Optimal Control of Tubular Reactors

K. S. CHANG and S. G. BANKOFF
Northwestern University, Evanston, Illinois

Part I. Computational Considerations

The necessary conditions for optimization of a system governed by a nonlinear vector first-order partial differential equation with two (space and time) independent variables, such as governs the unsteady behavior of tubular flow reactors, are derived. Rather general objective functionals and boundary conditions, such as the recycle of unconverted reactant with an appropriate time delay for separation and a free choice of final time, are allowed. A gradient technique in control space is formulated, and it is shown that distinct computational advantages can accrue from the use of the method of characteristics.

Despite the formidable computational problems, quite a bit of attention has been devoted to the optimal control of distributed parameter processes in recent years. Work on the problem was initiated by Butkovskii and Lerner (*1*), and necessary conditions were developed by Butkovskii (*2, 3*), Luré (*4*), Katz (*5*), and Wang and Tung (*6, 7*). The latter workers, in particular, discussed the observability, controllability, and stability of these systems. Schmaedeke (*8*) has established some existence theorems for optimal control. Butkovskii (*9*) suggested as computational methods spatial discretization of the partial differential equations, and the use of the method of moments (*10*), which requires a knowledge of the system eigenfunctions. A function space method has been applied by Chaudhuri (*11*) to a linear system. In the chemical field, Volin and Ostrovskii (*12*) and Jackson (*13*) have dealt with optimal control of tubular reactors with slowly decaying catalyst, Denn, et al. (*14*) with a reactor with radial temperature gradients, Jackson (*15*) with processes, such as absorption columns, which are approximately describable by partial differential equations, and Denn (*16*), as well as others, with temperature regulation of slabs. The present work may be looked upon as an extension of Sirazetdinov's (*17*) formulation of the necessary conditions for a process governed by a first-order partial differential equation to a system of these equations, such as arise in the description of a tubular reactor. Very general boundary conditions are permitted, such as the recycle of unconverted reactant, with an appropriate time delay, back to the inlet, perhaps after some separation has occurred. In addition, the final time of the optimal control period may be left free, and hence be chosen optimally. A steepest descent scheme in control space is evolved, and the method of characteristics examined to determine its applicability. In the case where the adjoint and state equation characteristics are identical, it offers considerable computational advantages.

STATEMENT OF PROBLEM

We begin by defining a very general optimal control problem for flow processes governed by a system of first-order partial differential equations (PDE) in space, x, and time, t. The system is

$$\frac{\partial v}{\partial t} = f(t, x, v, \mu, u) \tag{1}$$

where $v \equiv [v_1(t,x), v_2(t,x), \ldots, v_n(t,x)]'$,

$$\mu \equiv (\mu_1, \mu_2, \ldots, \mu_n)' = \left(\frac{\partial v_1}{\partial x}, \frac{\partial v_2}{\partial x}, \ldots, \frac{\partial v_n}{\partial x}\right)',$$

$$u \equiv [u_1(t,x), u_2(t,x), \ldots, u_m(t,x)]',$$

and $f \equiv (f_1, f_2, \ldots, f_n)'$ is a vector-valued function with

$$f_i = f_i^0(t, x, v, u) + f_i^1(t, x, u)\frac{\partial v_i}{\partial x} \qquad i = 1, 2, \ldots, n \tag{2}$$

The notation for transpose $(\ )'$ will be used only if there is danger of ambiguity in the context. The function f is defined on the rectangle $[t_o, t_f] \times [x_o, x_f] \equiv I \times E$ and $f \epsilon C_p^2$ (in the class of all twice piece-by-piece continuously differentiable functions). The vector $v = v(t,x)$ determines the state of the system at $(t, x) \epsilon I \times E$. If the vector $u = u(t,x)$ is a bounded piecewise continuous control on $I \times E$, and $u \epsilon \Omega$, it is termed admissible. Here Ω is the set defined by

$$\Omega = \{u | g(u) \leq 0\} \tag{3}$$

where $g = (g_1, g_2, \ldots, g_r)$ $(r \leq m)$ is a vector-valued function representing a set of r constraint hypersurfaces.

For a chemical reactor the components of the state vector will be quantities such as the extents of the various reactions and the temperature. In general, the initial conditions will be prescribed:

$$v(t_o, x) = \Phi(x); \quad x_o \leq x \leq x_f \tag{4}$$

where $\Phi \epsilon C^2$. If there is no product recycle, there will also be a specification of the inlet conditions as a function of time

$$v(t, x_o) = \Psi(t); \quad t_o \leq t \leq t_f \tag{5}$$

where $\Psi \epsilon C^2$ and continuity at (t_o, x_o): $\Phi(x_o^+) = \Psi(t_o^+)$ is not necessarily assumed.

More generally, however, some of the product stream may be recycled, after further processing, with a time lag t_r, back to the inlet stream, so a functional relationship of the form

$$F[v(t, x_o),\ v(t - t_r, x_f)] = 0; \quad t_o \leq t \leq t_f; \quad F \epsilon C^2 \tag{6}$$

exists where $v(t - t_r, x_f) = m(t)$, a known function, for $0 \leq t \leq t_r$. It is seen that Equation (5) is a special form of (6). Moreover, the final time, t_f, may be subject to variation, as in the case of minimum-time problems.

Subject to the foregoing equality and inequality constraints, we now wish to minimize a general functional, of the form

$$J = J_1 + J_2 + J_3 \tag{7}$$

where

$$J_1 = \int_{x_o}^{x_f} G_1[t_f, x, v(t_f, x), \mu(t_f, x), u(t_f, x)]dx$$

$$J_2 = \int_{t_o}^{t_f} G_2[t, x_f, v(t, x_f), \mu(t, x_f), u(t, x_f)]dt \tag{8}$$

$$J_3 = \int_{t_o}^{t_f}\int_{x_o}^{x_f} G_3[t, x, v(t,x), \mu(t,x), u(t,x)]dxdt$$

and G_1, G_2, and G_3 are scalar functions (ϵC^2) of the

K. S. Chang is at the University of Waterloo, Waterloo, Ontario, Canada.

vector variables.

In analogy with the interconvertibility of the Mayer and Bolza forms of the one-dimensional extremal problem, if G_1 and G_2 are not functions of μ and u, one can define functionals J_1' and J_2' of the form of J_3 by

$$J_1' = \int_{x_o}^{x_f}\int_{t_o}^{t_f} \frac{\partial G_1}{\partial t}\,dtdx = J_1 - \int_{x_o}^{x_f} G_1[t_o, x, v(t_o, x)]dx \tag{9}$$

$$J_2' = \int_{t_o}^{t_f}\int_{x_o}^{x_f} \frac{\partial G_2}{\partial t}\,dxdt = J_2 - \int_{t_o}^{t_f} G_2[t, x_o, v(t, x_o)]dt \tag{10}$$

where the integrals on the extreme right side are often specified by the initial and boundary conditions. It will thus be convenient to consider G_1 and G_2 to be independent of u and μ in the subsequent development.

NECESSARY CONDITIONS FOR OPTIMALITY

Define a Hamiltonian function

$$H(t, x, v, \mu, p, u) \equiv -G_3 + <p, f> \tag{11}$$

where $p \equiv (p_1, p_2, \ldots, p_n)$ is an adjoint, or auxiliary, vector function on $I \times E$. The scalar function G_3 is presumed to be linear in the gradients of the state variables:

$$G_3 = \alpha_o(t, x, v, u) + <\alpha, \mu> \tag{12}$$

where α_o is a scalar function, and $\alpha(t, x, u)$ is an n-dimensional vector-valued function.

Two general classes of minimization problems are considered: 1. $u \equiv u(t, x)$ and 2. $u \equiv u(t)$. The case where $u \equiv u(x)$ follows by direct analogy to the latter case. The problem is thus to find the admissible control $u(t, x)$ [or $u(t)$] which minimizes J. If $u \equiv u(t)$, define

$$\overline{H} \equiv \int_{x_o}^{x_f} H(t, x, v, \mu, p, u)dx \tag{13}$$

The admissible control $u \equiv u(t, x)$ $[u(t)]$ is said to satisfy the maximum condition if at any (t, x) $\epsilon I \times E$ $[t\epsilon I]$ the function $H[\overline{H}]$ attains its absolute maximum. The maximum principle is contained in the following theorems:

1. If an admissible control $u(t, x)$ $[u(t)]$ defined on $I \times E[I]$ minimizes J, then it satisfies the maximum condition almost everywhere on $I \times E[I]$.

2. If the system of PDE (1) and the objective functional (7) are both linear in v and μ, the condition of theorem 1 is both necessary and sufficient.

The proof of the theorems essentially lies in the derivation of a formula for the incremental change of the functional due to arbitrary incremental changes of the function arguments and in the estimates of the second-order increments of the functions involved, to some extent following Rozonoér (*18*) for ordinary vector differential equations and Sirazetdinov (*17*) for a scalar first-order partial differential equation. The complete proof is rather detailed, but fairly straightforward, and interested readers are referred to (*19*). We content ourselves with an exposition of the first-order terms, in order to develop the auxiliary boundary and final conditions. Substituting (11) into (7) and (8), one obtains

$$J = J_1 + J_2 + \int_{t_o}^{t_f}\int_{x_o}^{x_f} \left(<\frac{\partial v}{\partial t}, p> - H \right) dxdt \tag{14}$$

$$\begin{aligned}\Delta J &= \Delta J_1 + \Delta J_2 \\ &+ \int_{t_o}^{t_f}\int_{x_o}^{x_f} \Big\{ -\Delta_u H + <\frac{\partial v}{\partial t} - \nabla_p H, \Delta p> \\ &\quad + <-\frac{\partial p}{\partial t} - \nabla_v H + \frac{\partial}{\partial x}(\nabla_\mu H), \Delta v> \\ &\quad + \frac{\partial}{\partial t}<p, \Delta v> - \frac{\partial}{\partial x}<\nabla_\mu H, \Delta v> \Big\} dxdt \\ &+ \Delta t_f \int_{x_o}^{x_f} \left(<\frac{\partial v}{\partial t}, p> - H \right)_{t_f, x} dx\end{aligned} \tag{15}$$

where

$$\Delta_u H \equiv H(t, x, v, \mu, p, u + \Delta u) - H(t, x, v, \mu, p, u) \tag{16}$$

and

$$\Delta J_1 = \int_{x_o}^{x_f} \left\{ <\nabla_v G_1, dv>_{t_f, x} + \left(\frac{\partial G_1}{\partial t_f}\right)_{t_f, x} \Delta t_f \right\} dx \tag{17}$$

$$\Delta J_2 = \int_{t_o}^{t_f} \{<\nabla_v G_2, dv>_{t, x_f}\}\,dt + (G_2)_{t_f, x_f}\Delta t_f \tag{18}$$

The total differential notation implies the sum of the variations due to path and final time:

$$dv(t_f, x_f) = (\Delta v + f\Delta t_f)_{t_f, x_f} \tag{19}$$

The first inner product in the double integral in (15) vanishes, in virtue of (1) and (11). The second inner product also vanishes, providing that the auxiliary PDE is satisfied in $I \times E$:

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}(\nabla_\mu H) - \nabla_v H \tag{20}$$

Performing the indicated integrations in (15), and collecting terms with (17) and (18), one obtains for the integrated terms:

$$\begin{aligned}\Delta J = &\int_{t_o}^{t_f}\int_{x_o}^{x_f} \{-\Delta_u H\}\,dxdt + (G_2)_{t_f, x_f}\Delta t_f \\ &+ \int_{x_o}^{x_f} \left[<\nabla_v G_1, dv>_{t_f, x} + \left(\frac{\partial G_1}{\partial t_f}\right)_{t_f, x} \Delta t_f \right. \\ &\qquad \left. + <p, dv>\Big|_{t_o, x}^{t_f, x} \right] dx \\ &+ \int_{t_o}^{t_f} \left(<\nabla_v G_2, dv>_{t, x_f} - <\nabla_\mu H, dv>\Big|_{t, x_o}^{t, x_f} \right) dt \\ &+ \Delta t_f \int_{x_o}^{x_f} \left(<\frac{\partial v}{\partial t}, p> - H \right)_{t_f, x} dx\end{aligned} \tag{21}$$

Similarly, taking the first variation of the boundary condition (6), we obtain

$$\begin{aligned}dF = &<\nabla_{v(t, x_o)}F, dv(t, x_o)> \\ &+ <\nabla_{v(t-t_r, x_f)}F, dv(t - t_r, x_f)> = 0, \\ &\qquad t_o \leq t \leq t_f\end{aligned} \tag{22}$$

where

$$\begin{aligned} dv(t - t_r, x_f) &= 0 && t - t_r < t_o \\ dv(t, x_o) &= \Delta v(t, x_o) && t_o \leq t \leq t_f\end{aligned} \tag{23}$$

in view of the assumption that the initial time is fixed ($\Delta t_o = 0$). The integral with respect to time in (21) can therefore be made to vanish if

$$(\nabla_v G_2 - \nabla_\mu H)_{t, x_f} \begin{cases} = \gamma\,(\nabla_{v(t-t_r, x_f)}F)_{t, x_f} & t_o \leq t \leq t_f - t_r \\ = 0 & t_f - t_r < t \leq t_f \end{cases} \tag{24}$$

$$(\nabla_\mu H)_{t, x_o} = \gamma\,(\nabla_{v(t, x_o)}F)_{t, x_o} \qquad t_o \leq t \leq t_f \tag{25}$$

where γ is any nonzero constant. By taking inner products, the proportionality constant can be eliminated:

Therefore, the incremental change ΔJ due to the incremental change in arguments involved is readily shown to be, to the first order

$$< (\nabla_v G_2 - \nabla_\mu H)_{t,xf},\ \nabla_{v(t,x_o)} F\,[v(t,x_o), v(t-t_r,x_f)]> = \begin{cases} <(\nabla_\mu H)_{t,x_o},\ \nabla_{v(t-t_r,xf)} F\,[v(t,x_o), v(t-t_r,x_f)]> & \text{for } t_o \leq t \leq t_f - t_r \\ 0 & \text{for } t_f - t_r < t \leq t_f \end{cases} \quad (26)$$

This corresponds to a two-point boundary condition in one-dimensional problems, the value of the adjoint variable at x_o being related to that at x_f. Similarly, by choosing

$$(\nabla_v G_1 - p)_{tf,x} = 0 \qquad x_o \leq x \leq x_f \quad (27)$$

a final condition on p is obtained. Note that $dv(t_o, x) = 0$, in view of (4) and (23). Finally, the stopping condition

$$(G_2)_{tf,xf} + \int_{x_o}^{x_f} \left\{ \left(\frac{\partial G_1}{\partial t_f} \right)_{tf,x} + \left(<\frac{\partial v}{\partial t}, p> - H \right)_{tf,x} \right\} dx = 0 \quad (28)$$

determines t_f.

The recycle condition, together with free final time, make the computations extremely time consuming, and we shall in this work confine our attention to once-through reactors with fixed final time for the optimization interval. Furthermore, the objective functional will be taken of the form:

$$J = \int_{t_o}^{t_f} \int_{x_o}^{x_f} \Gamma(t, x, v, \mu, u)\, dxdt \quad (29)$$

where $G_3 = \Gamma$. Corresponding to (4) and (5), the auxiliary final and boundary conditions then become

$$p(t_f, x) = 0 \qquad x_o \leq x \leq x_f \quad (30)$$

$$(\nabla_\mu H)_{t,xf} = 0 \qquad t_o \leq t \leq t_f \quad (31)$$

COMPUTATIONAL ALGORITHM

Control Space Gradient Method

A relaxation procedure based on first-order variations, analogous to that of Merriam (*20*) for ordinary differential equations with an integral error index, can now be derived.

Suppose that the values from the ith iteration:

$$\frac{\partial v^{(i)}}{\partial t} = f^{(i)}\,[t, x, v^{(i)}, \mu^{(i)}, u^{(i)}] \quad (32)$$

are available. Since the state and auxiliary equations and initial and boundary conditions are to be satisfied at every iteration, it follows that the current value of the objective functional may be obtained from

$$J^{(i)} = \int_{t_o}^{t_f} \int_{x_o}^{x_f} \left\{ \Gamma^{(i)} - <p^{(i)}, \left[f^{(i)} - \frac{\partial v^{(i)}}{\partial t} \right] > \right\} dxdt \quad (33)$$

We wish to choose a new control, $u^{(i+1)}$, such that $J^{(i+1)} < J^{(i)}$, providing the step size is made sufficiently small. It is readily shown by expanding the integrand in $J^{(i+1)}$ in a Taylor series about the ith iteration that a sufficient set of conditions is given by

$$u_j^{(i+1)} = u_j^{(i)} + \epsilon_j^{(i)} \left(\frac{\partial H}{\partial u_j} \right)^{(i)}; \quad u = u(x, t) \quad (34)$$

$$u_j^{(i+1)} = u_j^{(i)} + \epsilon_j^{(i)} \int_{x_o}^{x_f} \left(\frac{\partial H}{\partial u_j} \right)^{(i)} dx; \quad u = u(t) \quad (35)$$

$$u_j^{(i+1)} = u_j^{(i)} + \epsilon_j^{(i)} \int_{t_o}^{t_f} \left(\frac{\partial H}{\partial u_j} \right)^{(i)} dt; \quad u = u(x) \quad (36)$$

for $\epsilon_j^{(i)}$ $(j = 1, 2, \ldots, m)$ are the positive step size factors, sufficiently small.

Numerical Scheme

If the objective functional has the form

$$J = \int_{t_o}^{t_f} \mathcal{G}(t, x_f, v(t, x_f))\, dt \quad (37)$$

the method of characteristics leads to a convenient numerical procedure. A discussion of the characteristic equations is given in Appendix I. Consider a particular point (σ, x_f) on the terminal line where $t_o \leq \sigma \leq t_f$. The set of characteristic curves $X_i(s'; \sigma, x_f)$ emanating backward from the point (σ, x_f) to the initial curve, where the initial data are given, is determined from (A5) in Appendix I. By forward integration with respect to the parameter s of the state equations of (A4) along the curves $X_i(s'; \sigma, x_f)$, one can then obtain $v_i(\sigma, x_f)$ $(i = 1, 2, \ldots, n)$, where the initial data $v_i[t(0), X_i(0; \sigma, x_f)]$ are the values of either $\Phi(x)$ or $\Psi(t)$ at $X_i(0; \sigma, x_f)$. The function $\mathcal{G}[\sigma, x_f, v(\sigma, x_f)]$ at any time σ can thus be computed, and the functional J can be evaluated from (37). Hence, if $\mathcal{G}$ is minimized for $\forall \sigma \epsilon [t_o, t_f]$, so is J. If the control vector $u[t(s), X_i(s; \sigma, x_f)]$ is involved in the integrands of both (A5) and (A4), the solution procedure for both $X_i(s'; \sigma, x_f)$ and $v_i(\sigma, x_f)$ is iterative, and it is doubtful that this method offers significant advantages. However, if the functions f_1^1 $(i = 1, 2, \ldots, n)$ are independent of the control vector u, the characteristic curves are determined by (A5) once and for all, and the iteration proceeds only on the system (A4). The auxiliary system of characteristic equations must also be integrated. Details of this development are given in (19). The computation involves a laborious interpolation procedure, and is hence distinctly unattractive unless the characteristic curves coincide. This is further simplified if all state equations have the same principal part. Fortunately, this is usually the case with single phase flow processes, and indeed with many multiphase processes. With this proviso, Equations (1) to (5) become identical and can be written as a scalar equation:

$$X_o(s'; \tau, \xi) = \xi - \int_{s(\tau,\xi)}^{s'} f_o^1\{t(s), u[t(s), X_o(s; \tau, \xi)]\}\, ds \quad (38)$$

Along this characteristic, Equation (A4) applies for $v(\tau, \xi)$ with similar expressions for $p(\tau, \xi)$ and $\mu(\tau, \xi)$.

In the selection of the next u, three cases arise depending on whether $u = u(t, x)$, $u = u(t)$, or $u = u(x)$. If $u = u(t, x)$, the value $u[t_1(s), X_o(s; \tau_1, \xi_1)]$ of u at $[t_1(s), X_o(s; \tau_1, \xi_1)]$ of $I \times E$ on a characteristic curve passing through (τ_1, ξ_1) is independent of the value $u[t_2(s), X_o(s; \tau_2, \xi_2)]$ of u at $[t_2(s), X_o(s; \tau_2, \xi_2)]$ on a neighboring nonintersecting characteristic curve. This has the great advantage of reducing the two-dimensional optimization calculation to a sequence of one-dimensional optimizations, each of which can be performed independently. The algorithm (34) is used for this case. However, if $u = u(t)$, $(\partial H/\partial u_j)^{(i)}$ must be evaluated at mesh points in the net along each characteristic curve, and the integral in (35) evaluated. A similar observation can be made when $u = u(x)$.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation (Grant GK-1126).

NOTATION

A = matrix

C_2 = space of twice continuously differentiable functions

d = total variation
E = distance interval, $[x_o, x_f]$
F = vector-valued function, $(F_1, F_2, \ldots, F_n)$
F = recycle function
f = vector-valued function, $(f_1, f_2, \ldots, f_n)$
f_i = scalar-valued function, $f_i^o + f_i^1 \partial v_i/\partial x$
f^o = vector-valued function, $(f_1^o, f_2^o, \ldots, f_n^o)$
f^1 = vector-valued function, $(f_1^1, f_2^1, \ldots, f_n^1)$
f_o^1 = scalar function, $(= f_1^1 = f_2^1 = \ldots = f_n^1)$
G_1, G_2, G_3 = integrands of functionals
$\mathcal{G}$ = integrand of J
g = vector-valued function, $(g_1, g_2, \ldots, g_r)$
H = function defined by $-G_3 + <p, f>$
$\overline{H}$ = integrated H with respect to x
I = identity matrix
I = time interval, $[t_o, t_f]$
$J, J_1, J_2, J_3, J_2', J_3'$ = functionals
$m(t)$ = vector-valued function
m = integer
p = auxiliary vector, $(p_1, p_2, \ldots, p_n)$
s = parameter along characteristic curve
t = time variable
t_o = initial time
t_f = final time
t_r = time lag
t_1, t_2 = value of t
u = control vector, $(u_1, u_2, \ldots, u_m)$
v = state vector, $(v_1, v_2, \ldots, v_n)$
$X_o(s; \tau, \xi)$ = scalar solution of characteristic equation for v passing through (τ, ξ) of $I \times E$
$X_i(s; \tau, \xi)$ = ith characteristic passing through (τ, ξ) of $I \times E$
x = distance variable
x_o = initial distance
x_f = final distance

Greek Letters

α = vector-valued function, $(\alpha_1, \alpha_2, \ldots, \alpha_n)$
α_o = scalar function
Γ = integrand of J
γ = nonzero constant
Δ = increment
ϵ = iteration step size
κ = characteristic root
λ = vector, $(\lambda_1, \lambda_2, \ldots, \lambda_n) \equiv (\partial v_1/\partial t, \partial v_2/\partial t, \ldots, \partial v_n/\partial t)$
μ = vector, $(\mu_1, \mu_2, \ldots, \mu_n) \equiv (\partial v_1/\partial x, \partial v_2/\partial x, \ldots, \partial v_n/\partial x)$
ξ, ξ_1, ξ_2 = values of x
π_i = initial data μ_i
σ = value of t
τ, τ_1, τ_2 = values of t
Φ = initial state vector, $(\Phi_1, \Phi_2, \ldots, \Phi_n)$
ϕ = vector-valued function, $(\phi_1, \phi_2, \ldots, \phi_n)$
Ψ = boundary state vector, $(\Psi_1, \Psi_2, \ldots, \Psi_n)$
Ω = region defined by Equation (3)

Subscripts

t, x, λ, μ, p, u = partial differentiation with respect to t, x, μ, p and u
i, j = ith and jth components of vectors

Superscripts

(i) = ith iteration

Symbols

$<,>$ = inner product
$(\,)'$ = transpose
∇_v, ∇_μ = gradient operators defined by $(\partial/\partial v_1, \partial/\partial v_2, \ldots, \partial/\partial v_n)'$ $(\partial/\partial \mu_1, \partial/\partial \mu_2, \ldots, \partial/\partial \mu_n)'$, ... on scalars
$\nabla_v f$ = Jacobian matrix

LITERATURE CITED

1. Butkovskii, A. G., and A. Y. Lerner, *Auto. Remote Control*, **21**, 472 (1960).
2. Butkovskii, A. G., *ibid.*, **22**, 13 (1961).
3. *Ibid.*, 1156 (1961).
4. Luré, K. A., *J. Appl. Math. Mech.*, **27**, 1284 (1963); G. Leitman, ed., "Topics in Optimization," Chapt. 5, Academic Press, New York (1967).
5. Katz, S., *J. Elec. Control*, **16**, 189 (1964).
6. Wang, P. K. C., and F. Tung, *J. Basic Eng.*, **D86**, 67 (1964).
7. ———, *Trans. Aut. Control*, **AC-9**, 13 (1964).
8. Schmaedeke, W., *SIAM Control*, **5**, 138 (1967).
9. Butkovskii, A. G., *Auto. Remote Control*, **22**, 1429 (1961).
10. *Ibid.*, **24**, 1106 (1963).
11. Chaudhuri, A. K., *Int. J. Control*, **2**, 365 (1965).
12. Volin, Y. M., and G. M. Ostrovskii, *Auto. Remote Control*, **25**, 1271 (1964).
13. Jackson, R., I.Ch.E.-AIChE Joint Meeting, London (1964).
14. Denn, M. M., R. D. Gray, Jr., and J. R. Ferron, *Ind. Eng. Chem. Fundamentals*, **5**, 59 (1966).
15. Jackson, R., *Int. J. Control*, **4**, 127, 585 (1966).
16. Denn, M. M., *ibid.*, **4**, 167 (1966).
17. Sirazetdinov, T. K., *Auto. Remote Control*, **25**, 431 (1964).
18. Rozonoér, L. I., *ibid.*, **20**, 1288, 1405, 1517 (1959).
19. Chang, K. S., Ph.D. thesis, Northwestern Univ., Evanston, Ill. (1967).
20. Merriam, G. W., III, *Inform. Control*, **8**, 215 (1965).
21. Courant, R., and D. Hilbert, "Methods of Mathematical Physics," Vol. II, Interscience, New York (1962).
22. Garabedian, P. R., "Partial Differential Equations," John Wiley, New York (1964).

Manuscript received October 26, 1967; revision received April 15, 1968; paper accepted April 27, 1968.

APPENDIX I. DEVELOPMENT OF CHARACTERISTIC EQUATIONS

The characteristic determinant of the system (1) which must vanish (*21, 22*) is

$$|A - \kappa I| = \begin{vmatrix} -f_1^1 - \kappa & & \circ \\ & -f_2^1 - \kappa & \\ & & \ddots \\ \circ & & -f_n^1 - \kappa \end{vmatrix} = 0 \qquad \text{(A1)}$$

where A is for this case a diagonal matrix. The set of characteristic curves is therefore given by

$$\frac{dx}{ds} = -f^1[t(s), x, u] \qquad \text{(A2)}$$

where $x = (x_1, x_2, \ldots, x_n)'$, $f^1 = (f_1^1, f_2^1, \ldots, f_n^1)'$, and s is a parameter measured along the characteristics. The solution of (A2) for each component of x, say x_i, gives the ith characteristic curve. Along the characteristics the following relationship holds

$$\frac{dv}{ds} = f^o[t(s), x, v, u] \qquad \text{(A3)}$$

where $f^o = (f_1^o, f_2^o, \ldots, f_n^o)'$. The independent scalar variable x becomes a vector-valued dependent variable in (A3) through the relationship (A2). The solution of (A2) passing through a point (τ, ξ) of $I \times E$ is designated by

$$x = X(s; \tau, \xi)$$

where $X = (X_1, X_2, \ldots, X_n)$. The integration of (A2) and (A3) along the characteristics through (τ, ξ) gives for the canonical (diagonal) system in this work:

$$v_i(\tau, \xi) = v_i[t(o), X_i(o; \tau, \xi)] + \int_0^{s(\tau,\xi)} f_i^o\{t(s'), X_i(s'; \tau, \xi), v[t(s'), X_i(s'; \tau, \xi)], u[t(s'), X_i(s'; \tau, \xi)]\}ds' \qquad \text{(A4)}$$

and

$$X_i(s';\tau,\xi) = \xi - \int_{s(\tau,\xi)}^{s'} f_i^1\{t(s), X_i(s;\tau,\xi), u[t(s), X_i(s;\tau,\xi)]\}ds \quad (A5)$$

$$i = 1,2,\ldots,n.$$

The initial data $v_i[t(0), X_i(0;\tau,\xi)]$ $(i = 1, 2, \ldots, n)$ of (A4) are the values of $v_i(i = 1, 2, \ldots, n)$ on the initial curve at the points $[t(0), X_i(0;\tau,\xi)]$ $(i = 1, 2, \ldots, n)$ of $I \times E$ located by tracing back $X_i(s';\tau,\xi)$ from (τ,ξ) according to the rule (A5). The values of $v_i(i = 1, 2, \ldots, n)$ (or the value of v at any point of the initial curve) are given by the initial and boundary conditions in (4) and (5).

The characteristics for the vector $\mu(= \partial v/\partial x)$ of the system (1) can be derived from the general system of the first-order equations:

$$F(t, x, v, \lambda, \mu, u) = 0 \quad (A6)$$

where $F = (F_1, F_2, \ldots, F_n)'$ and $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda n)' = \left(\frac{\partial v_1}{\partial t}, \frac{\partial v_2}{\partial t}, \ldots, \frac{\partial v_n}{\partial t}\right)'$. Partial differentiation of (A6) with respect to x yields

$$F_x + (\nabla_v F)\,\mu + (\nabla_\lambda F)\frac{\partial \lambda}{\partial x} + (\nabla_\mu F)\frac{\partial \mu}{\partial x} = 0 \quad (A7)$$

In view of the identity $\partial\lambda/\partial x = \partial\mu/\partial t$, it follows that

$$(\nabla_\lambda F)\frac{\partial \mu}{\partial t} + (\nabla_\mu F)\frac{\partial \mu}{\partial x} = -F_x - (\nabla_v F)\mu \quad (A8)$$

For the system (1), $\nabla_\lambda F = I$, and $\nabla_\mu F$ is a diagonal matrix, so that the system (A8) is in canonical form. Thus the characteristic equations for μ are

$$\frac{dx}{ds} = \nabla_\mu F$$

$$\frac{d\mu}{ds} = -F_x - (\nabla_v F)\mu \quad (A9)$$

In connection with the system (1), the first equation in (A10) is identical to (A2) for $\nabla_\mu F = -f^1$. The second equation becomes explicitly

$$\frac{d\mu}{ds} = \phi(t(s), x, v, \mu, u) \quad (A11)$$

where $\phi = (\phi_1, \phi_2, \ldots, \phi_n)' = f_x + (\nabla_v f)\mu$.

On the line $t = t_o$ the initial values μ are, from Equation (4),

$$\mu(t_o, x) = \frac{d\Phi(x)}{dx} \quad (A12)$$

On the remaining segment of the initial curve (the line $x = x_o$) the values of the initial data of μ can be shown, from Equation (5), to be given by

$$\mu_i[t(o), x_o] = \frac{\left[f_i^o\{t(o), x_o, \Psi_i[t(o)], u[t(o), x_o]\} - \frac{d\Psi_i(t)}{dt}\Big|_{s=0}\frac{dt}{ds}\Big|_{s=0} \right]}{[-f_i^1\{t(o), x_o, u[t(o), x_o]\}]}$$

$$= \Pi_i\left\{t(o), x_o, \Psi_i[t(o)], \frac{d\Psi_i[t(o)]}{dt}, \frac{dt(o)}{ds}, u[t(o), x_o]\right\} \quad (A13)$$

$$i = 1, 2, \ldots, n$$

The expressions (A13) are the initial data $\mu_i[t(o), X_i(o;\tau,\xi)]$ of (A11) along the segment (the line $x = x_o$). An interesting observation is that the initial data are functions of u. Here $f_i^1\{t(o), x_o, u[t(o), x_o]\} \neq 0$. If $f_i^1 = 0$, the initial curve ($x = x_o$) is a characteristic. This situation is excluded from the development.

Part II. Jacketed Configuration with Step Feed Disturbance

The optimal unsteady control of a jacketed tubular reactor with and without heat generation due to chemical reaction in response to a step disturbance in the feed stream concentration is computed. More rigorous control action was found to be necessary for the case with heat generation than without heat generation. Convergence of the method was also investigated with fixed and free final time.

STATEMENT OF PROBLEM

The optimal control of a jacketed tubular reactor is considered, in which the first-order consecutive reactions $A \xrightarrow{k_1} B \xrightarrow{k_2} C$ take place, both with and without heat generation. The shell-side fluid temperature, u, is the control, where $u = u(t)$ only. It is desired to obtain the optimal control, $\hat{u}(t)$, with a known step disturbance in the inlet concentration, such that the cumulative yield of the desired product B at the exit $x = L$ during a specified time interval $[0, t_f]$ is maximized. The system of partial differential equations for the process is

$$\frac{\partial v_1}{\partial t} = -k_1 v_1 - \nu\frac{\partial v_1}{\partial x}$$

$$\frac{\partial v_2}{\partial t} = k_1 v_1 - k_2 v_2 - \nu\frac{\partial v_2}{\partial x}$$

$$\frac{\partial v_3}{\partial t} = b_1 k_1 v_1 + b_2 k_2 v_2 + \frac{1}{\alpha}[u(t) - v_3] - \nu\frac{\partial v_3}{\partial x} \quad (1)$$

subject to the initial and boundary conditions

$$v_i(0, x) = \Phi_i(x), \quad v_i(t, 0) = \Psi_i(t) \quad i = 1, 2, 3 \quad (2)$$

where $k_i = k_{io}\exp(-E_i/RT)$ $(i = 1, 2)$, and $\Phi_i(x)$ and $\Psi_i(t)$ are prescribed functions of x and of t, respectively. This problem has been discussed earlier (*1*), where now the functional to be minimized is

$$-\int_o^{t_f} v_2(t, L)\,dt \quad \text{or} \quad J = \int_o^{t_f}\int_o^L \left(-\frac{\partial v_2}{\partial x}\right)dx dt \quad (3)$$

The numerical values for the physical constants involved are taken from the work of Bilous and Amundson (*2*): $k_{10} = 0.535 \times 10^{11}$ min.$^{-1}$, $k_{20} = 0.461 \times 10^{18}$ min.$^{-1}$, $E_1 = 18$ kcal./mole, $E_2 = 30$ kcal./mole, and $R = 2.0$. In addition, let $\alpha = 3.0$, $L = 1$ unit of length and $\nu = 0.1$ unit of length/min., thus fixing the residence time at 10 min.

Negligible Heat Generation

Suppose that it has been previously determined, by a gradient method, that the optimal steady state, with negligible heat of reaction and for inlet concentrations $v_1(0)$ and $v_2(0)$ of 0.95 and 0.5 g. mole/liter, is obtained by choosing $v_e(0) = 349.3$°K. and $u = 335.5$°K. up to the time $t = 0^-$. At $t = 0^+$, the inlet concentrations suddenly change to $v_1(0) = 0.65$ and $v_2(0) = 0.35$ g. mole/liter, and the feed temperature to $v_3(0) = 300$°K. It is required to find the control $u(t)$ such that the cumulative yield $v_2(1)$ is maximized during the time interval $[0, t_f]$. The original and new optimal steady states are plotted in Fig-