipulation, the following result is obtained:

$$\delta S = \int_0^1 [F_u{}^T - z^T]\, \delta u\, dx|_{t=t_f} + \int_0^{t_f} \int_0^1 \{[G_\theta{}^T + z^T f_\theta]\, \delta\theta + [z_t{}^T + z^T f_u - (z^T f_{u_x})_x + (z^T f_{u_{xx}})_{xx} + G_u{}^T]\, \delta u\}\, dx\, dt + \int_0^{t_f} [[z^T f_{u_x} - (z^T f_{u_{xx}})_x]\, \delta u + z^T f_{u_{xx}}\, \delta u_x]^1{}_{x=0}\, dt \quad (7)$$

In order to eliminate terms not depending explicitly on $\delta\theta$ from the second integrand of Equation (7), it is stipulated that each component of the adjoint vector z satisfy the following partial differential equation:

$$z_{i_t} = -z^T f_{u_i} + (z^T f_{u_{i_x}})_x - (z^T f_{u_{i_{xx}}})_{xx} - G_{u_i}, \quad i = 1, \ldots, s \quad (8)$$

The boundary conditions for Equation (8) are specified such that

$$[[z^T f_{u_x} - (z^T f_{u_{xx}})_x]\, \delta u + z^T f_{u_{xx}}\, \delta u_x]^1{}_{x=0} = 0 \quad (9)$$

This is accomplished by choosing the adjoint boundary conditions such that the coefficients of the unknown endpoint variations δu and δu_x vanish.

At this point, the system boundary conditions, Equations (3) and (4), are assumed to be linear and have the more explicit form

$$\phi_i(u, u_x) = 0 = u_{i_x} + a_i u_i + b_i \text{ at } x = 0,\ i = 1, \ldots, s \quad (10)$$

$$\psi_i(u, u_x) = 0 = u_{i_x} + c_i u_i + d_i \text{ at } x = 1,\ i = 1, \ldots, s \quad (11)$$

To set the coefficients of the δu_i equal to zero in Equation (7), the adjoint variables z_i are required to satisfy the following $2s$ conditions at $x = 0$ and $x = 1$:

$$z^T f_{u_{i_x}} - (z^T f_{u_{i_{xx}}})_x - a_i z^T f_{u_{i_{xx}}} = 0 \quad \text{at } x = 0,\ i = 1, \ldots, s \quad (12)$$

$$z^T f_{u_{i_x}} - (z^T f_{u_{i_{xx}}})_x - c_i z^T f_{u_{i_{xx}}} = 0 \quad \text{at } x = 1,\ i = 1, \ldots, s \quad (13)$$

The specification of the adjoint vector is completed by stipulating the transversality conditions

$$z_i = F_{u_i} \text{ at } t = t_f,\ i = 1, \ldots, s \quad (14)$$

This causes the first integral in Equation (7) to vanish, and it follows that δS can now be expressed explicitly in terms of the control perturbation $\delta\theta$:

$$\delta S = \int_0^{t_f} \int_0^1 [G_\theta{}^T + z^T f_\theta]\, \delta\theta\, dx\, dt \quad (15)$$

It is convenient to define a Hamiltonian function

$$H(u, u_x, u_{xx}, \theta, z, x, t) = G(u, \theta, x, t) + z(x, t)^T f(u, u_x, u_{xx}, \theta, x, t) \quad (16)$$

so that Equation (15) becomes

$$\delta S = \int_0^{t_f} \int_0^1 H_\theta{}^T\, \delta\theta\, dx\, dt \quad (17)$$

It follows by reasoning similar to that of Katz (*4*), who achieved a similar result for a more general and abstract class of problems, that the best choice of control action θ which minimizes the objective functional S is that which makes

$$\int_0^1 H\, dx \quad (18)$$

stationary over the interval $0 \leq t \leq t_f$ with respect to components of θ lying interior to the admissible control region and a minimum for those lying on the boundary. For control components interior to the region, this means

$$\int_0^1 H_\theta\, dx = 0 \quad (19)$$

Essentially this is an infinite dimensional or functional equivalent of Pontryagin's maximum principle (*2, 3*) for finite dimensional (lumped) systems.

OPTIMAL STARTUP CONTROL OF A JACKETED TUBULAR REACTOR CONSTRAINT ON MAXIMUM REACTION TEMPERATURE

The distributed maximum principle discussed in the preceding section will be applied to control of a tubular, continous flow chemical reactor in which an exotherm reaction is taking place. It is assumed that the reaction is first order and reversible ($A \rightleftarrows B$). Since the reaction rate is temperature dependent, it follows that the yield can be controlled by varying the reaction temperature. In this example, the reaction temperature and thus yield are controlled by manipulation of the reactor wall temperature.

The mathematical model for the system is based upon the following assumptions:

1. System parameters are uniform and constant with respect to time.
2. Wall temperature is a function of time only.
3. Axial heat and mass dispersion and mixing are significant inside the reactor.
4. Concentration, temperature, and velocity of the stream are constant with respect to radial distance.

A mass balance taken over a differential element along the reactor yields for component A:

$$\frac{\partial c_A}{\partial \tau} = D_m \frac{\partial^2 c_A}{\partial l^2} - v \frac{\partial c_A}{\partial l} + R_A \quad (20)$$

where, for the case of a first-order, reversible $A \rightleftarrows B$ reaction the rate of production of A, R_A, is given by the Arrhenius expression

$$R_A = -[k_1 c_A - k_2 c_B] = -[k_{10} \exp(-E_1/RT) c_A - k_{20} \exp(-E_2/RT) c_B]$$

A heat balance on the differential section of the reactor yields

$$\frac{\partial T}{\partial \tau} = \frac{k_{\text{eff}}}{C_p \rho} \frac{\partial^2 T}{\partial l^2} - v \frac{\partial T}{\partial l} + \frac{(-\Delta H)}{C_p \rho} R_A - \frac{2h}{C_p \rho r} (T - T_W) \quad (21)$$

It is assumed that the manner of mixing is such that the effective mass and thermal diffusivities are equal, that is

$$D_m = \frac{k_{\text{eff}}}{C_p \rho} \equiv D$$

The boundary conditions for this problem are those first suggested by Danckwerts (*9*):

$$\frac{\partial c_A(0, \tau)}{\partial l} = \frac{V}{D} [c_A(0, \tau) - c_A{}^f] \text{ at } l = 0 \quad (22)$$

$$\frac{\partial c_A(L,\tau)}{\partial l} = 0 \qquad \text{at} \quad l = L \qquad (23)$$

$$\frac{\partial T(0,\tau)}{\partial l} = \frac{V}{D}[T(0,\tau) - T^f] \qquad \text{at} \quad l = 0 \qquad (24)$$

$$\frac{\partial T(L,\tau)}{\partial l} = 0 \qquad \text{at} \quad l = L \qquad (25)$$

These boundary conditions are based on the consideration that mass and energy are neither created nor destroyed in the infinitesimal region $l = 0^-$ to $l = 0^+$.

In dimensionless form, the system equations become

$$\frac{\partial u_1}{\partial t} = \frac{1}{\beta}\frac{\partial^2 u_1}{\partial x^2} - \frac{\partial u_1}{\partial x} - \tau_r\phi(u_1, u_2) \qquad (26)$$

$$\frac{\partial u_2}{\partial t} = \frac{1}{\beta}\frac{\partial^2 u_2}{\partial x^2} - \frac{\partial u_2}{\partial x} - Q\tau_r\phi(u_1, u_2) - K\tau_r(u_2 - \theta) \qquad (27)$$

where

$$\phi = k_{10}\exp(-P_1/u_2)u_1 - k_{20}\exp(-P_2/u_2)(1 - u_1)$$

The dimensionless boundary conditions are

$$\frac{\partial u_1(0,t)}{\partial t} = \beta[u_1(0,t) - u_1^f] \qquad \text{at} \quad x = 0 \qquad (28)$$

$$\frac{\partial u_1(1,t)}{\partial t} = 0 \qquad \text{at} \quad x = 1 \qquad (29)$$

$$\frac{\partial u_2(0,t)}{\partial t} = \beta[u_2(0,t) - u_1^f] \qquad \text{at} \quad x = 0 \qquad (30)$$

$$\frac{\partial u_2(1,t)}{\partial t} = 0 \qquad \text{at} \quad x = 1 \qquad (31)$$

The following numerical values are assumed:

$$\beta = 5,\ \tau_r = 0.05\ \text{hr.},\ Q = -200,\ K = 30\ \text{hr.}^{-1}$$

$$k_{10} = 2.51 \times 10^5\ \text{hr.}^{-1},\ k_{20} = 1.995 \times 10^7\ \text{hr.}^{-1}$$

$$P_1 = 5.03,\ P_2 = 10.06,\ T_r = 1{,}000°\text{R.},$$
$$u_1^f = 0.9,\ u_2^f = 0.6$$

Initially the concentration and temperature profiles are assumed to be constant throughout the length of the reactor and at the values of the inlet conditions, that is

$$u_1(x,0) = u_1^f \qquad \text{at} \quad t = 0 \qquad (32)$$

$$u_2(x,0) = u_2^f \qquad \text{at} \quad t = 0 \qquad (33)$$

It is presupposed that a steady state operating point has been determined which is optimal with respect to some performance criterion (for example, maximum yield). The startup policy, in turn, is to be determined such that by controlling the addition or removal of heat, the process is driven from the initial state toward the final steady state in some optimal fashion. It is desired to minimize the spatial integral of the weighted sum of the squared concentration and temperature deviation from the desired steady state profiles, $u_{1d}(x)$ and $u_{2d}(x)$, integrated over a transient startup period of fixed length. The performance functional may thus be written as

$$S = \int_0^{t_f}\int_0^1 \{\mu[u_1(x,t) - u_{1d}(x)]^2 + \nu[u_2(x,t) - u_{2d}(x)]^2\}\,dx\,dt \qquad (34)$$

where μ and ν are suitably chosen constant weighting coefficients. The manipulated variable is the dimensionless wall temperature θ, which is considered to be a function of time only and lies within the range

$$\theta_{\min} \leq \theta(t) \leq \theta_{\max} \qquad (35)$$

The Hamiltonian, as defined by Equation (16), is

$$H = \mu[u_1(x,t) - u_{1d}(x)]^2 + \nu[u_2(x,t) - u_{2d}(x)]^2 + z_1\left[\frac{1}{\beta}u_{1xx} - u_{1x} - \tau_r\phi(u_1,u_2)\right] + z_2\left[\frac{1}{\beta}u_{2xx} - u_{2x} - Q\tau_r\phi(u_1,u_2) - K\tau_r(u_2 - \theta)\right] \qquad (36)$$

With reference to Equation (8), the adjoint partial differential equations corresponding to Equations (26) and (27), respectively, are

$$z_{1t} = -\frac{1}{\beta}z_{1xx} - z_{1x} + \tau_r\phi_{u_1}(u_1,u_2)z_1 + Q\tau_r\phi_{u_1}(u_1,u_2)z_2 - 2\mu[u_1(x,t) - u_{1d}(x)] \qquad (37)$$

$$z_{2t} = -\frac{1}{\beta}z_{2xx} - z_{2x} + \tau_r\phi_{u_2}(u_1,u_2)z_1 + [Q\tau_r\phi_{u_2}(u_1,u_2) + K_r]z_2 - 2\nu[u_2(x,t) - u_{2d}(x)] \qquad (38)$$

where

$$\phi_{u_1} = \frac{\partial\phi}{\partial u_1} = k_{10}\exp(-P/u_2) = k_{20}\exp(-p_2/u_2) \qquad (39)$$

$$\phi_{u_2} = \frac{\partial\phi}{\partial u_2} = \frac{1}{u_2^2}[P_1k_{10}\exp(-P_1/u_2)u_1 - P_2k_{20}\exp(-P_2/u_2)(1 - u_1)] \qquad (40)$$

Since the boundary conditions given by Equations (28) through (31) correspond to the general forms, Equations (10) and (11), the adjoint boundary conditions correspond to Equations (12) and (13), respectively. Thus

$$z_{ix}(0,t) = 0 \qquad \text{at} \quad x = 0,\ i = 1,2 \qquad (41)$$

$$z_{ix}(1,t) + \beta_i z_i(1,t) = 0 \qquad \text{at} \quad x = 1,\ i = 1,2 \qquad (42)$$

The final conditions on the adjoint variables corresponding to Equation (14) are

$$z_i(x,t_f) = 0 \qquad \text{at} \quad t = t_f,\ i = 1,2 \qquad (43)$$

The solution of Equations (26) and (27) forward in time from $t = 0$ to $t = t_f$ is accomplished by the use of quasilinearization (*10, 11*) together with an implicit difference scheme. The details of the computational method are presented in reference *12*.

To apply the quasilinearization technique, the nonlinear terms $\phi(u_1, u_2)$ in Equations (26) and (27) are first linearized by means of a first-order Taylor series expanded about the $(k-1)^{\text{th}}$ iterative solution $\mathbf{u}^{(k-1)}$ as follows:

$$\phi^{(k)} = \phi^{(k-1)} + \phi_{u_1}^{(k-1)}[u_1^{(k)} - u_1^{(k-1)}] + \phi_{u_2}^{(k-1)}[u_2^{(k)} - u_2^{(k-1)}] \qquad (44)$$

Substitution of Equation (44) into Equations (26) and (27) yields the following linearized recurrence relationship:

$$u_{1_t}^{(k)} = \frac{1}{\beta}u_{1_{xx}}^{(k)} - u_{1_x}^{(k)} - \tau_r[\phi^{(k-1)} + \phi_{u_1}^{(k-1)}(u_1^{(k)} - u_1^{(k-1)}) + \phi_{u_2}^{(k-1)}(u_2^{(k)} - u_2^{(k-1)})] \qquad (45)$$

$$u_{2_t}^{(k)} = \frac{1}{\beta} u_{2_{xx}}^{(k)} - u_{2_x}^{(k)} - Q\tau_r \left[\phi^{(k-1)} + \phi_{u_1}^{(k-1)}(u_1^{(k)} - u_1^{(k-1)}) + \phi_{u_2}^{(k-1)}(u_2^{(k)} - u_2^{(k-1)})\right] - K\tau_r[u_2^{(k)} - \theta] \quad (46)$$

The solution of Equations (45) and (46) is greatly simplified by "decoupling" the component equations. This is done in the i^{th} equation by setting

$$u_j^{(k)} - u_j^{(k-1)} = \begin{cases} u_i^{(k)} - u_i^{(k-1)}, & j = i \\ 0, & j \neq i \end{cases} \quad (47)$$

Then Equations (45) and (46) become

$$u_{1_t}^{(k)} = \frac{1}{\beta} u_{1_{xx}}^{(k)} - u_{1_x}^{(k)} - \tau_r \{\phi^{(k-1)} + \phi_{u_1}^{(k-1)}(u_1^{(k)} - u_1^{(k-1)})\} \quad (48)$$

$$u_{2_t}^{(k)} = \frac{1}{\beta} u_{2_{xx}}^{(k)} - u_{2_x}^{(k)} - Q\tau_r \{\phi^{(k-1)} + \phi_{u_2}^{(k-1)}(u_2^{(k)} - u_2^{(k-1)})\} - K\tau_r [u_2^{(k)} - \theta] \quad (49)$$

This is valid, for as convergence is attained, $u_i^{(k-1)}$ approaches $u_i^{(k)}$. The equations are still implicitly coupled as $\phi^{(k-1)}$ contains the solutions $u_1^{(k-1)}$ and $u_2^{(k-1)}$.

Equations (48) and (49) are most conveniently solved by finite-difference methods. The implicit scheme due to Crank and Nicolson (*13*) is considered here.

In applying the Crank-Nicolson method the spatial axis is discretized into M increments of equal length Δx so that Δx is $1/M$. Time discretization is effected by solving the difference equations at equal time increments Δt. The solution $u(m, n)$ denotes the value of the dependent variable at the spatial location $(m-1)\Delta x$ and at time $(n-1)\Delta t$.

The partial time derivatives are approximated by taking forward differences between the $(n-1)^{th}$ and n^{th} time steps, that is

$$u_t \approx \frac{1}{\Delta t}[u(m, n) - u(m, n-1)] \quad (50)$$

For spatial discretization, implicit difference operators are constructed for the first and second spatial partial derivatives by taking central differences, averaged over the $(n-1)^{th}$ and n^{th} time steps, that is

$$u_{xx} \approx \frac{M^2}{2}[u(m+1, n) - 2u(m, n) + u(m-1, n) + u(m+1, n-1) - 2u(m, n-1) + u(m-1, n-1)] \quad (51)$$

$$u_x \approx \frac{M}{4}[u(m+1, n) - u(m-1, n) + u(m+1, n-1) - u(m-1, n-1)] \quad (52)$$

The above difference operators have a discretization error on the order of $(\Delta x)^2$. The dependent variable u is also averaged over the $(n-1)^{th}$ and n^{th} time steps:

$$u \approx \frac{1}{2}[u(m, n) + u(m, n-1)] \quad (53)$$

The difference approximations for the first derivative terms occurring in the boundary conditions at $x = 0$ and $x = 1$ are taken to be three-point forward and backward differences, respectively:

$$u_x|_{x=0} \approx \frac{M}{2}[-u(3, n) + 4u(2, n) - 3u(1, n)] \quad (54)$$

$$u_x|_{x=1} \approx \frac{M}{2}[3u(M+1, n) - 4u(M, n) + u(M-1, n)] \quad (55)$$

Substitution of the above difference operators into the recurrence relations, Equations (48) and (49), and the boundary conditions, gives rise to the set of linearized difference equations for the k^{th} iteration:

$$\left[B_1^{(k-1)}(2, n) + \frac{4M}{2\beta + 3M} A_1\right] u_1^{(k)}(2, n) + \left[C_1 - \frac{M}{2\beta + 3M} A_1\right] u_1^{(k)}(3, n) = -A_1 u_1(1, n-1) - D_1^{(k-1)}(2, n) u_1(2, n-1) - C_1 u_1(3, n-1) - E_1^{(k-1)}(2, n) - \frac{2\beta u_1^f}{2\beta + 3M} A_1$$

$$A_1 u_1^{(k)}(m-1, n) + B_1^{(k-1)}(m, n) u_1^{(k)}(m, n) + C_1 u_1^{(k)}(m+1, n) = -A_1 u_1(m-1, n-1) - D_1^{(k-1)}(m, n) u_1(m, n-1) - C_1 u_1(m+1, n-1) - E_1^{(k-1)}(m, n) \quad m = 3, \ldots, M-1$$

$$\left[A_1 - \frac{1}{3} C_1\right] u_1^{(k)}(M-1, n) + \left[B_1^{(k-1)}(M, n) + \frac{4}{3} C_1\right] u_1^{(k)}(M, n) = -A_1 u_1(M-1, n-1) - D_1^{(k-1)}(M, n) u_1(M, n-1) - C_1 u_1(M+1, n-1) - E_1^{(k-1)}(M, n) \quad (56)$$

$$\left[B_2^{(k-1)}(2, n) + \frac{4M}{2\beta + 3M} A_2\right] u_2^{(k)}(2, n) + \left[C_2 - \frac{M}{2\beta + 3M} A_2\right] u_2^{(k)}(3, n) = -A_2 u_2(1, n-1) - D_2^{(k-1)}(2, n) u_2(2, n-1) - C_2 u_2(3, n-1) - E_2^{(k-1)}(2, n) - \frac{2\beta u_2^f}{2\beta + 3M} A_2$$

$$A_2 u_2^{(k)}(m-1, n) + B_2^{(k-1)}(m, n) u_2^{(k)}(m, n) + C_2 u_2^{(k)}(m+1, n) = -A_2 u_2(m-1, n-1) - D_2^{(k-1)}(m, n) u_2(m, n-1) - C_2 u_2(m+1, n-1) - E_2^{(k-1)}(m, n) \quad m = 3, \ldots, M-1$$

$$\left[A_2 - \frac{1}{3} C_2\right] u_2^{(k)}(M-1, n) + \left[B_2^{(k-1)}(M, n) + \frac{4}{3} C_2\right] u_2^{(k)}(M, n) = -A_2 u_2(M-1, n-1) - D_2^{(k-1)}(M, n) u_2(M, n-1) - C_2 u_2(M+1, n-1) - E_2^{(k-1)}(M, n) \quad (57)$$

where

$$A_1 = A_2 = \frac{M^2}{2\beta} + \frac{M}{4}$$

$$B_1^{(k-1)}(m, n) = -\frac{1}{2}\tau_r \phi_{u_1}^{(k-1)}(m, n) - \frac{M^2}{\beta} - \frac{1}{\Delta t}$$

$$B_2^{(k-1)}(m, n) = -\frac{1}{2} Q \tau_r \phi_{u_2}^{(k-1)}(m, n) - \frac{1}{2} K \tau_r - \frac{M^2}{\beta} - \frac{1}{\Delta t}$$

$$C_1 = C_2 = \frac{M^2}{2\beta} - \frac{M}{4}$$

$$D_1^{(k-1)}(m,n) = -\frac{1}{2}\tau_r\,\phi_{u1}^{(k-1)}(m,n) - \frac{M^2}{\beta} + \frac{1}{\Delta t}$$

$$D_2^{(k-1)}(m,n) = -\frac{1}{2}Q\,\tau_r\,\phi_{u2}^{(k-1)}(m,n) - \frac{1}{2}K\,\tau_r - \frac{M^2}{\beta} + \frac{1}{\Delta t}$$

$$E_1^{(k-1)}(m,n) = -\tau_r\,[\phi^{(k-1)}(m,n) - u_1^{(k-1)}(m,n)\phi_{u1}^{(k-1)}(m,n)]$$

$$E_2^{(k-1)}(m,n) = -Q\,\tau_r\,[\phi^{(k-1)}(m,n) - u_2^{(k-1)}(m,n)\phi_{u2}^{(k-1)}(m,n)] + K\,\tau_r\,\theta$$

The end-point values are determined from the following boundary equations:

$$u_1^{(k)}(1,n) = -\frac{M}{2\beta+3M}u_1^{(k)}(3,n) + \frac{4M}{2\beta+3M}u_1^{(k)}(2,n) + \frac{2\beta}{2\beta+3M}u_1^{f} \quad (58)$$

$$u_1^{(k)}(M+1,n) = \frac{4}{3}u_1^{(k)}(M,n) - \frac{1}{3}u_1^{(k)}(M-1,n) \quad (59)$$

$$u_2^{(k)}(1,n) = -\frac{M}{2\beta+3M}u_2^{(k)}(3,n) + \frac{4M}{2\beta+3M}u_2^{(k)}(2,n) + \frac{2\beta}{2\beta+3M}u_2^{f} \quad (60)$$

$$u_2^{(k)}(M+1,n) = \frac{4}{3}u_2^{(k)}(M,n) - \frac{1}{3}u_2^{(k)}(M-1,n) \quad (61)$$

Similarly the backward solution of Equations (37) and (38) from $t = t_f$ to $t = 0$ is accomplished by solving the following sets of implicit difference equations from $n = N$ to $n = 1$:

$$\begin{aligned}
&\left[B_3(2,n) + \frac{4}{3}A_3\right] z_1^{(k)}(2,n) + \left[C_3 - \frac{1}{3}A_3\right] z_2^{(k)}(3,n) \\
&\quad = -A_3\,z_1(1,n+1) - D_3(2,n+1)\,z_1(2,n+1) - C_3\,z_1(3,n+1) + E_3^{(k-1)}(2,n) \\
&A_3\,z_1^{(k)}(m-1,n) + B_3(m,n)z_1^{(k)}(m,n) + C_3\,z_1^{(k)}(m+1,n) = -A_3\,z_1(m-1,n+1) \\
&\quad - D_3(m,n+1)\,z_1(m,n+1) - C_3\,z_1(m+1,n+1) + E_3^{(k-1)}(m,n) \\
&\quad m = 3,\ldots,M-1 \\
&\left[A_3 - \frac{M}{2\beta+3M}C_3\right] z_1^{(k)}(M-1,n) + \left[B_3(M,n) + \frac{4M}{2\beta+3M}C_3\right] z_1^{(k)}(M,n) \\
&\quad = -A_3\,z_1(M-1,n+1) - D_3(M,n+1)\,z_1(M,n+1) + E_3^{(k-1)}(M,n)
\end{aligned} \quad (62)$$

$$\begin{aligned}
&\left[B_4(2,n) + \frac{4}{3}A_4\right] z_2^{(k)}(2,n) + \left[C_4 - \frac{1}{3}A_4\right] z_2^{(k)}(3,n) \\
&\quad = -A_4\,z_2(1,n+1) - D_4(2,n)\,z_2(2,n+1) - C_4\,z_2(3,n+1) + E_4^{(k-1)}(M,n) \\
&A_4\,z_2^{(k)}(m-1,n) + B_4(m,n)\,z_2^{(k)}(m,n) + C_4\,z_2^{(k)}(m+1,n) = -A_4\,z_2(m-1,n) \\
&\quad - D_4(m,n)\,z_2(m,n+1) - C_4\,z_2(m+1,n+1) + E_4^{(k-1)}(m,n) \\
&\quad m = 3,\ldots,M-1 \\
&\left[A_4 - \frac{M}{2\beta+3M}C_4\right] z_2^{(k)}(M-1,n) + \left[B_4(M,n) + \frac{4M}{2\beta+3M}C_4\right] z_2^{(k)}(M,n) \\
&\quad = -A_4\,z_2(M-1,n+1) - D_4(M,n)\,z_2(M,n+1) - C_4\,z_2(M+1,n+1) + E_4^{(k-1)}(M,n)
\end{aligned} \quad (63)$$

where

$$A_3 = A_4 = \frac{M^2}{2\beta} - \frac{M}{4}$$

$$B_3(m,n) = -\frac{1}{2}\tau_r\,\phi_{u1}(m,n) - \frac{M^2}{\beta} - \frac{1}{\Delta t}$$

$$B_4(m,n) = -\frac{1}{2}Q\,\tau_r\,\phi_{u2}(m,n) - \frac{1}{2}K\,\tau_r - \frac{M^2}{\beta} - \frac{1}{\Delta t}$$

$$C_3 = C_4 = \frac{M^2}{2\beta} + \frac{M}{4}$$

$$D_3(m,n) = -\frac{1}{2}Q\,\tau_r\,\phi_{u1}(m,n) - \frac{M^2}{\beta} + \frac{1}{\Delta t}$$

$$D_4(m,n) = -\frac{1}{2}Q\,\tau_r\,\phi_{u2}(m,n) - \frac{1}{2}K\,\tau_r - \frac{M^2}{\beta} + \frac{1}{\Delta t}$$

$$E_3^{(k-1)}(m,n) = Q\,\tau_r\,\phi_{u1}(m,n)\,z_2^{(k-1)}(m,n) - 2\mu\,[u_1(x,t) - u_{1d}(x)]$$

$$E_4^{(k-1)}(m,n) = \tau_r\,\phi_{u2}(m,n)\,z_1^{(k-1)}(m,n) - 2\nu\,[u_2(x,t) - u_{2d}(x)]$$

The end-point values for the adjoint variables are determined, using the following boundary equations:

$$z_1^{(k)}(1,n) = -\frac{1}{3}z_1^{(k)}(3,n) + \frac{4}{3}z_1^{(k)}(2,n) \quad (64)$$

$$z_1^{(k)}(M+1,n) = \frac{4M}{2\beta+3M}z_1^{(k)}(M,n) - \frac{M}{2\beta+3M}z_1^{(k)}(M-1,n) \quad (65)$$

$$z_2^{(k)}(1,n) = -\frac{1}{3}z_2^{(k)}(3,n) + \frac{4}{3}z_2^{(k)}(2,n) \quad (66)$$

$$z_2^{(k)}(M+1,n) = \frac{4M}{2\beta+3M}z_2^{(k)}(M,n)$$

$$-\frac{M}{2\beta + 3M} z_2^{(k)}(M-1, n) \quad (67)$$

In order to minimize the objective functional S with respect to $\theta(t)$, a functional gradient technique (*1, 5 to 7*) similar to that of Bryson and Denham (*8*) for lumped systems is used. From Equation (17) the maximum rate of decrease of S with respect to $\theta(t)$ is in the direction H_θ. Using the Hamiltonian defined in Equation (36), one gets the resulting change in control $\delta\theta(t)$ (*12*)

$$\delta\theta(t) = -\alpha \frac{\int_0^1 K\tau_r z_2(x,t)\,dx}{\left[\int_0^{t_f}\left[\int_0^1 K\tau_r z_2(x,t)\,dx\right]^2 dt\right]^{1/2}} \quad (68)$$

where α can be considered as a perturbation coefficient representing step length. Using Simpson's integration scheme, one can use values of $z_2(m, n)$, $m = 1, \ldots, M+1$, $n = 1, \ldots, N+1$, obtained by backward solution of the adjoint equations, in Equation (68) to compute $\delta\theta(n)$, $n = 1, \ldots, N$. The new control function $\theta(n)$ $n = 1, \ldots, N$ is then found from the relation

$$\theta(n)|_{\text{new}} = \theta(n)|_{\text{old}} + \delta\theta(n) \quad (69)$$

The computational scheme for obtaining the optimal control policy via the necessary conditions of the distributed maximum principle and the functional gradient technique can now be summarized as follows:

1. An initial control policy $\theta(t)$ is assumed.
2. Using the assumed policy in the system equations, we solve these forward in time from $t = 0$ to $t = t_f$ and retain the transient solutions.
3. The performance functional S is evaluated using the values of the state variables computed in step 2.
4. Using final values of the state variables to compute the final conditions on the adjoint variables and inserting computed values of the state variables where required in the adjoint equations, we solve these backward in time from $t = t_f$ to $t = 0$ and retain the transient solutions.
5. An improved control policy is calculated using the gradient technique with values of the adjoint variables computed in step 4.
6. Steps 2 through 5 are repeated iteratively until the objective functional converges to within a specified tolerance.

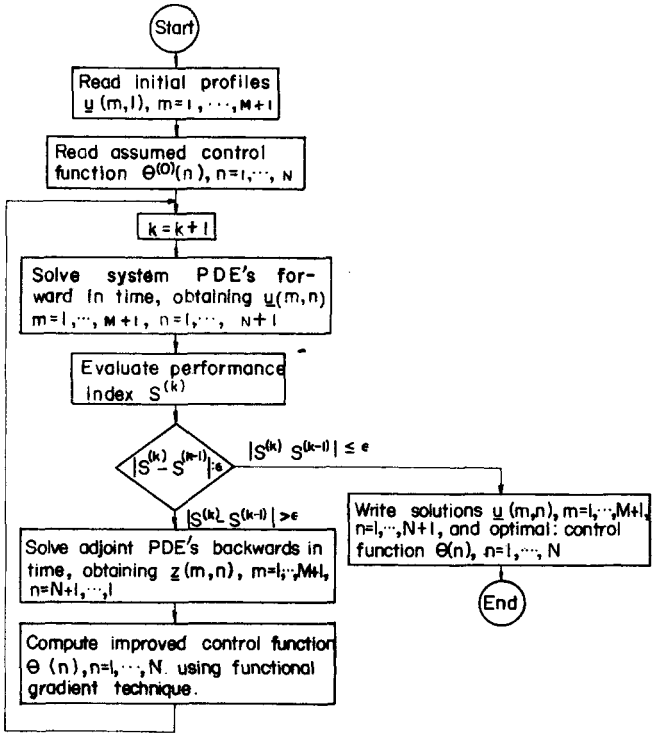

Fig. 1. Flow diagram for optimal control using the distributed maximum principle.

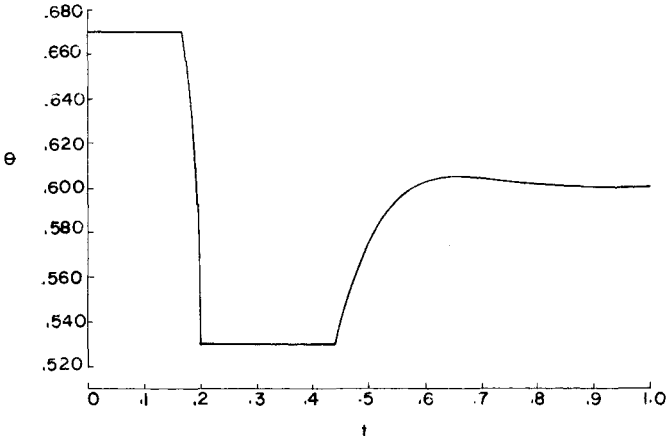

Fig. 2. Optimal startup control policy for tubular reactor using distributed maximum principle.

Computation was performed using a time increment of 0.02 residence time and a spatial increment of 0.05 dimensionless distance unit. The limits on the control were assumed to be $\theta_{\max} = 0.670$ and $\theta_{\min} = 0.530$ dimensionless temperature unit. The performance index S was evaluated using a terminal time of one residence time and the weighting coefficients μ and ν were each taken to be unity. The desired steady state profiles $u_{1d}(x)$ and $u_{2d}(x)$ were chosen to correspond to a control value of $\theta = 0.6$ dimensionless temperature unit.

With regard to the selection of the initial approximation, it should be noted that the steady state value of the control must be known and used in the initial assumed trajectory at the terminal time t_f. This is because the algorithm is incapable of shifting the control at terminal time, as seen from Equation (68) and the transversality conditions, Equation (43). Thus for convenience the initial control trajectory approximation was taken to be $\theta(t) = 0.6$ dimensionless temperature unit, the steady state value.

Figure 2 shows the optimal control trajectory obtained after 30 iterations, using a perturbation coefficient of $\alpha = 0.1$. This value, determined by trial and error, provided a reasonable rate of convergence of the performance index to a minimum without oscillations. The policy is seen to approach a bang-bang trajectory with maximum wall temperature applied to the system initially. At about 0.20 residence time a switch to minimum wall temperature occurs, followed by a singular approach to the steady state value starting at about 0.44 residence time.

The resulting transient concentration and temperature profiles, $u_1(x, t)$ and $u_2(x, t)$, obtained using the optimal startup policy, are shown in Figures 3 and 4, respectively, with time as a parameter. Shown in dashed lines for comparison are the transient profiles resulting from using steady state control. The value of the performance index obtained for the optimally controlled case was 0.046557 compared to 0.050435 with steady state control.

Since the optimal control policy so closely resembled bang-bang control, a purely bang-bang policy was considered. This consisted of starting with maximum effort, switching to minimum effort, and finally switching to the steady state control level. These two switch points were approximated from the optimal trajectory to be 0.20 and 0.50 residence times, respectively. The state equations were solved using the bang-bang policy, and the performance index Equation (34) was computed, yielding a value of 0.046555. Thus the performance index remained essentially unchanged using the bang-bang approximation and, because of its simplicity to implement, a bang-bang policy would probably be preferred for this application.

Constraint on Maximum Reaction Temperature

Suppose it is desired to determine the optimal startup control trajectory which minimizes the performance criterion Equation (34), while at the same time holding the maximum reaction temperature at or below a specified upper limit. This inequality constraint can be written as

$$u_2(x,t) - u_{2\max} \leqq 0 \qquad (70)$$

The constraint is introduced into the performance index as a penalty by means of a weighting coefficient σ as follows:

$$S = \int_0^{t_f} \int_0^1 \{\mu[u_1(x,t) - u_{1d}(x)]^2 + \nu[u_2(x,t) - u_{2d}(x)]^2 + \sigma[u_2(x,t) - u_{2\max}]^2 \hat{h}\,[u_2(x,t) - u_{2\max}]\}\,dx\,dt \qquad (71)$$

where $\hat{h}\,[u_2(x,t) - u_{2\max}]$ is the Heavyside unit step function. Thus a penalty is not invoked until a constraint is violated. The weighting coefficient σ is increased iteratively, stopping the iteration when the maximum temperature has converged to within a specified tolerance of the constraint boundary.

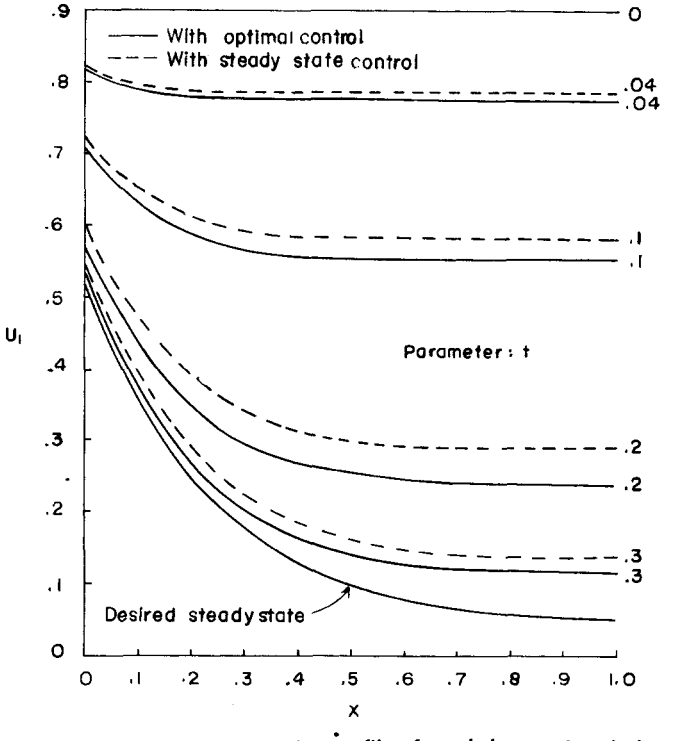

Fig. 3. Transient concentration profiles for tubular reactor startup.

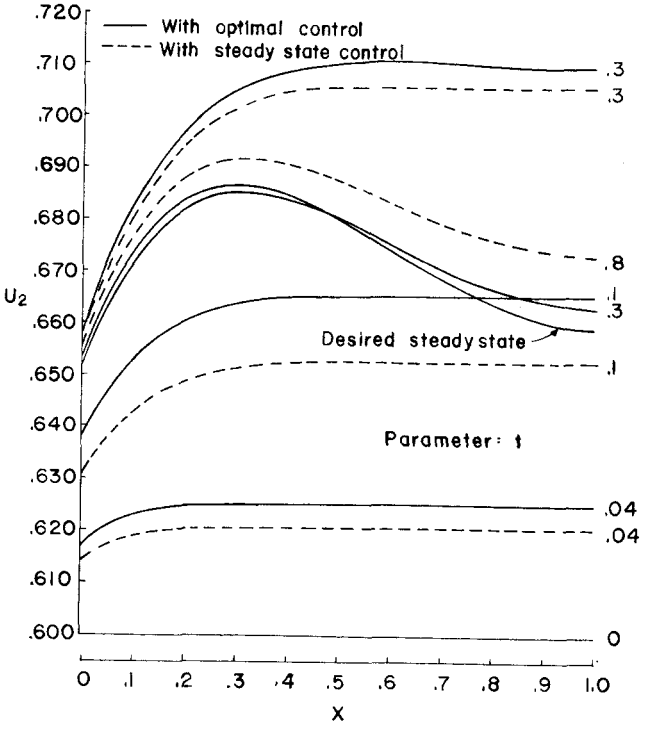

Fig. 4. Transient temperature profiles for tubular reactor startup.

State variable constraints such as this require a corresponding modification of the adjoint equations, since an additional term involving a state variable is introduced into the performance index. Thus the adjoint equations, Equations (37) and (38), become

$$z_{1t} = -\frac{1}{\beta} z_{1xx} - z_{1x} + \tau_r \phi_{u_1}(u_1, u_2) z_1 + Q\,\tau_r\,\phi_{u_1}(u_1, u_2) z_2 - 2\mu[u_1(x,t) - u_{1d}(x)] \qquad (72)$$

$$z_{2t} = -\frac{1}{\beta} z_{2xx} - z_{2x} + \tau_r \phi_{u_2}(u_1, u_2) z_1 + [Q\,\tau_r\,\phi_{u_2}(u_1, u_2) + K\tau_r] z_2 - 2\nu[u_2(x,t) - u_{2d}(x)] - 2\sigma\,[u_2(x,t) - u_{2\max}]\,\hat{h}\,[u_2(x,t) - u_{2\max}] \qquad (73)$$

Figure 5 shows the optimal control policy for the case of an upper constraint on maximum reaction temperature, $u_{2\max} = 0.700$ dimensionless temperature unit. The switch point from maximum to minimum wall temperature occurs earlier than for the case of an unconstrained state, thus reducing the amount of temperature overshoot. The following table lists the penalty weights σ and the corresponding maximum temperatures that resulted after each ascent:

σ	Maximum reaction temperature
(Unconstrained)	0.7165
1	0.7147
10	0.7093
10^2	0.7048
10^3	0.7014
10^4	0.7005

The iteration was stopped for $\sigma = 10^4$ as the resulting maximum temperature was considered in close enough proximity to the constraint boundary. Each time the value of σ was increased, it was found necessary to adjust the perturbation coefficient α downward in order to prevent

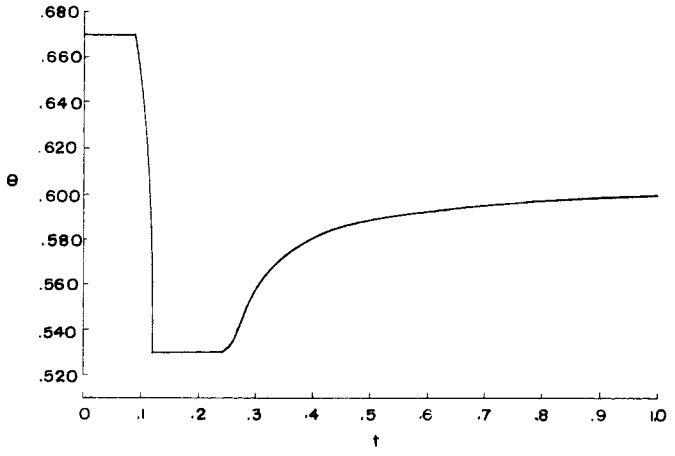

Fig. 5. Optimal startup control policy for tubular reactor using distributed maximum principle (with constraint on maximum reaction temperature; $U_{2\max} = 0.700$).

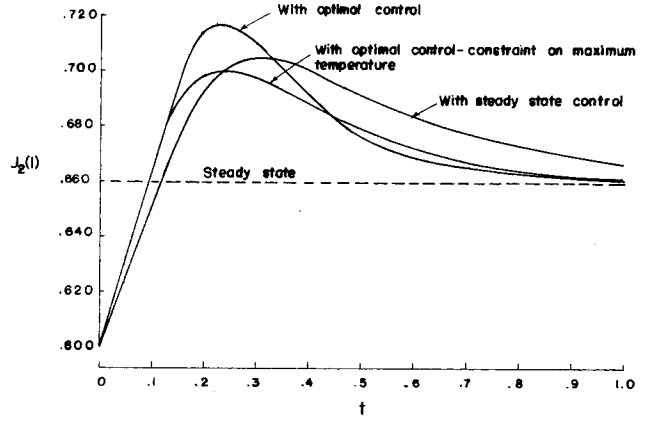

Fig. 6. Exit temperature trajectories for tubular reactor during startup.

oscillations and instability of the performance index.

Exit temperature trajectories for optimal control, with and without a state variable constraint, and for steady state control, are shown in Figure 6.

CONCLUSIONS

An optimal startup policy of a jacketed tubular reactor in which a first-order, reversible, exothermic reaction takes place is determined. A distributed maximum principle is presented for determining weak necessary conditions for optimality of diffusional distributed parameter systems.

Optimization of the two-point boundary value system of second-order, nonlinear, parabolic partial differential equations presents a formidable computational problem. An approximate numerical method which is iterative in nature, involving repeated numerical integration of the performance and adjoint equations, combined with the use of a functional gradient technique to improve the control vector, is introduced to overcome computational difficulties. A convenient method for handling inequality constraints involving state variables is also presented.

ACKNOWLEDGMENT

This study was supported in part by NASA Grant No. NGR-17-001-034 and by the Air Force Office of Scientific Research, Office of Aerospace Research, U.S. Air Force, under Contract F44620-68-0020 (Themis Project).

NOTATION

c_i = concentration of i^{th} component, lb.-mole/cu.ft.
C_p = specific heat, B.t.u./(lb.)(°R.)
D_m = mass diffusivity, sq.ft./hr.
E_i = activation energy of i^{th} component, B.t.u./lb.-mole
h = film coefficient, B.t.u./(sq.ft.)(°R.)(hr.)
k_{eff} = effective thermal conductivity, B.t.u./(ft.)(°R.)(hr.)
k_i = rate constant, hr.^{-1}
k_{i0} = frequency factor, hr.^{-1}
l = axial distance, ft.
L = length of reactor, ft.
r = radius of reactor, ft.
R = gas constant, B.t.u./(lb.-mole)(°R.)
T = reaction temperature, °R.
T_r = reference temperature, °R.
T_w = wall temperature, °R.
v = velocity of flow, ft./hr.
ΔH = heat of reaction, B.t.u./lb.-mole
ρ = density, lb./cu.ft.
τ = time, hr.

Parameters

τ_r = L/v hr. = mean residence time
β = vL/D = axial Peclet number
t = τ/τ_r = dimensionless time
x = l/L = dimensionless axial distance
u_1 = $C_A/(C_A + C_B)$ = dimensionless concentration of A
u_2 = T/T_r = dimensionless reaction temperature
θ = T_w/T_r = dimensionless wall temperature

Other Parameters

$$Q = \frac{\Delta H\,(C_A + C_B)}{C_p\,\rho\,T_r}$$

$$K = \frac{2h}{C_p\,\rho\,r},\ \text{hr.}^{-1}$$

$P_1 = E_1/R\,T_r$
$P_2 = E_2/R\,T_r$

LITERATURE CITED

1. Denn, M. M., R. D. Gray, Jr., and J. R. Ferron, *Ind. Eng. Chem. Fundamentals*, **7**, 286-295 (1968).
2. Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidge, and E. F. Mishchenko, "The Mathematical Theory of Optimal Processes," English translation by K. N. Trirogoff, Interscience, New York (1962).
3. Fan, L. T., "The Continuous Maximum Principle," Wiley, New York (1966).
4. Katz, S., *J. Electronics Control*, **16**, 189-222 (1964).
5. Jackson, R., *Proc. Intern. Chem. Eng.*, **4**, 32-38 (1965).
6. Seinfeld, J. H., Ph.D. thesis, Princeton Univ., N. J. (1967).
7. Courant, R., and D. Hilbert, "Methods of Mathematical Physics," Vol. 1, Interscience, New York (1965).
8. Bryson, A. E., and W. F. Denham, *J. Appl. Mech.*, **29**, 247-257 (1962).
9. Danckwerts, P. V., *Chem. Eng. Sci.*, **2**, 1-13 (1953).
10. Bellman, R. E., and R. E. Kalaba, "Quasilinearization and Nonlinear Boundary Value Problems," American Elsevier, New York (1965).
11. Sage, A. P., and S. P. Chandhuri, *Intern. J. Control*, **6**, 81-98 (1967).
12. Hahn, D. R., Ph.D. thesis, Kansas State Univ., Manhattan (1969).
13. Forsythe, G. E., and W. R. Wasow, "Finite Difference Methods for Partial Differential Equations," Wiley, New York (1965).

Manuscript received February 16, 1970; revision received July 18, 1970; paper accepted November 16, 1970.