1.0

1.1

1.25    1.4    1.6

2.8    2.5

3.2    2.2

3.6

4.0    2.0

1.8

MICROCOPY RESOLUTION TEST CHART

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

| REPORT DOCUMENTATION PAGE | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|

| 1. REPORT NUMBER | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
|---|---|---|
| AFOSR-TR- 80-0634 | AD-A089 845 | |

| 4. TITLE *(and Subtitle)* | 5. TYPE OF REPORT & PERIOD COVERED |
|---|---|
| COMPUTATIONAL DIFFICULTIES IN THE IDENTIFICATION AND OPTIMIZATION OF CONTROL SYSTEMS | Interim |
| | 6. PERFORMING ORG. REPORT NUMBER |

| 7. AUTHOR(s) | 8. CONTRACT OR GRANT NUMBER(s) |
|---|---|
| H. T. Banks | AFOSR 76-3092 |

| 9. PERFORMING ORGANIZATION NAME AND ADDRESS | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
|---|---|
| Brown University Division of Applied Mathematics Providence, RI 02912 | 61102F 2304/A1 |

| 11. CONTROLLING OFFICE NAME AND ADDRESS | 12. REPORT DATE |
|---|---|
| Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332 | January 1980 |
| | 13. NUMBER OF PAGES |
| | 21 |

| 14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)* | 15. SECURITY CLASS. *(of this report)* |
|---|---|
| | UNCLASSIFIED |
| | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

As more realistic models for resource management are developed, the need for efficient computational techniques for parameter estimation and optimal control involving nonlinear vector systems will grow. We discuss some of the difficulties associated with such computational schemes and also report on results available for identification and control of several classes of systems which are of increasing importance in a number of areas of applications.

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73

**LEVEL**

③

# COMPUTATIONAL DIFFICULTIES IN THE IDENTIFICATION AND OPTIMIZATION OF CONTROL SYSTEMS*

BY

H. T. BANKS

LEFSCHETZ CENTER FOR DYNAMICAL SYSTEMS
DIVISION OF APPLIED MATHEMATICS
BROWN UNIVERSITY
PROVIDENCE, RHODE ISLAND    02912

JAN          80

# COMPUTATIONAL DIFFICULTIES IN THE IDENTIFICATION AND OPTIMIZATION OF CONTROL SYSTEMS

H. T. BANKS[+]

Lefschetz Center for Dynamical Systems
Division of Applied Mathematics
Brown University
Providence, Rhode Island 02912

## ABSTRACT

As more realistic models for resource management are developed, the need for efficient computational techniques for parameter estimation and optimal control involving nonlinear vector systems will grow. We discuss some of the difficulties associated with such computational schemes and also report on results available for identification and control of several classes of systems which are of increasing importance in a number of areas of applications.

## 1. INTRODUCTION

In this presentation we consider computational methods for <u>vector</u> dynamical systems of the type

$$\dot{x}(t) = f(t, x(t), u(t), \beta), \qquad 0 \le t \le T,$$
$$x(0) = x_0 \tag{1}$$

or

$$x(i+1) = F(i, x(i), u(i), \beta), \quad i = 0, 1, 2, \ldots, M - 1,$$
$$x(0) = x_0, \tag{2}$$

with $x = (x_1, \ldots, x_n) \in R^n$, $u = (u_1, \ldots, u_m) \in R^m$, and $\beta = (\beta_1, \ldots, \beta_\nu) \in R^\nu$.

The methods discussed are for the related problems of parameter estimation (identification) and optimization (optimal control) in "state" models which
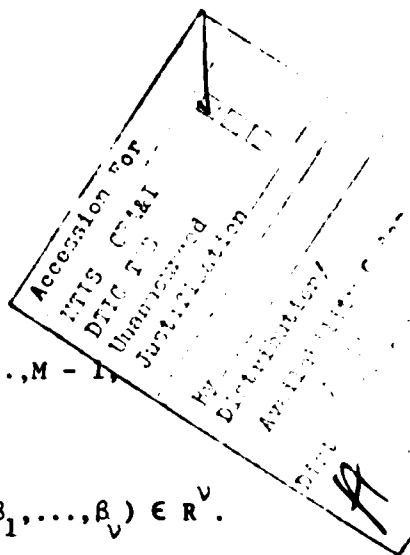
are assumed to be based on accepted (or hypothesized) mechanisms for growth, competition, harvesting, etc. in resource management. As one considers increasingly realistic models that include multi-species effects, age structure, multiple trophic levels, etc., (e.g. see (May et al. 1979), (Clark 1976), (Clark 1976a)) the analytic techniques employed in early studies of scalar or two dimensional vector models (see (Clark 1976)) will undoubtedly prove inadequate and one should find computational procedures such as those outlined below quite useful even in investigations of qualitative features of the more sophisticated models.

Our discussions here will be divided into two main sections. First, we give a brief review of standard ideas and techniques for identification and optimization problems involving ordinary differential equation models such as (1). Least squares techniques entailing such standard iterative techniques as gradient and conjugate gradient procedures will be outlined for the parameter identification problem for (1). Brief mention of maximum likelihood estimator ideas will be made. We next turn to necessary conditions for optimal control problems governed by (1). Our emphasis here will be on the celebrated Pontryagin maximum principle and the resulting two point boundary value problem that must be solved.

The second part of our presentation will be devoted to an indication of recent results (the development of which was mostly motivated by needs in other areas of applications such as aerodynamics, biochemistry and radiation biology) for special systems that we feel should be of use in resource management studies. We discuss the difficulties associated with identification and control of models with delays (hereditary systems) and offer tested ideas to alleviate some of these difficulties. Finally, procedures for optimization problems with underlying discrete systems such as (2) will be briefly outlined. Special problems in which the time intervals $\Delta t_i$ between the control actions $u(i)$ are themselves part of

the optimization parameters appear to be of interest in certain terrestrial management endeavors and these will be discussed.

## 2. PARAMETER IDENTIFICATION

Consider the equation (1) in which the vector function u represents input to the system, $\beta$ is a vector of unknown parameters (e.g., time constants, growth rates, etc.) and we are able to observe the "output" vector $y(t) = Cx(t)$. Here $y = (y_1, y_2, \ldots, y_p) \in R^p$ and C is a known $p \times n$ matrix. In practice, we make observations $\{\hat{y}^1, \hat{y}^2, \ldots, \hat{y}^K\}$ of the output $y$ at times $t_1 < t_2 < \ldots < t_K$ (with or without perturbing the system via the input u). If we denote by $x(\cdot\,;\beta)$ the corresponding solution of (1), $\hat{y}^i$ is then an observation for $y(t_i;\beta) = Cx(t_i;\beta)$ and a <u>Least Squares Estimate</u> (LSE) of $\beta$ is a solution to the problem of choosing $\beta \in \mathscr{B} \subset R^\nu$ so as to minimize

$$(3) \qquad J(\beta) = \sum_{i=1}^{K} |\hat{y}^i - y(t_i;\beta)|^2.$$

Here $\mathscr{B}$ is some given parameter set determined by physical, biological, economic, chemical, etc., restraints on the system.

For nonlinear vector systems, the problem of finding a $\beta^* \in \mathscr{B}$ to minimize (3) is not, in general, solvable by analytic methods and some type of iterative numerical scheme is called for to produce estimates $\{\beta^j\}$ that hopefully converge to a "best fit" parameter $\beta^*$. Such a procedure can be formally stated as:

<u>(i)</u> Guess an initial estimate $\beta^0$. (To make a good initial guess is important and one usually relies on his "knowledge" of the parameters although "apparent" values for parameters—which may have little relation to actual physical or biological limits—often play an important role in biological and ecological models!)

(ii) Given an estimate $\beta^j$, generate a next estimate $\beta^{j+1}$ by some formula $\beta^{j+1} = \mathscr{F}(\beta^j)$. (For example, many popular methods are based on the iterate formula $\beta^{j+1} = \beta^j + \alpha_j d^j$ where $d^j$ is a "direction" and $\alpha_j$ is a "step-size" parameter.)

(iii) If $|\beta^{j+1} - \beta^j|$ is less than some given error tolerance, stop the procedure, accepting $\beta^{j+1}$ as our value for $\beta^*$. Otherwise, set $\beta^j = \beta^{j+1}$ and return to step (ii) to generate a next estimate.

Among the most popular methods for the choice in (ii) above are the so-called descent methods (see Chapter X of (Banks and Palatt 1975)) which include the gradient and conjugate-gradient techniques as special cases. Letting $\nabla J = \dfrac{\partial J}{\partial \beta}$ denote the gradient (partial derivative) of $J$ with respect to $\beta$, the <u>gradient</u> method employs the iterative formula

$$\beta^{j+1} = \beta^j + \alpha_j \left[ -\nabla J(\beta^j) \right]$$

where $\alpha_j$ is determined by a one-dimensional search on $\alpha$ in $\beta^j - \alpha\nabla J(\beta^j)$ to minimize $J$. The <u>conjugate-gradient</u> or <u>conjugate-directions</u> methods modify this choice of directions $d^j$ (one no longer chooses $d^j = -\nabla J(\beta^j)$, but a choice related to this direction is made) in an attempt to insure that "independent" directions are chosen on subsequent steps in the algorithm. This is especially profitable in problems where narrow, steep "valleys" are present in the surface J as a function of the parameters. Among the more popular conjugate directions methods are those of Fletcher-Reeves and Daniel (Ortega and Rheinboldt 1970). In practice, hybrid methods combining gradient (G) and conjugate-gradient (CG) steps in some pattern (e.g., G, CG, CG, CG, G, CG, CG, CG, G, CG,...) often are found to perform better than an algorithm employing a single type of formula at each step.

A somewhat different approach to the least squares fit to data utilizes the method of quasilinearization. We shall not discuss these ideas here, but interested parties may consult (Banks and Groome 1973), (Bellman and Kalaba 1965), or Chapter X of (Banks and Palatt 1975).

An alternative to the LSE as formulated above is the <u>Maximum Likelihood</u> <u>Estimator</u> (MLE) for $\beta$. This procedure is based on the following considerations. In many problems, the observations $\{\hat{y}^i\}$ are corrupted by noise so we might write $\hat{y}^i = y(t_i; \beta^*) + Z_i$ where $Z_i$ is the random measurement noise at time $t_i$ and $\beta^*$ is the true parameter value. If one assumes that $Z_1, \ldots, Z_K$ are independent random variables with identical probability density functions $h$ (an assumption that is often <u>not</u> true in practice), then the joint probability density function is given by $\tilde{h}(z_1, z_2, \ldots, z_K) = \prod_{i=1}^{K} h(z_i)$. Loosely speaking, the function $\tilde{h}$ has its maximum at those values of $(z_1, \ldots, z_K)$ that are most likely to occur. Therefore, we might devise a procedure for estimating $\beta^*$ on the basis that the observed values of $Z_i = \hat{y}^i - y(t_i; \beta^*)$ correspond to those that are most likely to occur, i.e., $\max \tilde{h}(z_1, \ldots, z_K) = h(Z_1, \ldots, Z_K)$. If we thus consider the function $F(\beta) \equiv \tilde{h}(\hat{y}^1 - y(t_1; \beta), \ldots, \hat{y}^K - y(t_K; \beta))$, we might expect this function to be maximized at the value $\beta = \beta^*$. An MLE for $\beta^*$ is then defined to be a value $\bar{\beta}$ which yields a maximum for $F(\beta)$. Instead of maximizing $F$, we define the <u>likelihood</u> function

$$L(\beta) = \ell n \, F(\beta) = \sum_{i=1}^{K} \ell n \, h(\hat{y}^i - y(t_i; \beta))$$

and equivalently seek a maximum for this function. This usually reduces to employing a procedure to determine $\bar{\beta}$ so that $\frac{\partial L}{\partial \beta}(\bar{\beta}) = 0$. A more detailed discussion of MLE's can be found in almost any standard text on statistics or

estimation Deutsch [4] , Good and Gaskins [4] , Cramer [4] , Lindgren [4] .

We note that in certain cases the MLE reduces to the LSE (Deutsch 1965, p. 136). For example, assume that $Z$ is scalar, Gaussian with mean zero and variance $\sigma^2$. Then $h(z) = (1/\sqrt{2\pi} \, \sigma) \exp(-z^2/2\sigma^2)$ and

$$
L(\beta) = \sum_{i=1}^{K} \ln \frac{e^{-(\hat{y}^i - y(t_i;\beta))^2/2\sigma^2}}{\sqrt{2\pi} \, \sigma}
$$

$$
= \sum_{i=1}^{K} \left[ -\ln(\sqrt{2\pi} \, \sigma) - \frac{[\hat{y}^i - y(t_i;\beta)]^2}{2\sigma^2} \right]
$$

$$
= -K\ln(\sqrt{2\pi} \, \sigma) - \sum_{i=1}^{K} \frac{1}{2\sigma^2} [\hat{y}^i - y(t_i;\beta)]^2 .
$$

Therefore maximizing $L$ is equivalent to minimizing the least squares function $J$ defined in (3).

## 3. OPTIMAL CONTROL

For this class of problems we assume that the parameters $\beta$ in (1) are known and define $g(t,x,u) \equiv f(t,x,u,\beta)$. In addition to the initial conditions in (1) we impose terminal or target conditions $x(T) \in \mathscr{T}_1$ where $\mathscr{T}_1$ is some given desired (smooth) subset of $R^n$. The control system is thus defined by

$$
\dot{x}(t) = g(t,x(t),u(t)), \qquad 0 \leq t \leq T,
$$

(4) $\qquad x(0) = x_0$

$$
x(T) \in \mathscr{T}_1
$$

and control functions  u: [0,T] → U, U  a given restraint set, are to be

chosen from a prescribed set  $\mathcal{U}$  of admissible controls.  The optimal control

problem consists of choosing  u ∈ $\mathcal{U}$  so as to minimize

$$I(u) = g^0(x(T)) + \int_0^T f^0(s,x(s),u(s))ds$$

subject to (4).  Here  $g^0, f^0$  are given "payoff" or "cost" functions.

The necessary conditions for  $(x^*,u^*)$  to be a solution of this problem

that we present here are the first order conditions that are analogues to the

conditions  $F'(\zeta) = 0$  employed in calculus when minimizing a scalar function

F  of one variable.  More correctly, they are, roughly speaking, the function

space analogue to the well-known Lagrange Multiplier Rule for constrained

minimization in multivariable calculus problems (see (Luenberger 1969), (Kirk 1970),

(Hestenes 1975), (Bryson and Ho 1969)).  Numerical computations are almost

always essential for solving these necessary conditions in the case of vector

systems that realistically model biological or engineering phenomena.  While

the method of dynamic programming has received wide spread publicity, we shall

not discuss it here since for vector systems of dimension greater than two it is

often very difficult, if not impossible, to implement this method.

For  $\psi_0 \in R^1$  and  $\psi = (\psi_1,\ldots,\psi_n) \in R^n$, we define the "Hamiltonian"

function (scalar-valued) by

$$H(t,x,u,\psi_0,\psi) = \psi_0 f^0(t,x,u) + \psi \cdot g(t,x,u).$$

The Pontryagin Maximum Principle may then be stated (for a careful statement

and proof, see (Berkovitz 1974) or (Fleming and Rishel 1975)) as follows:

__PMP__: If $(x^*, u^*)$ is a solution to the above problem, there exist a nontrivial $n+1$ vector function $t \to (\lambda_0, \lambda(t)) = (\lambda_0, \lambda_1(t), \ldots, \lambda_n(t))$ such that

(a) $\lambda_0$ = constant $\leq 0$,

$$\dot{\lambda}(t) = -\frac{\partial H}{\partial x}(t, x^*(t), u^*(t), \lambda_0, \lambda(t)), \qquad 0 \leq t \leq T,$$

(b) $\lambda(T)$ is orthogonal to $\mathcal{T}_1$ at $x^*(T)$,

(c) $H(t, x^*(t), u^*(t), \lambda_0, \lambda(t)) = \max_{v \in U} H(t, x^*(t), v, \lambda_0, \lambda(t))$ for $0 \leq t \leq T$.

The boundary conditions on $\lambda$ given in (b) are called transversality conditions and the condition in (c) is of course the "maximum condition" that the optimal pair $(x^*, u^*)$ must satisfy. The equations in (a), which can be equivalently written

$$\dot{\lambda}(t) = -\lambda_0 \frac{\partial f^0}{\partial x}(t, x^*(t), u^*(t)) - \lambda(t) \frac{\partial g}{\partial x}(t, x^*(t), u^*(t)),$$

are called the "costate" or "adjoint" or "multiplier" equations. These equations, with the boundary conditions from (b), are coupled with the system (4) and together they form a two-point boundary value problem (TPBVP) for a 2n-vector system (i.e. (4), (a), (b) must be solved simultaneously). Note, however, that these equations involve the (unknown but sought after) control function $u^*$ which must be determined from the condition (c). But condition (c) involves the functions $x^*$ and $\lambda$ which are to be determined from (4), (a), and (b). Thus, conditions (a), (b), (c) taken with (4) constitute a TPBVP with a coupled maximization condition. There is, in general, no hope of solving this analytically and some type of computational scheme is required.

Just as in the case of the identification problems of section 2, one can

develop iterative schemes based on gradient, conjugate-gradient, etc. ideas to generate a sequence of pairs $(x^j, u^j)$ that (hopefully) converge to $(x^*, u^*)$. However, in this case the iterations are made in function space. The gradient of I with respect to u is an operator with kernel $\frac{\partial H}{\partial u}$ ; that is,

$$\nabla I(u; \delta u) \doteq \int_0^T \frac{\partial H}{\partial u} \delta u \ dt$$

and the descent type procedures will thus use $\frac{\partial H}{\partial u}$ as the gradient "direction". For example, a procedure based on the gradient method is given by the steps:

[1] Choose an initial estimate $u^0$;

[2] Given the $j^{th}$ estimate $u^j$, generate the next estimate $u^{j+1}$ by:

   (i) Use $u^j$ in $\dot{x} = g(t,x,u)$, $x(0) = x_0$, to compute $x^j$.

   (ii) Use $u^j, x^j$ with $\dot{\lambda} = \frac{-\partial H}{\partial x}(t,x,u,\lambda_0,\lambda)$, $\lambda(T) \perp \mathscr{T}_1$ at $x^j(T)$ to compute $\lambda^j$.

   (iii) Compute $D^j = \frac{\partial H}{\partial u}(t,x^j,u^j,\lambda_0,\lambda^j)$.

   (iv) Put $u^{j+1} = u^j + \alpha_j D^j$ where $\alpha_j$ is determined by a one-dimensional search to minimize $I(u)$;

[3] If $u^{j+1} \approx u^j$, terminate the procedure, accepting $u^{j+1}$ as the estimate for $u^*$; otherwise, set $u^j = u^{j+1}$ and return to [2] above.

There are many variations on the iterative procedure outlined here, some of which involve different choices for the directions $D^j$. For further discussions, one can consult (Kirk 1970), (Polak 1973), (Banks and Palatt 1975). An excellent survey of other methods can also be found in (Polak 1973).

## 4. SYSTEMS WITH DELAYS

Most investigations of resource management are, by necessity, based on models for "population" behavior (resource increase, utilization, etc.). Due to the complexity in ecological chains, it is often desirable to consider multiple trophic levels (e.g., vegetation-herbivore-carnivore systems) in the models used. Other factors that play an important role are age structure, environment recovery, and delay in recruitment. All of these considerations lead, in any realistic modeling effort (May et al. 1979, p. 275), (Cushing 1977), (Clark 1976), (May 1973, p. 73), (May 1976, p. 6) to systems with delays. Such systems have played an increasingly important role in engineering and other areas of applications and in recent years efforts to develop mathematical techniques to aid in investigating delay systems have grown. We report on some of the recent results obtained from these efforts in hopes that these techniques will also prove of value to investigators of models for renewable resources.

Consider for the moment the delay system

$$
(5) \quad
\begin{aligned}
\dot{x}(t) &= f(t,x(t),x(t-r),u(t),\beta), & 0 \leq t \leq T \\
x(\theta) &= \phi(\theta), & -r \leq \theta \leq 0,
\end{aligned}
$$

which, even though we include only a discrete delay term, is an infinite dimensional "state" system (similar to a partial differential equation). This poses immediate added difficulties for the identification and control problems discussed in Sections 2 and 3 above. However, in the case of certain parameter estimation problems, there are even more serious questions that must be entertained. Often, in addition to the parameters $\beta$ in (5), one also wishes to estimate the

delay $r$ so that the parameter identification problem consists of estimating $\gamma = (\beta,r)$ or, in the case of multiple delay systems, $\gamma = (\beta,r_1,\ldots,r_\mu)$. A moment's reflection (see (3) and the related discussions) will reveal that in order to use the standard techniques discussed above, one needs to be able to compute $\frac{\partial x}{\partial \beta}(t_i;\beta)$ and, in this case, $\frac{\partial x}{\partial r}(t_i;r)$. However, this derivative does not, in general, even exist! Consider the following example:

$$\dot{x}(t) = x(t-r), \qquad t > 0,$$

$$x(\theta) = \begin{cases} 0 & -1 \leq \theta \leq 0 \\ 1 & \theta \leq -1. \end{cases}$$

For $r \geq -1$, the solution is given by $x(t;r) = 0$ for $t \geq 0$. For $r < -1$, say $r = -1 - \varepsilon$, we find $x(t;r) = t$ for $0 \leq t \leq \varepsilon$, and $x(t;r) = \varepsilon$ for $\varepsilon \leq t \leq 1 + \varepsilon$. It is then very easy to show that for $0 < t_i < 1$, $\frac{\partial x}{\partial r}(t_i;-1)$ does not exist. The techniques we summarize here will overcome this difficulty as well as those presented by the infinite-dimensional state aspects of (5).

We sketch the ideas for the simplest linear delay system. Results are available for the most general linear systems, including those with distributed delay terms (e.g. $\int_{t-r}^{t} k(\theta)x(\theta)d\theta$) - see (Banks and Burns 1978), (Banks et al. 1979), (Banks and Kappel 1979), (Banks et al. 1979b)-as well as for a rather general class of nonlinear systems (Banks 1979), (Banks 1980).

Consider the simplest linear delay system

$$\dot{x}(t) = A_0(\beta)x(t) + A_1(\beta)x(t-r) + Bu(t) \qquad t > 0,$$

(6)
$$x(\theta) = \phi(\theta), \qquad -r \leq \theta \leq 0,$$

$$y(t) = Cx(t).$$

Here $x \in R^n$ and $A_0, A_1$ are $n \times n$ matrix functions that are assumed to depend on the parameters $\beta$ in a smooth manner. Let $\gamma = (\beta, r)$ with $\Gamma$ some given admissible parameter set. We wish to choose $\gamma \in \Gamma$ so as to minimize

$$J(\gamma) = \sum_{i=1}^{K} |\hat{y}^i - y(t_i; \gamma)|^2$$

where the $\hat{y}^i$ are observations for $y(t_i; \gamma) = Cx(t_i; \gamma)$. The methods we propose entail approximating (6) by high order differential equations in which the parameters $\gamma$ appear smoothly so that we may apply standard techniques. That is, we have an approximating identification problem:

Find $\gamma^N \in \Gamma$ so as to minimize

$$J^N(\gamma) = \sum_{i=1}^{K} |\hat{y}^i - C^N w^N(t_i; \gamma)|^2$$

where

(7)
$$\dot{w}^N(t) = A^N(\beta, r)w^N(t) + B^N u(t)$$
$$w^N(0) = w_0^N(\varphi).$$

Here (7) is an approximating system for (6) that is finite dimensional $(w^N \in R^{\rho(N,n)})$ with the dimension $\rho$ depending on the index of approximation N and the dimension n of the original system (6). One can, for the schemes we have developed, argue that $\gamma^N$ converges (as $N \to \infty$) rapidly to a solution $\gamma^*$ of the original estimation problem. For details see (Banks et al. 1979b). We shall not discuss this here since the arguments are quite technical. Essentially, one treats (6) as an abstract system in a Hilbert space and employs classical Ritz type ideas: Approximate the problem on finite-dimensional

subspaces. Convergence properties and error estimates can then be obtained using approximation results from linear semigroup theory.

These approximation ideas (for both the parameter estimation and optimal control problems) have been developed to date for two specific classes of schemes:

(I) The "averaging" approximations (Banks and Burns 1978): These approximations are based on approximations of states $\psi \in L_2^n(-r,0)$ by step functions $\psi^N = \sum_{j=1}^{N} a_j \chi_j$ where $\chi_j$ is the characteristic function for $I_j^N = [\frac{-jr}{N}, \frac{-(j-1)r}{N}]$, $j = 1,2,\ldots,N$, and $a_j$ is the integral average of $\psi$ over the subinterval $I_j^N$; i.e.

$$a_j = \frac{N}{r} \int_{I_j^N} \psi(s)ds.$$

The matrix $A^N(\beta,r)$ in (7) turns out to be very simple to use in computations, being given by



Here $\mathcal{I}_n$ is the $n \times n$ identity matrix.

(II) Spline based approximations (Banks and Kappel 1979): These approximations are based on best $L_2$ approximations of states in subspaces consisting of spans of standard spline basis elements. The theory

has been developed for arbitrary order splines and has been tested numerically for piecewise linear and cubic elements. The matrix $A^N$ in these cases is not as simple as in the case of the averaging approximations, but the methods still are quite easily implemented.

We have carried out extensive testing of these methods as applied to identification and optimization problems. These findings (Banks et al. 1979a) indicate that our theoretical estimates are supported by computational efficacy of the methods. We present here a sample of the numerical findings in a parameter estimation problem. The example is typical of tubular column reactor problems discussed in (Banks et al. 1979b) in which one designs experiments to determine a transport coefficient $\beta$ for the product of an enzyme catalyzed reaction taking place in the column. Enzyme pellets are packed in the column and it is also desired to estimate the "diffusion delay" $r$ associated with diffusion from the interior of the pellet to the exterior flow region of the column. Mathematically, we wish to estimate $\beta$ and $r$ in the system

$$\dot{z}_1(t) = E_0(\beta)z_1(t) + E_1(\beta)z_1(t-r) + G$$
$$0 \le t \le 8,$$
$$\dot{z}_i(t) = E_0(\beta)z_i(t) + E_1(\beta)z_i(t-r) + E_2 z_{i-1}(t-2)$$
$$i = 2,3,4,$$
$$z_1(\theta) = (1,0)^T$$
$$-2 \le \theta \le 0,$$
$$z_i(\theta) = (0,0)^T$$

where

$$E_0(\beta) = \begin{bmatrix} -3-\beta & 0 \\ 2\beta & 0 \end{bmatrix} \quad E_1(\beta) = \begin{bmatrix} 0 & \beta \\ 0 & -2\beta \end{bmatrix} \quad E_2 = \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix} \quad G = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

In this example we have each vector $z$ is 2-dimensional so that the example is an 8 vector system $x = (z_1, z_2, z_3, z_4)^T$. It is assumed that only the first component of $z_4$ can be observed. Data was generated using the "true" values $\beta^* = -3$ and $r^* = 2$. A least squares criterion as in (3) was used along with a standard iterative package (an IMSL package). The approximate problem for (7) with $N = 24$ was investigated for both the averaging and linear spline schemes. Start-up values in the iteration were $\gamma^0 = (\beta^0, r^0) = (-4, 3)$. For the averaging approximations, satisfactory convergence was not obtained, while the iterates for the spline scheme are given in Table 1. (This illustrates dramatically one difference between theory and practice-in theory both schemes should produce good approximations-only the spline scheme does in this example.)

| $\beta$ | $r$ |
|---------|------|
| -4.0 | 3.0 |
| -3.717 | 2.341 |
| -3.535 | 1.954 |
| -3.436 | 1.929 |
| -3.172 | 1.903 |
| -3.109 | 1.921 |
| -3.069 | 1.950 |
| -3.039 | 1.987 |
| -3.004 | 1.994 |

TABLE 1

Use of these approximation ideas in optimization problems as described in Section 3 has also been investigated both theoretically and numerically. Briefly, if one wishes to minimize $I(u)$ (see Section 3) over $\mathcal{U}$ subject to (6), one again approximates by a problem for a high dimension ordinary differential equation (7). One then seeks to find a $u^N \in \mathcal{U}$ that minimizes

$$I^N(u) = g_N^0(w^N(T)) + \int_0^T f_N^0(s,w^N(s),u(s))ds$$

subject to (7). Here $g_N^0, f_N^0$ are appropriately chosen approximations for $g^0, f^0$.
Again, one can argue (see Banks and Burns 1978, Banks and Kappel 1979, Banks
et al. 1979) that solutions $(x^N, u^N)$ of the approximating problems converge
as $N \to \infty$ to $(x^*, u^*)$, a solution of the original problem, for a large class
of payoff functions $g^0, f^0$. Extensive numerical results can be found in
(Banks et al. 1979a), while theoretical results for nonlinear systems can be
found in (Banks 1979, Banks 1980).

## 5. DISCRETE SYSTEMS

Parameter estimation and optimal control problems for discrete systems
(2) can be formulated in much the same way as indicated above for the con-
tinuous systems (1). We outline briefly the optimization problem and results
because we feel that these have direct applications in resource management and
because the associated necessary conditions are not simple analogues to those
for the continuous system problem in the case of many nonlinear systems.

Assuming that the parameter $\beta$ is known, we seek a set of values
$u = \{u(0), u(1), \ldots, u(M-1)\}$ from a given constraint set $U \subset R^m$ so as to
minimize

$$I(u) = g^0(x(M)) + \sum_{i=0}^{M-1} f^0(i,x(i),u(i))$$

subject to (2). The necessary conditions (see PMP above) are again defined in
terms of a Hamiltonian function

$$H_i(v) = f^0(i,x(i),v) + \lambda(i+1) \cdot F(i,x(i),v,\beta), \quad i = 0,1,\ldots,M-1,$$

and an adjoint or costate equation

$$\lambda(i) = -\frac{\partial H_i}{\partial x(i)}(u^*(i)), \quad i = 1,\ldots,M-1, \tag{8}$$

with appropriate boundary conditions (see (Almquist and Banks 1976) for a more precise statement). However, the global maximum condition (the analogue of (c) in PMP above)

$$H_i(u^*(i)) = \max_{v \in U} H_i(v)$$

is valid for nonlinear system problems only under very strong convexity conditions on F and U (which are not usually satisfied in applications). Rather there is a <u>local</u> maximum condition

$$\frac{\partial H_i}{\partial u(i)}(u^*(i))\cdot[u^*(i)-v] \geq 0, \quad v \in U \tag{9}$$

or, in some cases,

$$\frac{\partial H_i}{\partial u(i)}(u^*(i)) = 0, \tag{10}$$

that is valid for most nonlinear system problems of interest. For a precise statement of the global and local conditions one can consult Chapter VII, Section 4 of (Neustadt 1976). The last chapter (Notes and Historical Comments) of (Neustadt 1976) contains an account of the interesting historical development (marked with some confusion and incorrect claims) of necessary conditions for discrete systems control problems along with a rather complete bibliography. Computational techniques based on (9) or (10) that are modifications of the

iterative schemes (e.g. descent methods) discussed above have been developed
and employed on a wide variety of discrete system problems. For examples
see (Almquist and Banks 1976), (Bryson and Ho 1969), (Kirk 1970).

In the usual applications of discrete system theory in engineering and
in some biological models-see (Almquist and Banks 1976), (Clark 1976a)-one
has an underlying continuous time process in which "discontinuous" events or
phenomena (e.g. treatments, harvests, etc.) occur. That is, one has a process
evolving in time with major events (controls $u(i)$) occurring at time intervals
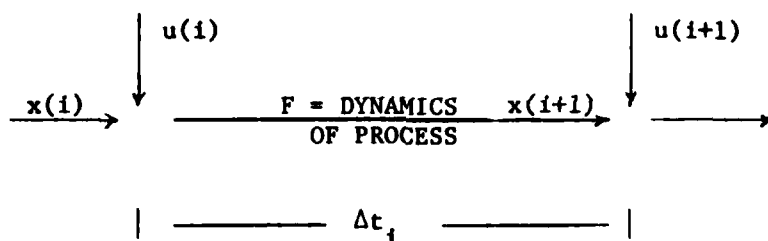$\Delta t_0, \Delta t_1, \ldots, \Delta t_{M-1}$ apart as depicted in Figure 1.

$$
\begin{array}{c}
\quad\quad\quad \downarrow u(i) \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \downarrow u(i+1) \\[4pt]
\xrightarrow{\;x(i)\;} \quad \underline{\begin{array}{c} F = \text{DYNAMICS} \quad x(i+1) \\ \text{OF PROCESS} \end{array}} \xrightarrow{\phantom{xxxx}} \\[8pt]
|\text{———————}\; \Delta t_i \;\text{———————}|
\end{array}
$$

FIGURE 1

In the optimization problem formulated above, it is assumed that $\Delta t_i$,
the time between control actions $u(i)$ and $u(i+1)$, is a fixed parameter.
However, in our efforts to extend the work of (Almquist and Banks 1976) on
fractionated therapy for tumors, we have been obliged to consider optimization
problems for discrete systems in which the times $\Delta t_i$ are additional para-
meters over which one wishes to optimize. That is, in addition to dealing
with variable levels of discrete control actions, one wishes to allow variable
times between the control actions. We have recently tested computational ideas

such as those described in (Almquist and Banks 1976) for problems where one
seeks optimal "controls" $w(i) = (u(i), \Delta t_i)$ for problems with systems of the form
$x(i+1) = G(i, x(i), w(i))$. Our preliminary findings (details will appear in a
manuscript in the future) suggest that efficient methods can be readily
developed for discrete time variable-control-time problems via careful
modification and extension of existing schemes.

We believe that such techniques could be of use in resource management
studies in which one considers variable times between harvesting, etc. For
example, they could be important in forestry management problems since it has
been observed (Aber et al. 1978, Aber et al. 1979) that rotation length (time
between control actions) is perhaps in some cases a more important parameter
than harvesting intensity (level of control action). Simulation and
optimization studies of various cutting regimes must, due to the complexity of
the underlying ecological models (Botkin et al. 1972), ultimately involve
computational procedures for vector models.

REFERENCES

Aber, J.D., Botkin, D.B., and Melillo, J.M. 1978. Predicting the Effects of
Different Harvesting Regimes on Forest Floor Dynamics in Northern Hardwoods,
Can. J. Forest Res., 8, 306-315.

Aber, J.D., Botkin, D.B., and Melillo, J.M. 1979. Predicting the Effects of
Different Harvesting Regimes on Productivity and Yield in Northern Hardwoods,
Can. J. Forest Res., 9, 10-14.

Almquist, K.J., and Banks, H.T. 1976. A Theoretical and Computational Method
for Determining Optimal Treatment Schedules in Fractionated Radiation Therapy,
Math. Biosci., 29, 159-179.

Banks, H.T. 1979. Approximation of Nonlinear Functional Differential Equation
Control Systems, J. Optimization Theory Appl., 29, 383-408.

Banks, H.T. 1980. Identification of Nonlinear Delay Systems, to appear.

Banks, H.T., and Burns, J.A. 1978. Hereditary Control Problems: Numerical
Methods Based on Averaging Approximations, SIAM J. Control and Optimization,
16, 169-208.

Banks, H.T., Burns, J.A., and Cliff, E.M. 1979. Spline-based Approximation
Methods for Control and Identification of Hereditary Systems, in International
Symposium on Systems Optimization and Analysis, A. Bensoussan and J.L. Lions,
eds., Lecture Notes in Control and Info. Sci., Vol. 14, Springer, Heidelberg,
314-320.

Banks, H.T., Burns, J.A., and Cliff, E.M. 1979a. A Comparison of Numerical
Methods for Identification and Optimization Problems Involving Control Systems
with Delays, Brown University LCDS Tech. Rep. 79-7, Providence.

Banks, H.T., Burns, J.A., and Cliff, E.M. 1979b. Parameter Estimation and
Identification for Systems with Delays, SIAM J. Control and Optimization,
submitted.

Banks, H.T., and Groome, G.M.,Jr. 1973. Convergence Theorems for Parameter
Estimation by Quasilinearization, J. Math. Anal. Appl., 42, 91-109.

Banks, H.T, and Kappel, F. 1979. Spline Approximations for Functional
Differential Equations, J. Diff. Eq., 34, 496-522.

Banks, H.T., and Pallatt, P.J. 1975. Mathematical Modeling in the Biological
Sciences, Brown University LCDS Lec. Notes 75-1, Providence.

Berkovitz, L.D. 1974. Optimal Control Theory, Springer-Verlag, New York.

Botkin, D.B., Janak, J.F., and Wallis, J.R. 1972. Some Ecological Consequences
of a Computer Model of Forest Growth, J. Ecology, 60, 849-872.

Bryson, A.E., and Ho, Y.C. 1969. Applied Optimal Control, Blaisdell Publ., Waltham.

Bellman, R., and Kalaba, R. 1965. Quasilinearization and Nonlinear Boundary-Value Problems, American Elsevier, New York.

Clark, C.W. 1976. Mathematical Bioeconomics, John Wiley, New York.

Clark, C.W. 1976a. A Delayed Recruitment Model of Population Dynamics with an Application to Baleen Whale Populations, J. Math. Biol., 3, 381-391.

Cramér, H. 1946. Mathematical Methods of Statistics, Princeton Univ. Press, Princeton.

Cushing, J.M. 1977. Integrodifferential Equations and Delay Models in Population Dynamics, Springer Lecture Notes in Biomath., 20, New York.

Deutsch, R. 1965. Estimation Theory, Prentice-Hall, Englewood Cliffs.

Fleming, W.H., and Rishel, R.W. 1975. Deterministic and Stochastic Optimal Control, Springer-Verlag, New York.

Hestenes, M.R. 1975. Optimization Theory, John Wiley, New York.

Kirk, D.E. 1970. Optimal Control Theory, Prentice-Hall, Englewood Cliffs.

Lindgren, B.W. 1962. Statistical Theory, Macmillan, New York.

Luenberger, D.G. 1969. Optimization by Vector Space Methods, John Wiley, New York.

May, R.M. 1973. Stability and Complexity in Model Ecosystems, Princeton Univ. Press, Princeton.

May, R.M. 1976. Theoretical Ecology: Principles and Applications, Blackwell, Oxford.

May, R.M., Beddington, J.R., Clark, C.W., Holt, S.J., and Laws, R.M. 1979. Management of Multispecies Fisheries, Science, 205, 267-277.

Mood, A.M., and Graybill, F.A. 1963. Introduction to the Theory of Statistics, McGraw-Hill, New York.

Neustadt, L.W. 1976. Optimization, Princeton Univ. Press, Princeton.

Ortega, J.M., and Rheinboldt, W.C. 1970. Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York.

Polak, E. 1973. An Historical Survey of Computational Methods in Optimal Control, SIAM Rev., 15, 553-584.