Some concepts of modern control theory

Riordon, J. S.

Publisher's version / Version de l'éditeur:
https://doi.org/10.4224/21277227
Report (National Research Council of Canada. Radio and Electrical Engineering Division : ERB), 1968-08

ERB-792

UNCLASSIFIED

NATIONAL RESEARCH COUNCIL OF CANADA

RADIO AND ELECTRICAL ENGINEERING DIVISION

SOME CONCEPTS OF MODERN CONTROL THEORY

- J. S. RIORDON -

OTTAWA

AUGUST 1968

ABSTRACT

This report introduces some of the fundamental ideas of the current theory of automatic control to readers familiar with the 'classical' methods of Laplace transforms, root loci, etc. The mathematical approach is intuitive rather than rigorous. A selected bibliography is appended.

CONTENTS

FIGURES

SOME CONCEPTS OF MODERN CONTROL THEORY

– J.S. Riordon –

1. INTRODUCTION

About ten years ago, the traditional transform approach to the study of dynamic systems began to give way to the more fundamental time domain approach. With this change there grew a whole new vocabulary -- or jargon -- which perhaps obscured the fact that this "modern" approach is in many ways just a different method of describing the original system. However, its champions claim that it is much more than this -- that it is a powerful new technique which unifies heretofore diverse concepts, and allows the introduction of systematized computer-based design methods that were formerly not possible. The purpose of this report is to introduce some of the fundamental ideas of the current theory of automatic control to readers familiar with the "classical" methods of Laplace transforms, root loci, etc. The mathematical approach is intuitive rather than rigorous; for those readers who wish to pursue the subject in greater depth, a selected bibliography is appended.

2. CONTINUOUS-TIME LINEAR SYSTEMS

2.1 Transform Approach

A linear lumped parameter time-invariant dynamic system with one input u(t) and one output c(t) may be described by the ordinary differential equation

$$b_0 u + b_1 \dot{u} + \ldots + b_m u^{(m)} = a_0 c + a_1 \dot{c} + \ldots + a_n c^{(n)} \qquad (1)$$

where $\dot{u} = \dfrac{du}{dt}$ $\qquad u^{(m)} = \dfrac{d^m u}{dt^m}$

$\dot{c} = \dfrac{dc}{dt}$ $\qquad c^{(n)} = \dfrac{d^n c}{dt^n}$

In real physical processes $n > m$, and (1) is an n^{th} order differential equation; the system itself is said to be n^{th} order. By taking Laplace transforms, we obtain the transfer function

$$\frac{C}{U}(s) = \frac{b_0 + b_1 s + b_2 s^2 + \ldots + b_m s^m}{a_0 + a_1 s + a_2 s^2 + \ldots + a_n s^n} \qquad (2)$$

An analogue compute realization of this system is shown in fig. 1. Note that at least n integrators are required to simulate an n^{th} order system.

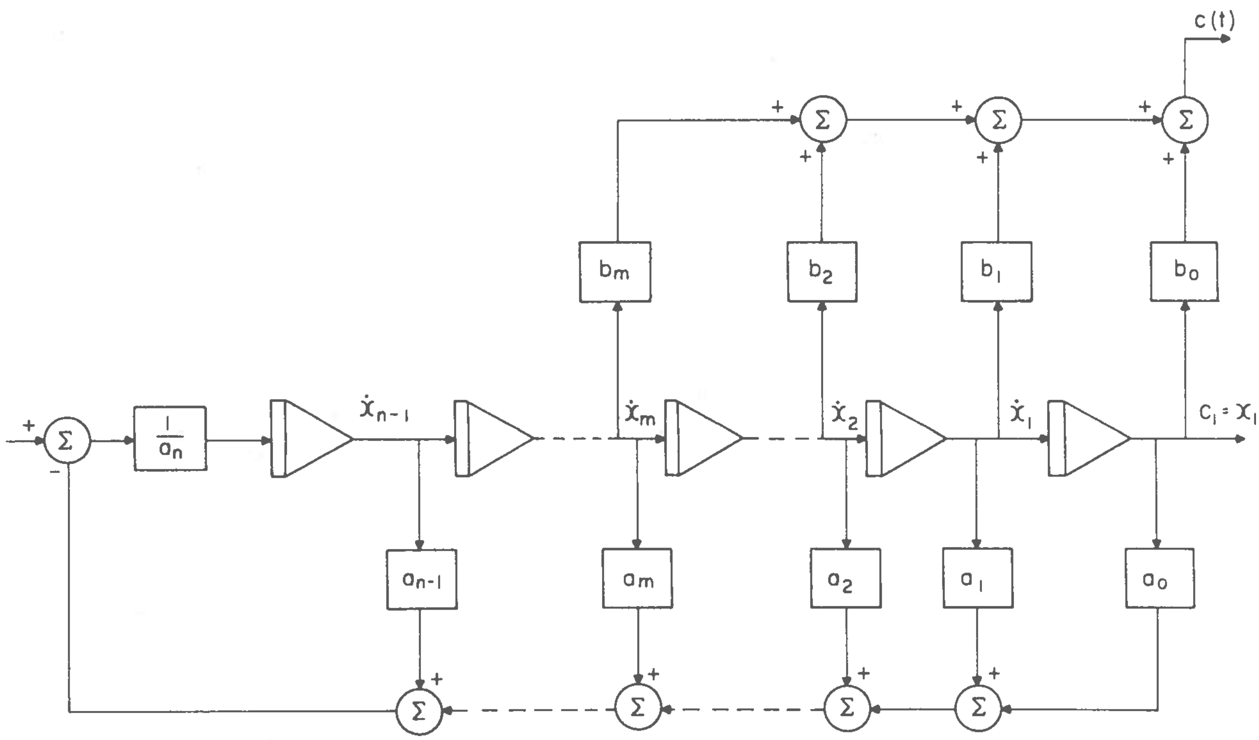

Figure 1 Realization of $\frac{C}{U}(s)$

The Laplace transform approach is useful, since it replaces the operation of convolution in the time domain with multiplication in the transform domain. Moreover, system stability can be determined by examination of the roots of the characteristic equation

$$a_0 + a_1 s + a_2 s^2 + \ldots + a_n s^n = 0 \tag{3}$$

Laplace transform techniques also suffer from several disadvantages:

i) the input $u(t)$ must be a Laplace-transformable function;

ii) design techniques are empirical, and tend to break down altogether for multi-input, multi-output systems;

iii) nonlinear or time-varying systems cannot be handled.

2.2 The State Space Concept

An n^{th} order differential equation may be solved numerically by transformation into n simultaneous first order differential equations. Consider the system of fig. 1. Let the variable c_1 be denoted x_1, let its first time derivative $\dot{c}_1$ be x_2, its second, x_3, and so on. Inspection of fig. 1 shows that

$$\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= x_3 \\
&\vdots \\
\dot{x}_{n-1} &= x_n \\
\dot{x}_n &= -\frac{a_0}{a_n} x_1 - \frac{a_1}{a_n} x_2 - \ldots \frac{a_{n-1}}{a_n} x_n + \frac{u}{a_n}
\end{aligned} \tag{4}$$

or in vector-matrix notation

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & & 0 & 1 \\ -\frac{a_0}{a_n} & -\frac{a_1}{a_n} & -\frac{a_2}{a_n} & \cdots & -\frac{a_{n-2}}{a_n} & -\frac{a_{n-1}}{a_n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \\ 0 \\ \frac{1}{a_n} \end{bmatrix} u \qquad (5)$$

The output, c, is given by

$$c = b_0 x_1 + b_1 x_2 + \cdots + b_m x_{m+1} \qquad (6)$$

In more general form, the dynamic equations of a linear system may be expressed as:

$$\boxed{\begin{aligned} \dot{\underline{x}}(t) &= A\,\underline{x}(t) + B\,\underline{u}(t) \qquad (7) \\ \underline{y}(t) &= C\,\underline{x}(t) \qquad (8) \end{aligned}}$$

where $x(t)$ is an n-vector (the state vector)

A is an n x n matrix

B is an n x q matrix (q control inputs)

$\underline{u}$ is a q-vector (the control vector)

$\underline{y}$ is an r-vector (the output vector of r elements)

C is an r x n matrix

In the system of (5) and (6) $q = r = 1$, so that $\underline{u} = u$ is a scalar and $B = \underline{b}$ is an n-vector.

Example: A system has the transfer function

$$\frac{C}{U}(s) = \frac{1}{s^2 + 3s + 2} \tag{9}$$

Express the system dynamics in the form of equations (7) and (8).

We observe that the differential equation of this system is

$$\ddot{c}(t) + 3\,\dot{c}(t) + 2\,c(t) = u(t) \tag{10}$$

Letting $x_1 = c$ and $x_2 = \dot{c}$, we can write

$$\dot{\underline{x}} = A\,\underline{x} + \underline{b}\,u \tag{11}$$

as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \tag{12}$$

Equation (8) becomes

$$y = [1 \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \tag{13}$$

The state $\underline{x}(t)$ of a system at time t may be represented by a point whose coordinates on a set of n orthogonal axes are $(x_1(t), x_2(t), \ldots x_n(t))$. The set of axes defines a state space -- just an extension of the phase plane concept -- and the path $\underline{x}(t)$, $t_0 < t < t_1$ traced out over a given time interval (t_0, t_1) is called the trajectory of the system in state space. It should be pointed out that successive state variables need not be time derivatives of each other as in (4) and (5). Any other linear combination of these variables will do just as well, providing all the combinations are linearly independent.

Some advantages of the state space formulation are:

i. a set of first order differential equations is often easier to solve than one n^{th} order equation;

ii. initial conditions are automatically included;

iii. the formulation applies to systems which are linear or nonlinear, continuous or time sampled, deterministic or stochastic;

iv. multi-input, multi-output systems may be handled;

v. computer solution of (7) is straightforward;

vi. design criteria may be stated in a manner suitable for computer calculation (this calculation may still be difficult to implement, though);

vii. for a given state $\underline{x}(t_1)$ at time t_1, the system behaviour at any previous time $t < t_1$, has no effect on future system response. In a sense the whole history of the system's evolution is summed up in the state vector.

2.3 System Stability

By analogy with the scalar equation

$$\dot{x}(t) = a\, x(t) \tag{14}$$

we might reasonably expect that the autonomous vector equation

$$\dot{\underline{x}}(t) = A\, \underline{x}(t) \tag{15}$$

represents a stable system if the matrix A is negative in some sense. In fact, its eigenvalues must be negative. Moreover, it can be shown that <u>the eigenvalues of A are simply the roots of the characteristic equation (3)</u>.

An eigenvalue of a matrix A is a scalar λ such that for some particular vector $\underline{x}$ (an eigenvector)

$$A\,\underline{x} = \lambda\,\underline{x} \tag{16}$$

Re-writing (16) as

$$A\,\underline{x} = \lambda\,I\,\underline{x} \tag{17}$$

we obtain

$$(A - \lambda\,I)\,\underline{x} = \underline{0}, \qquad \underline{x} \neq \underline{0} \tag{18}$$

where I is the identity matrix, with unit diagonal elements, and zero off-diagonal elements.

Equation (18) holds if the determinant $|A - \lambda I|$ is zero; i.e.,

$$\left|A - \lambda\,I\right| = 0 \tag{19}$$

Example: To find the eigenvalues of (9)

From (12), $A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$

$$\left|A - \lambda\,I\right| = \begin{vmatrix} 0 - \lambda & 1 \\ -2 & -3 - \lambda \end{vmatrix} = -\lambda(-3-\lambda)-(-2) = 0$$

i.e., $\lambda^2 + 3\,\lambda + 2 = 0$

$$\lambda_1 = -1;\ \lambda_2 = -2; \tag{20}$$

Note that (20) is the characteristic equation of (9) and that λ_1 and λ_2 are the corresponding poles in the s-plane. Obviously the system is stable.

Again considering the analogy of (14) and (15) we might reasonably assume that (15) has the solution

$$\underline{x}(t) = \exp\ [\cdot At]\ \underline{x}\,(0) \tag{21}$$

where $\underline{x}(0) = x(t{=}0)$ = initial conditions on integrators.

This assumption turns out to be true. The exponential of matrix A is itself a matrix denoted $\Phi(t)$ and given by

$$\Phi(t) = \exp\ [At] = I + At + \frac{1}{2!}(At)^2 + \frac{1}{3!}(At)^3 + \ldots \tag{22}$$

$\Phi(t)$ is known as the <u>state transition matrix.</u>

2.4 Calculation of $\Phi(t)$ Using Laplace Transforms

A number of methods exist for calculating $\Phi(t)$. The one we shall consider here, although not necessarily the easiest, relates the state transition matrix to Laplace transform analysis. If X(s) is the Laplace transform of state vector $\underline{x}(t)$, then (15) may be written as

$$s\,X(s) - \underline{x}(0) = A\,X(s) \tag{23}$$

$$\therefore\ X(s) = \left[sI - A\right]^{-1}\underline{x}(0)$$

$$\text{i.e.,}\ \underline{x}(t) = \mathcal{L}^{-1}\left\{\left[s\,I - A\right]^{-1}\right\}\underline{x}(0) \tag{24}$$

Comparing (24) with (21) and noting (22), we see that

$$\Phi(t) = \mathcal{L}^{-1}\left\{\left[sI - A\right]^{-1}\right\} \tag{25}$$

<u>Example:</u> Calculate $\Phi(t)$ for the system of equation (9).

$$\left[sI - A\right] = \begin{bmatrix} s & -1 \\ 2 & s+3 \end{bmatrix}$$

$$\left[sI - A\right]^{-1} = \begin{bmatrix} \dfrac{s+3}{(s+1)(s+2} & \dfrac{1}{(s+1)(s+2)} \\ \dfrac{-2}{(s+1)(s+2)} & \dfrac{s}{(s+1)(s+2)} \end{bmatrix}$$

$$\therefore \quad \Phi(t) = \begin{bmatrix} 2\,e^{-t} - e^{-2t} & e^{-t} - e^{-2t} \\ 2\,e^{-2t} - 2\,e^{-t} & 2\,e^{-2t} - e^{-t} \end{bmatrix} \tag{26}$$

In this second order system we have

$$\left.\begin{aligned} x_1(t) &= \phi_{11}\, x_1(0) + \phi_{12}\, x_2(0) \\ x_2(t) &= \phi_{21}\, x_1(0) + \phi_{22}\, x_2(0) \end{aligned}\right\} \tag{27}$$

Note the interpretation which may be put on elements $\phi_{ij}(t)$ of the transition matrix. By inspection of (27), we have

$\phi_{ij}(t)$ = value of state variable x_i at time t resulting from a unit initial value of x_j when all other state variables except x_j have zero initial value.

A useful property of the state transition matrix is that

$$\Phi\,(t_1+t_2)\,\underline{x} = \Phi\,(t_1)\quad[\Phi\,(t_2)\,\underline{x}\,]$$

$$\text{i.e.,}\qquad \Phi\,(t_1+t_2) = \Phi\,(t_1)\,\Phi\,(t_2) \tag{28}$$

This is known as the semigroup property of the operator $\Phi(t)$; it underlies the usefulness of the state transition method in sampled data systems and optimal control.

2.5 Systems with Control Input

Consider the scalar equation

$$\dot{x}(t) = a\,x(t) + b\,u(t) \tag{29}$$

If h(t) is the system impulse response then use of the convolution integral shows that the solution of (29) is

$$x(t) = x(0)\,e^{at} + \int_0^t h(t-\tau)\,bu(\tau)\,d\tau \tag{30}$$

But since this is a first order system

$$h(t) = e^{at} \tag{31}$$

Again the analogy between (29) and (7) correctly suggests that the solution of (7) is

$$\boxed{\underline{x}(t) = \Phi(t)\,\underline{x}(0) + \int_0^t \Phi(t-\tau)\, B\, \underline{u}(\tau)\, d\tau} \tag{32}$$

3. DISCRETE TIME LINEAR SYSTEMS

3.1 Transform Approach

In the theory of sampled data systems it is assumed that the continuous function f(t) shown in fig. 2a is modulated by a train of impulse functions

$$\sum_{k=-\infty}^{k=\infty} \delta(t-kT)$$

of "zero" width, "infinite" amplitude, and unit area, each pulse being spaced by a period T, the sampling time (see fig. 2b). The modulator output $f^*(t)$, shown in fig. 2c, has the form

$$f^*(t) = f(t) \sum_{k=-\infty}^{\infty} \delta\,(t - kT) \tag{33}$$

If $f(t) = 0$ for $t < 0$, then

$$f^*(t) = \sum_{k=0}^{\infty} f(kT)\, \delta\,(t-kT) \tag{34}$$

and its Laplace transform is

$$\mathcal{L}\,[f^*(t)\,] = F^*(s) = \int_0^\infty f^*(t)\; e^{-st}\; dt$$

$$= \int_0^\infty \sum_{k=0}^{\infty} f(kT)\;\delta\;(t-kT)\; e^{-st}\; dt$$

i.e., $$F^*(s) = \sum_{k=0}^{\infty} f(kT)\; e^{-kTs} \tag{35}$$

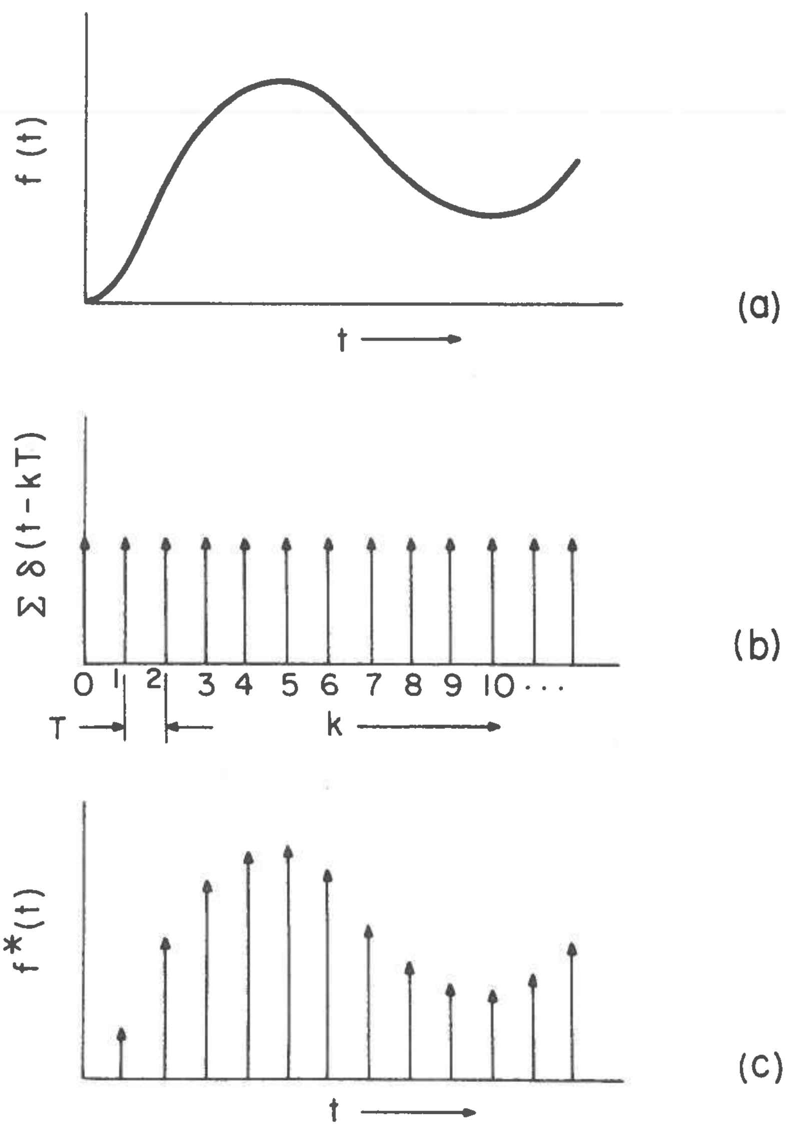

Figure 2 Impulse modulation

We now substitute the variable

$$z = e^{sT} \tag{36}$$

and obtain the z transform F(z)

$$F(z) = F^*(s) \Big|_{s = \frac{1}{T} \log z} = Z\,[f(t)]$$

defined by

$$\boxed{F(z) = \sum_{k=0}^{\infty} f(kT)\, z^{-k}} \tag{37}$$

Here the symbol Z (·) indicates "the z transform of".

The z transform is the discrete time equivalent of the Laplace transform, and the theorems of linearity, superposition, convolution, etc. all have discrete time counterparts. Equation (36) is a conformal mapping which maps the infinite number of poles and zeros of $F^*(s)$ in the s plane into a finite, and therefore manageable, set in the z plane. Similarly the stable left hand side of the s plane is mapped into the interior of the unit circle in the z plane.

Since the Laplace transform $1/(s+a)$ has an equivalent z transform $z/(z-e^{-aT})$, the z transform of a system may be obtained from a partial fraction expansion of its Laplace transform.

<u>Example</u>: Determine $\frac{C}{U}(z)$ when $\frac{C}{U}(s)$ is given by (9).

$$\frac{C}{U}(s) = \frac{1}{(s+1)(s+2)} = \left[\frac{1}{s+1} - \frac{1}{s+2}\right]$$

$$\therefore \quad \frac{C}{U}(z) = \frac{z}{z - e^{-T}} - \frac{z}{z + e^{-2T}}$$

$$\frac{C}{U}(z) = \frac{(\alpha_1 - \alpha_2)\, z}{(z-\alpha_1)(z-\alpha_2)}$$

where $\alpha_1 = e^{-T}; \quad \alpha_2 = e^{-2T}$ (38)

Suppose the sampling time, T, is 0.2 seconds. Then $\alpha_1 = 0.819$ and $\alpha_2 = 0.670$, so that

$$\frac{C}{U}(z) = \frac{0.149\ z}{z^2 - 1.489\ z + 0.549} \tag{39}$$

Note that α_1 and α_2 are the roots of the characteristic equation; since both of them lie within the unit circle, the system is stable.

One important property of the z transform is expressed by the shift theorem, which states that if a function f(t) has a value f(k) at t = kT, then

$$Z\,[\,f(k-1)\,] = z^{-1}\ Z\,[\,f(k)\,] \tag{40}$$

We may re-write (39) as

$$C(z)\ (1-1.489\ z^{-1} + 0.549\ z^{-2}) = U(z)(0.149\ z^{-1}) \tag{41}$$

In view of (40), (41) is seen to be

$$Z\,[c(k)] - 1.489\ Z\,[c(k-1)] + 0.549\ Z\,[c(k-2)]$$
$$= 0.149\ Z\,[u(k-1)\,]$$

so that the difference equation of this system is evidently

$$c(k) - 1.489\ c(k-1) + 0.549\ c(k-2) = 0.149\ u(k-1) \tag{42}$$

Equation (42) is a recursion relationship which allows the computation of the output sequence $\{c(k)\}$ for any given input sequence $\{u(k)\}$, k = 0, 1, 2, ... ∞.

In practice the system input function is not a series of impulses, but usually a piecewise constant (staircase) function. This may be considered as the output of a zero-order hold (sample-and-hold) circuit whose input is a modulated impulse train. In such a case the z transform becomes

$$\frac{C}{U}(z) = \frac{z-1}{z} \cdot Z\left[\frac{1}{s} \cdot \frac{C}{U}(s)\right] \tag{43}$$

<u>Example:</u> previous case, with sample-and-hold circuit added.

$$\frac{C}{U}(z) = \frac{z-1}{z} \cdot Z\left[\frac{1}{s(s+1)(s+2)}\right]$$

$$= \frac{z-1}{z} \cdot Z\left[\frac{1}{2s} - \frac{1}{s+1} + \frac{1}{2(s+2)}\right]$$

$$= \frac{z-1}{z}\left[\frac{z}{2(z-1)} - \frac{z}{(z-\alpha_1)} + \frac{z}{2(z-\alpha_2)}\right]$$

$$\frac{C}{U}(z) = \frac{0.5\ (1-2\ \alpha_1 + \alpha_2)\ z + 0.5\ (\alpha_1 + \alpha_1\ \alpha_2 - 2\ \alpha_2)}{(z - \alpha_1)\ (z-\alpha_2)} \tag{44}$$

For T = 0.2 seconds, (44) becomes

$$\frac{C}{U}(z) = \frac{0.0164\ z + 0.0134}{z^2 - 1.489\ z + 0.549} \tag{45}$$

The corresponding difference equation is

$$c(k) - 1.489\ c(k-1) + 0.549\ c(k-2) = 0.0164\ u(k-1) + 0.0134\ u(k-2) \qquad \ldots (46)$$

Comparing (45) with (39) we see that the poles are unchanged, but that the zero has moved from the origin to $z = -(0.0134/0.0164) = -0.817$, and the d.c. gain has increased. To observe the latter effect, we may invoke the final value theorem

$$\lim_{k\to\infty} c(kT) = \lim_{z\to 1} (z-1)\ C(z) \tag{47}$$

which allows the calculation of the steady state response for a fixed input. If the input is a unit pulse train with z transform

$$U(z) = \frac{z}{z-1}$$

then the steady state response of the system without a sample-and-hold circuit is

$$\lim_{k\to\infty} c(k) = \lim_{z\to 1} (z-1)\, U(z)\, \frac{C}{U}(z)$$

$$= \lim_{z\to 1} (z-1) \left(\frac{z}{z-1}\right)\left(\frac{0.149z}{z^2 - 1.489z + 0.549}\right)$$

$$\lim_{k\to\infty} c(k) = 2.50 \tag{48}$$

while the steady state response with the sample-and-hold circuit must be the same as the response of the continuous system to a step function.

$$\lim_{k\to\infty} c(k) = \lim_{z\to 1} (z-1)\left(\frac{z}{z-1}\right)\left(\frac{0.0164z + 0.0134}{z^2 - 1.489z + 0.549}\right)$$

$$\lim_{k\to\infty} c(k) = 0.50 \tag{49}$$

which agrees with the continuous calculation with $U(s) = 1/s$; i.e.,

$$\lim_{t\to\infty} c(t) = \lim_{s\to 0} s\, U(s)\, \frac{C}{U}(s)$$

$$= \lim_{s\to 0} s.\ \frac{1}{s}\ .\ \frac{1}{(s+1)(s+2)} = 0.50$$

Like the Laplace transform, the z transform is a powerful analytical tool; however, it suffers from similar disadvantages. In particular, it is not ideally suited for synthesis; design techniques tend to be empirical, and the extension to nonlinear, multi-input, and stochastic systems is difficult.

3.2 State Space Analysis of Sampled Data Systems

Since the state space approach works directly in the time domain and does not rely on transform methods, the introduction of time sampling involves little change in the system equations. In fact, it is only necessary to specify T and the form of u (impulse train or piecewise constant) in (32), and we move in a single painless step into the discrete time domain.

Example: Previous case; $T = 0.2$ seconds; $B\underline{u} = \underline{b}\, u\, \delta\,(t-kT)$ is an impulse train of time-varying magnitude $u(t)$; $\underline{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ as in (12).

From (26), with $t = T = 0.2$,

$$\Phi(0.2) = \begin{bmatrix} 0.968 & 0.149 \\ -0.298 & 0.521 \end{bmatrix} \tag{50}$$

and in (32)

$$\int_0^T \Phi\,(T-\tau)\ \underline{b}\, u(\tau)\ \delta\,(t-kT)\ d\tau = \underline{b}\, u(T)$$

so that

$$\underline{x}(1) = \Phi(T)\ \underline{x}(0) + \underline{b}\ u(1)$$

and in general

$$\underline{x}(k+1) = \Phi(T)\ \underline{x}(k) + \underline{b}\ u(k+1) \tag{51}$$

i.e.;

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}(k+1) = \begin{bmatrix} 0.968 & 0.149 \\ -0.298 & 0.521 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k+1) \quad \ldots \tag{52}$$

For a given impulse sequence $\left\{u(t)\ \delta\ (t-kT)\right\}$, the output sequence $\left\{x_1(k)\right\}$ calculated with (52) is identical to the sequence $\left\{c(k)\right\}$ calculated from the difference equation (42). If the input is piecewise constant instead of an impulse train, the final term of (32) may be adjusted accordingly. Let $u(t)$ take on a series of values $\left\{u(k),\ kT < t \leq (k+1)\ T\right\}$. Then

$$\int_0^T \Phi(T-\tau\ )\ \underline{b}\ u(k)\ d\tau = D\ \underline{b}\ u(k) \tag{53}$$

where D is an n x n matrix whose elements $d_{ij}(T)$ are given by

$$d_{ij}(T) = \int_0^T \phi_{ij}(T-\tau)\, dt \tag{54}$$

In the example considered previously, the matrix $\Phi(T-\tau)$ is obtained from (26) with $t = T - \tau$, and may be integrated term by term to yield

$$D = D(T) = \begin{bmatrix} 1.5 - 2\,\alpha_1 + 0.5\,\alpha_2 & 0.5 - \alpha_1 + 0.5\,\alpha_2 \\ -1 + 2\,\alpha_1 - \alpha_2 & \alpha_1 - \alpha_2 \end{bmatrix} \quad \ldots \quad (55)$$

where α_1 and α_2 are defined by (38). For T = 0.2, (55) becomes

$$D\,(0.2) = \begin{bmatrix} 0.198 & 0.0164 \\ -0.0329 & 0.149 \end{bmatrix}$$

so that the state equation corresponding to difference equation (46) is

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}(k+1) = \begin{bmatrix} 0.968 & 0.149 \\ -0.298 & 0.521 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}(k) + \begin{bmatrix} 0.0164 \\ 0.149 \end{bmatrix} u(k) \tag{56}$$

In the case of a multiple input $\underline{u}(k)$, the general form of (56) is

$$\underline{x}\,(k+1) = \Phi(T)\,\underline{x}(k) + D(T)\,B\,\underline{u}(k) \tag{57}$$

3.3 System Stability

The relationship of $\Phi(T)$ to the characteristic equation in z is equivalent to that of A (equation 7) to the characteristic equation in s. Thus the eigenvalues λ_i of $\Phi(t)$ are identical to the roots of the characteristic z equation (check equation (52) against (39) in this respect). A necessary and sufficient condition for stability of a linear sampled data system is that

$$|\lambda_i| < 1$$

4. NONLINEAR SYSTEMS

A general deterministic n^{th} order nonlinear system with state, control, and output vectors $\underline{x}$, $\underline{u}$, and $\underline{y}$ respectively may be represented by the state description

$$\boxed{\begin{aligned} \dot{\underline{x}}(t) &= \underline{f}(\underline{x}, \underline{u}, t) \qquad (58) \\ \underline{y}(t) &= \underline{g}(x,t) \qquad (59) \end{aligned}}$$

where $\underline{f} = \begin{bmatrix} f_1(\underline{x}, \underline{u}, t) \\ f_2(\underline{x}, \underline{u}, t) \\ \vdots \\ f_n(\underline{x}, \underline{u}, t) \end{bmatrix}$ = vector of n functions

and $\underline{g}$ is similarly a column vector of r functions.

Equations (58) and (59) are the nonlinear equivalents of (7) and (8).

Example: If θ is the electrical angle between the stator and rotor fields of a synchronous motor, then the torque equation is

$$\ddot{\theta} + \alpha\,\dot{\theta} + \beta \sin \theta = L \qquad (60)$$

where L is normalized torque, and α and β are constants. Letting $x_1 = \theta$, $x_2 = \dot{\theta}$, and $u = L$, we obtain the nonlinear state equation

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -\beta \sin x_1 - \alpha\, x_2 + u \end{bmatrix} \qquad (61)$$

If the angle but not its derivative is observable, then

$$\underline{y} = [1 \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \qquad (62)$$

There is no simple solution to (58). However, if $\underline{f}$ is known, then the n equations may be integrated numerically to obtain the state trajectory for specific initial conditions and a specific control sequence. Often it is desirable to linearize the system equations about some (perhaps time-varying) operating point, so that

$$\underline{\dot{x}}(t) = A(t)\, \underline{x}(t) + B(t)\, \underline{u}(t) \tag{63}$$

where A(t) is the Jacobian matrix

$$A(t) = \frac{\partial \underline{f}}{\partial \underline{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \tag{64}$$

and

$$B(t) = \frac{\partial \underline{f}}{\partial \underline{u}} \tag{65}$$

In the case of discrete time nonlinear systems, the general formulation is

$$\boxed{\begin{aligned} \underline{x}(k+1) &= \underline{f}\,(\underline{x}(k), \underline{u}(k), k) && (66) \\ \underline{y}(k) &= \underline{g}\,(\underline{x}(k), k) && (67) \end{aligned}}$$

For second order systems, or those which may be approximated as such, considerable work has been done on <u>phase plane analysis</u> -- singular points, limit cycles, etc. For control loops containing a separable non-linearity, the extension of linear frequency response methods has led to the concept of the <u>describing function.</u> Sufficient conditions for stability of nonlinear systems may sometimes be determined by the application of Lyapunov's second method of stability analysis. All of these approaches are limited, though, and no general theory of nonlinear systems exists. Though further progress will no doubt be made, it is probable that no general theory will ever exist, because nonlinear systems are by definition the complement of a set; i.e., they are the hodge-podge left over after one removes from the set of all systems the well defined set of linear systems.

5. OPTIMIZATION TECHNIQUES: THE DESIGN PROBLEM

5.1 Performance Criteria

If we are to design optimum systems, we must first define the term "optimum". To do this we introduce the concept of a performance criterion, also known as a performance index (PI) or cost function. As the name suggests, it is simply a measure of the system performance; an optimum system with respect to a given cost function is one which causes that cost function to assume an extremum value (maximum or minimum, as desired). For example, the well known mean squared error criterion imposes a quadratic cost on system errors. The quadratic form is widely used, as it corresponds in many physical situations to a measure of power. In some cases the cost function of a system changes as circumstances are altered, so that the optimum solution today is different from that of yesterday. Thus under normal circumstances a merchant ship may steam at about 16 knots to minimize a cost which is a function of fuel consumption, wages, and delivery time. The optimum speed for the same ship in time of war may be 24 knots; the difference is due to the presence in the second instance of an additional speed-dependent cost term, "probability of being sunk times resultant cost".

In addition, all physical systems are subject to constraints, although these may sometimes be ignored. Suppose in the previous example that the ship's maximum speed is 24 knots. Under normal circumstances, this constraint may be ignored in calculating optimum speed, since its presence does not affect the result. In the second case with the altered cost function, however, the constraint is operative, and must be considered (the optimum speed with no constraint might be, say, 60 knots).

5.2 Static Optimization

In general system cost V is a function of a number of variables $x_1, x_2, \ldots x_n$ which may be manipulated either directly or indirectly. If we regard the set of x_i's as a vector $\underline{x}$, then the optimization problem is to choose a vector $\underline{x}^* = (x_1^*, x_2^*, \ldots x_n^*)^T$ to minimize the cost function

$$V = V(x_1, x_2, \ldots x_n) = V(\underline{x}) \tag{68}$$

The problem is said to be static if $\underline{x}$ is not a function of time. The classical approach to this problem is indirect; that is, instead of seeking $\underline{x}^*$ directly, we form a set of equations

$$V_{\underline{x}} = \frac{\partial V}{\partial \underline{x}} = \left(\frac{\partial V}{\partial x_1}, \frac{\partial V}{\partial x_2}, \ldots \frac{\partial V}{\partial x_n} \right) = \underline{0} \tag{69}$$

whose solution yields $\underline{x}^*$(subject to suitable conditions on the second partial derivatives).

Frequently it is either not feasible to differentiate V or else (69) has no analytic solution. In such a case a direct method may be used to determine $\underline{x}^*$. This technique usually involves some form of hill-climbing (for minimization problems, the term, "valley descending" might be more appropriate), the most common being that of steepest descent. The gradient $V_{\underline{x}}$ is calculated or estimated (by small perturbations about $\underline{x}$) and an improved value of V is found by moving in $\underline{x}$-space along the gradient vector (the steepest path down the hill). Thus given an initial point $\underline{x}_0$ with gradient $V_{\underline{x}}(0)$ we determine a new point $\underline{x}_1$ by

$$\underline{x}_1 = \underline{x}_0 + \rho \; V_{\underline{x}}^T(0) \tag{70}$$

where ρ is a scalar (the superscript T is a transpose operator).

If $\underline{x}_0$ is far from $\underline{x}^*$, then almost any move along $V_{\underline{x}}$ represents an improvement (reduction in V). However, as the optimum point is approached, baffling features may arise. One problem is the determination of a suitable value for ρ so that successive values of $\underline{x}$ neither creep slowly towards $\underline{x}^*$ nor overshoot it altogether. In addition the hill itself may have a fairly exotic shape, such as a steep four-dimensional curving ridge. A variety of ingenious methods for dealing with such problems appears in the literature.

Constraints introduce a further complication. Inequality constraints may be handled frequently by adaptations of hill-climbing methods. For linear cost functions and constraints, the highly developed method of linear programming is an extremely powerful tool. Equality constraints may in simple cases be handled by direct substitution.

Example: Minimize

$$V = x_1 + 2\,x_2 \tag{71}$$

subject to

$$x_1\,x_2 = 8 \tag{72}$$

Direct substitution of (72) into (71) shows that

$$V = x_1 + 16\,x_1^{-1}$$

and the solution of (69) yields

$$x_1 = -\,4$$

$$x_2 = -\,2$$

An alternative method is the use of the Lagrange multiplier technique. Let the m equality constraints be expressed in the form

$$\phi_i(\underline{x}) = 0, \qquad i = 1,2, \ldots m \tag{73}$$

We now form a new cost function $V(\underline{x}, \underline{\lambda})$ such that

$$V(\underline{x}, \underline{\lambda}) = V(\underline{x}) + \sum_{i=1}^{m} \lambda_i \phi_i \tag{74}$$

It can be shown that choice of the (n+m) variables $(x_1, x_2, \ldots x_n, \lambda_1, \lambda_2, \ldots \lambda_m)$ to minimize (74) with no constraints solves the original constrained minimization problem.

Example: equations (71) and (72). Note that m = 1 and

$$\phi_1 = \phi = x_1 x_2 - 8$$

so that

$$V(\underline{x}, \lambda) = x_1 + 2 x_2 + \lambda (x_1 x_2 - 8) \tag{75}$$

Partial differentiation of (75) with respect to x_1, x_2, and λ gives three simultaneous equations

$$\left.\begin{aligned} 1 + \lambda x_2 &= 0 \\ 2 + \lambda x_1 &= 0 \\ x_1 x_2 - 8 &= 0 \end{aligned}\right\} \tag{76}$$

For a minimum

$$x_1 = - 4, \; x_2 = - 2, \qquad \lambda = 0.5$$

In this example direct substitution is simpler than the use of a Lagrange multiplier. Suppose that it is necessary to minimize (71) subject to the constraint

$$\int_{-\infty}^{\infty} \exp \left[-\frac{1}{2} (x_1-3)^2 \right] \int_{x_1}^{\infty} \exp \left[-\frac{1}{2} (x_2-5)^2 \right] dx_1 \, dx_2 = 6$$

Here ϕ (x_1, x_2) is so complicated that direct substitution is not feasible. A computer solution with the Lagrange multiplier approach is quite straightforward, however.

5.3 Dynamic Optimization

In a dynamic system the control action taken at any time affects the future evolution of the system trajectory so that a sequence of control inputs constitutes a set of interacting variables. The system cost function measures performance over the whole of a given time interval (which may be infinite if desired), and so usually takes the form of a time integral. Thus if power is the instantaneous parameter of interest, optimization might imply the design of a minimum energy system. If it is desired to minimize the amount of fuel required to place a satellite in a given orbit (so that payload is a maximum), then the integral of fuel flow rate must be minimized. Mathematically, then, the problem is as follows:

Given the system dynamics

$$\dot{\underline{x}} = \underline{f} \, (\underline{x}, \underline{u}, t) \tag{77}$$

with initial conditions

$$\underline{x} \, (0) = \underline{x}_0 \tag{78}$$

choose a control sequence $\left\{ \underline{u}(t),\ 0 \leq t < T \right\}$ or a feedback controller $\left\{ \underline{u}(\underline{x},t),\ 0 \leq t < T \right\}$ to act in a given time interval $[0,T)$ which will minimize the (given) cost function

$$V = \int_0^T L\ (\underline{x},\ \underline{u},\ t)\ dt\ + V_T\ (\underline{x}(T)\ ) \tag{79}$$

$L(\underline{x},\ \underline{u},\ t)$ is a continuous scalar function (such as a quadratic) and V_T represents end-point costs (for instance, the cost of errors in the final satellite trajectory). In any real system there will also exist various constraints, but these will be ignored for the present.

5.4 Calculus of Variations

The classical approach to this problem is through the calculus of variations developed by Euler, Lagrange, Bernoulli and others over the past 300 years or so. It is desired to choose a function $\underline{x}(t)$ such that the performance index

$$V(\underline{x},\ \dot{\underline{x}}) = \int_0^T L(\underline{x},\ \dot{x},\ t)\ dt \tag{80}$$

is minimized. It can be shown that a necessary condition for minimization is given by the celebrated Euler-Lagrange equation

$$\boxed{\frac{\partial L}{\partial \underline{x}} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = 0} \tag{81}$$

We can fit (77) - (79) to this mould by letting the q control variables represent q additional state variables $x_{n+1}, \ldots x_{n+q}$, and by minimizing the cost function (79) subject to the constraint condition that the dynamic equation (77) be satisfied. To ensure this condition, we introduce n Lagrange multipliers $\lambda_1(t), \dots \lambda_n(t)$, just as in the static case, except that now the λ's are time-varying. The Euler-Lagrange equation is applied to the modified cost function formed by adjoining the dynamic constraints. Boundary conditions on the λ's are determined from transversality conditions arising from the theory of the calculus of variations.

Example: Given an interval [0, T), minimize

$$V = \int_0^T (x_1^2 + a\,u^2)\,dt \qquad (82)$$

for the system of equation (9).

Letting $u = x_3$ and putting the constraints in the form of (73), we obtain from (12)

$$\begin{aligned} \phi_1 &= \dot{x}_1 - x_2 = 0 \\ \phi_2 &= \dot{x}_2 + 2\,x_1 + 3\,x_2 - x_3 = 0 \end{aligned} \qquad (83)$$

so that the adjoined cost function $L(\underline{x}, \lambda)$ is

$$L(\underline{x}, \underline{\lambda}) = x_1^2 + a\,x_3^2 + \lambda_1\,(\dot{x}_1 - x_2) + \lambda_2\,(\dot{x}_2 + 2\,x_1 + 3\,x_2 - x_3) \qquad (84)$$

If (81) is applied to (84) and the variable x_3 is eliminated by substitution, four equations result:

$$\left.\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= -2\,x_1 - 3\,x_2 + \frac{1}{2a}\,\lambda_2 \\
\dot{\lambda}_1 &= 2\,x_1 + 2\,\lambda_2 \\
\dot{\lambda}_2 &= -\,\lambda_1 + 3\,\lambda_2
\end{aligned}\right\} \tag{85}$$

Note that the variables are interconnected. Inspection of (85) shows that the system takes the form shown in fig. 3. At the top is

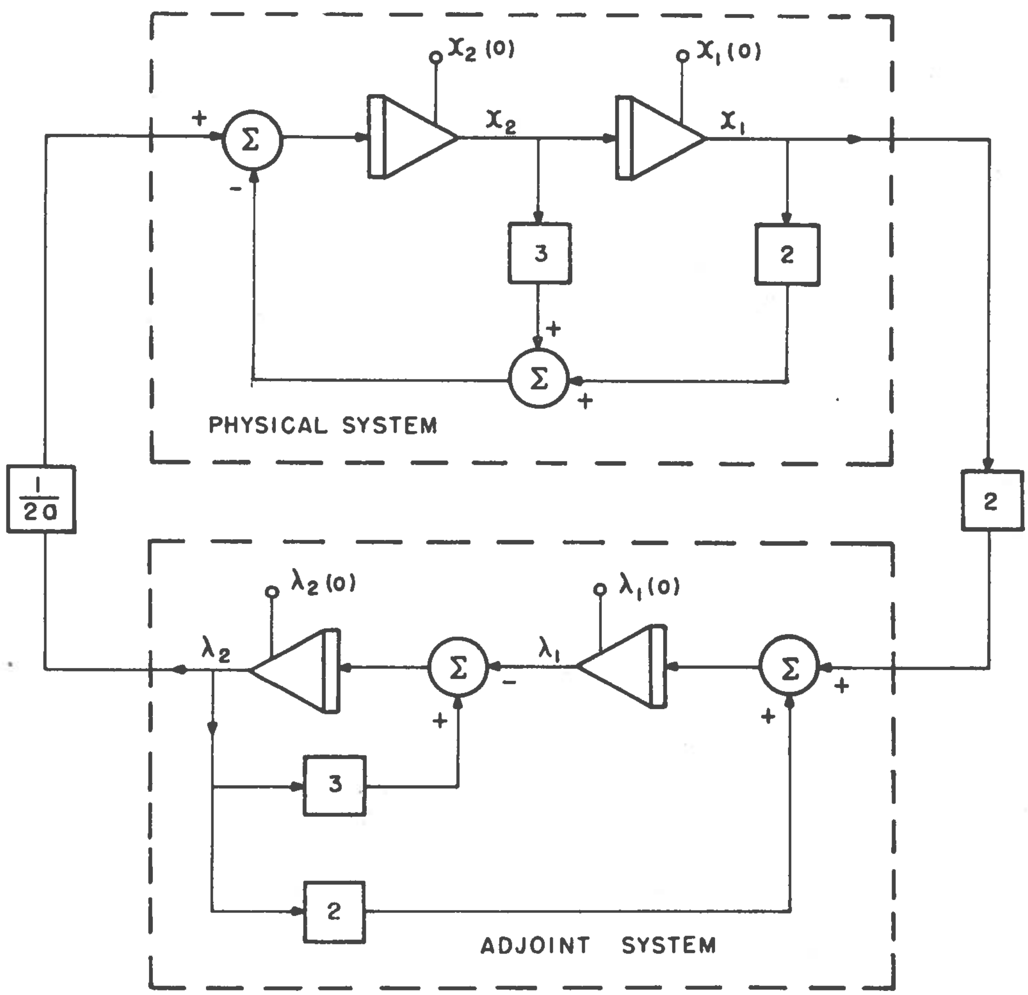

Figure 3 Realization of Euler-Lagrange equations

the original system, while below it is another dynamic system known as the adjoint system, which generates an optimum (minimum cost) control signal. Note the very important fact that, although no a priori assumption of feedback was made, the result indicates that a feedback system is optimum. There is, however, quite a problem in implementing suitable control. Equations (85) constitute four differential equations for which four boundary conditions are known. But-- and this is the fly in the ointment -- initial conditions are known for $\underline{x}$, while final conditions ($\lambda_1(T) = \lambda_2(T) = 0$ in this case) are known for $\underline{\lambda}$. Thus, whether the equations are integrated either forwards or backwards in time, two of the boundary conditions must be guessed. Just to make matters worse, the equations governing the missing conditions are unstable; i.e., the $\underline{x}$ equations are unstable in reverse time, and the $\underline{\lambda}$ equations in forward time. Thus even a small error in the guess will cause an enormous error at the opposite boundary. This difficulty, which is a feature of most dynamic optimization problems, is known as the two-point boundary value problem (TPBVP). A solution of (85) may be obtained numerically using successive approximations; however, as the order of the system increases, this becomes a very time-consuming task.

5.5 The Maximum Principle

One of the limitations of the calculus of variations is that the Euler-Lagrange equation is not valid when the magnitudes of control inputs u_i are limited so that

$$u_i \leq U_i \tag{86}$$

as is frequently the case. In 1956 the Russian mathematician Pontryagin formulated a new approach to the optimization problem which circumvented this difficulty. Given a system of the form (77) and (78), it is desired to choose $u(t)$, $0 \leq t < T$ to minimize a cost function

$$V(\underline{x}) = \sum_{i=1}^{n} c_i x_i \tag{87}$$

where parameters c_i are known constants. Now let a function H, the Hamiltonian function be defined as

$$H = \langle \underline{p}, \underline{f} \rangle = \sum_{i=1}^{n} p_i f_i \tag{88}$$

where $\underline{f}$ is defined by (77) $p_i(t)$ are the adjoint or co-state variables. The relationship between $\underline{f}$ and $\underline{p}$ is given by the canonical Hamiltonian equations

$$\boxed{\begin{aligned} \dot{\underline{x}} &= H_p \\ \dot{\underline{p}} &= -H_x \end{aligned}} \qquad \begin{aligned} (89) \\ (90) \end{aligned}$$

where
$$H_p \equiv \frac{\partial H}{\partial \underline{p}} \quad ; \quad H_x \equiv \frac{\partial H}{\partial \underline{x}}$$

with boundary conditions on $\underline{p}$

$$p_i(T) = - c_i \tag{91}$$

if $\underline{x}(T)$ is free (not specified).

The Maximum Principle states that (87) is minimized if $\underline{u}$ is adjusted at all times to maximize the Hamiltonian H given by (88).

Example: as in previous section

Let $$\dot{x}_3 = x_1^2 + a\,u^2 \tag{92}$$

so that $$V = c_1x_1 + c_2x_2 + c_3x_3 \tag{93}$$

where $c_1 = c_2 = 0;\ \ c_3 = 1$

The Hamiltonian is

$$H = p_1x_2 - p_2(2\,x_1 + 3\,x_2 - u) + p_3(x_1^2 + a\,u^2) \tag{94}$$

Application of (89) yields the original system equations together with (92), while (90) becomes

$$\dot{p}_1 = 2\,p_2 - 2\,x_1p_3 \tag{95}$$

$$\dot{p}_2 = -p_1 + 3\,p_2 \tag{96}$$

$$\dot{p}_3 = 0 \tag{97}$$

Boundary conditions (91) are

$$p_1(T) = p_2(T) = 0; \qquad p_3(T) = -1$$

In view of (97), it can be seen that $p_3(t) = -1$ for all t. Substituting this value in (95), we see that the adjoint equations (95) and (96) are identical to those of (85). If u is unconstrained, maximization of H implies

$$\frac{\partial H}{\partial u} = H_u = p_2 + 2\,a\,p_3u = 0$$

$$\text{i.e.; } u = \frac{1}{2a}\ p_2 \tag{98}$$

and the state equations become

$$\left.\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\,2\,x_1 - 3\,x_2 + \frac{1}{2a}\,p_2 \end{aligned}\right\} \tag{99}$$

If there were a limitation

$$|u| \leq U$$

on control effort, then (98) would be

$$\left.\begin{aligned} u &= \frac{1}{2a} p_2 , & |p_2| &\leq 2 a U \\ u &= +U , & p_2 &> 2 a U \\ u &= - U , & p_2 &< (-2 a U) \end{aligned}\right\} \qquad (100)$$

and the four simultaneous differential equations in x_1, x_2, p_1, p_2 would be nonlinear.

A comparison of (99), (95), and (96) with (85) shows that in the unconstrained case $\underline{\lambda}$ and $\underline{p}$ are equivalent. Note that the TPBVP is still present as well.

In a general linear system the relationship of the dynamics of $\underline{x}$

$$\dot{\underline{x}} = A \underline{x} + B \underline{u}$$

to those of $\underline{p}$ is clearly given by (90), which shows that

$$\dot{\underline{p}} = - A^T \underline{p} - \left(\frac{\partial L}{\partial \underline{x}} \right)^T \qquad (101)$$

where $\underline{p}$ is treated as a column vector. Equation (101) shows that if $\underline{x}$ is stable in forward time, then $\underline{p}$ is not, as has been noted previously.

5.6 Dynamic Programming: A Discrete Time Example

Concurrent with the development in the Soviet Union of Pontryagin's maximum principle was the formulation by Bellman in the United States of the method of dynamic programming. The idea behind this technique is to imbed the control problem relating to a particular initial condition $\underline{x}_0$ within a wider problem in which the initial and final conditions may take on any values within prescribed limits. If all such problems are solved, then the TPBVP ceases to exist.

Control of a dynamic system may be regarded as a multi-stage decision process, in which each decision depends upon the current state of the system. Thus control actions are dependent upon state; this is equivalent to saying that feedback control is to be used. The optimal control takes the form of a control policy $\underline{u}^*(\underline{x},t)$, rather than an essentially open-loop control schedule $\underline{u}^*(t)$ applying only to a particular value of $\underline{x}_0$. Underlying this concept is the principle of optimality:

> An optimal policy has the property that whatever the initial state and the initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

For a sampled data system in which control actions take place at a series of discrete stages or time intervals, the application of the optimality principle may be seen in the following example:

Let the system be first order, with dynamics

$$x(k+1) = c\ x(k) + d\ u(k) \tag{102}$$

It is desired to determine a control policy which minimizes the cost function

$$V_N(0) = \sum_{i=0}^{N-1} (x^2(i) + a\,u^2(i)) \tag{103}$$

Note that because the process operates in discrete time intervals, the integral cost function is replaced by a summation.

To begin, suppose the process runs for only one stage. Then we wish to choose $u(0)$ to minimize $x^2(0) + a\,u^2(0)$. The optimal choice is $u^*(0) = 0$, and the resulting cost is

$$V_1(0) = x^2(0) \tag{104}$$

Now suppose that $N = 2$. We must choose $u(0)$ and $u(1)$ so that

$$V_2(0) = \underset{u(0),u(1)}{\mathrm{Min}} \left[x^2(0) + a\,u^2(0) + x^2(1) + a\,u^2(1) \right] \quad \ldots \tag{105}$$

Since the first two terms on the right are not affected by u_1, (105) becomes

$$V_2(0) = \underset{u(0)}{\mathrm{Min}} \left[x^2(0) + a\,u^2(0) + \underset{u(1)}{\mathrm{Min}} (x^2(1) + a\,u^2(1)) \right] \quad \ldots \tag{106}$$

The second minimization (for a single-stage system) has already been performed. Replacing 0 by 1 in (104), we may re-write (106) as

$$V_2(0) = \underset{u(0)}{\mathrm{Min}} \left[x^2(0) + a\,u^2(0) + x^2(1) \right] \tag{107}$$

with $u^*(1) = 0$. Substitution of (102) into (107) gives

$$V_2(0) = \underset{u(0)}{\mathrm{Min}} \left[x^2(0) + a\,u^2(0) + (c\,x(0) + d\,u(0))^2 \right] \quad \ldots \tag{108}$$

Differentiation w.r.t. $u(0)$ yields

$$u^*(0) = -\frac{c\,d}{a + d^2}\, x(0) \tag{109}$$

Substitution of (109) in (108) shows that $V_2(0)$ is proportional to $x^2(0)$. Similarly at any stage k of an N-stage process

$$V_N(k) = g(k)\, x^2(k) \tag{110}$$

where g(k) is the constant of proportionality at stage k.

Now

$$V_N(k) = \min_{u(k) \ldots u(N-1)} \left[\sum_{i=k}^{N-1} (x^2(i) + a\, u^2(i)\,) \right] \tag{111}$$

Application of the principle of optimality allows us to express (111) in the recursive form

$$V_N(k) = \min_{u(k)} \left[x^2(k) + a\, u^2(k) + V_N(k+1) \right] \tag{112}$$

From (102) and (110) we have

$$g(k)\, x^2(k) = \min_{u(k)} \left[x^2(k) + a\, u^2(k) + g(k+1)(c\, x(k) + d\, u(k)\,)^2 \right] \quad \ldots \tag{113}$$

Minimization of (113) yields

$$u^*(k) = -\, h(k)\, x(k) \tag{114}$$

where $$h(k) = \frac{g(k+1)\ c\ d}{a + g(k+1) d^2} \tag{115}$$

The combination of (114) and (113) shows that

$$g(k) = 1 + a\, h^2(k) + g(k+1) \left[c - d\, h(k) \right]^2 \tag{116}$$

Equations (115) and (116) constitute a recursive method of calculating successive values of h(k) and g(k), working backwards from the values h(N-1) = 0 (since $u^*(N-1) = 0$) and g(N-1) = 1. The set $\left\{ h(k),\ k = 0, 1, 2, \ldots N-1 \right\}$, together with (114) defines an optimal control policy. Note the following points:

i. (114) defines a simple linear feedback control scheme;

ii. the solution applies to any initial value of x.

iii. there is no TPBVP.

A similar approach may be used when (102) is a vector-matrix equation describing an n^{th} order system. The development is identical in concept but all equations are in vector form.

5.7 Dynamic Programming: Continuous Time

Consider a system

$$\dot{\underline{x}} = \underline{f} = \underline{f}(\underline{x}, \underline{u}, t) \tag{117}$$

in which it is desired to determine $\underline{u}^*(\underline{x},t)$ to minimize a cost function

$$\int_0^T L(\underline{x}, \underline{u}, t)\, dt + V_T(x(t)\ )$$

$$\text{Let } V(\underline{x},t) = \operatorname*{Min}_{\underline{u}} \left[\int_t^T L(\underline{x}, \underline{u}, t)\, dt + V_T \right] \tag{118}$$

For a small time interval Δt, the principle of optimality may be applied to give

$$V(\underline{x},t) = \operatorname*{Min}_{\underline{u}} \left[L\,\Delta t + V(\underline{x} + \Delta x,\ t + \Delta t) \right] \tag{119}$$

If we assume that V is a continuous function of $\underline{x}$ and t, then a first-order expansion of the last term of (119) gives

$$V(\underline{x},t) = \operatorname*{Min}_{\underline{u}} \left[L\,\Delta t + V(\underline{x},t) + \frac{\partial V}{\partial t}\Delta t + \left(\sum_{i=1}^{n} \frac{\partial V}{\partial x_i}\ \frac{d\ x_i}{dt} \right) \Delta t \right] \tag{120}$$

Let

$$\frac{\partial V}{\partial \underline{x}} = \left(\frac{\partial V}{\partial x_1}, \frac{\partial V}{\partial x_2}, \quad \cdots \frac{\partial V}{\partial x_n} \right) = V_x \tag{121}$$

Now, letting $\Delta t \rightarrow 0$ and noting that the final term of (120) may be expressed as

$$< V_x, \underline{f} >$$

we have

$$\boxed{-\frac{\partial V}{\partial t} = \underset{\underline{u}}{\text{Min}} \left[L(\underline{x}, \underline{u}, t) + < V_x, \underline{f} > \right]} \tag{122}$$

which is known as Bellman's equation. The importance of this equation lies in the fact that if a function $V(\underline{x},t)$ can be found which satisfies (122), then an optimal feedback control policy $\underline{u}^*(\underline{x},t)$ can be derived from it; note that system linearity need not be assumed. No TPBVP is involved, as only the boundary condition $V(\underline{x},T) = V_T(\underline{x}(T))$ is required.

At the risk of being repetitious we shall emphasize again the difference between a pre-calculated optimal open-loop schedule $\left\{\underline{u}^*(t),\ 0 \leq t < T\right\}$, which results from the application of the calculus of variations or the maximum principle, and an optimal feedback control policy, $\left\{\underline{u}^*(\underline{x}, t),\ 0 \leq t < T\right\}$, obtained through dynamic programming. The latter is nearly always preferable since it is much less sensitive to noise (this, after all, is one of the prime reasons why any sort of feedback control is used). One seldom gets anything free, however, and the sad fact is that (122) is usually exceedingly difficult to solve, so much so that one is frequently forced to abandon it in favour of the lesser evil (computationally speaking) of the maximum principle with its TPBVP.

Equation (122) is a nonlinear partial differential equation. Its solution would be simpler if it could be modified to form an ordinary differential equation. To do this, we note from (118) that

$$\frac{dV}{dt} = - L\ (\underline{x},\ \underline{u}^*,\ t) \tag{123}$$

but

$$\frac{dV}{dt} = \frac{\partial V}{\partial t} + \frac{\partial V}{\partial \underline{x}}\ \frac{d\underline{x}}{dt} = \frac{\partial V}{\partial t} + \langle V_x,\ \underline{f} \rangle \tag{124}$$

Partial differentiation of (123) and (124) w.r.t. $\underline{x}$ yields a row-vector equation

$$\frac{\partial}{\partial \underline{x}} \left(\frac{\partial V}{\partial t} \right) + \underline{f}\ \ V_{xx} + V_x\ \underline{f}_x = - L_x \tag{125}$$

If the order of differentiation is immaterial, then the first two terms of (125) equal the time derivative of V_x. Thus

$$\dot{V}_x = - L_x - V_x\ \underline{f}_x \tag{126}$$

where $\underline{f}_x$ is the Jacobian matrix (64), and V_{xx} is the matrix of second partial derivatives of V. Now let us define a function

$$H'(\underline{x},\ V_x,\ t) = L\ (\underline{x},\ \underline{u}^*,\ t) + \langle V_x(\underline{x},t),\ \underline{f}(\underline{x},\underline{u}^*,t) \rangle \quad \ldots \tag{127}$$

so that (126) becomes

$$\dot{V}_x = - H'_x \tag{128}$$

Observe that (128) is identical in form to the Hamilton canonical equation (90). In Pontryagin's version of the maximum principle, though, $p_{n+1} = - 1$ (cf. (97) et seq.) and the Hamiltonian is

$$H = - L(\underline{x},\ \underline{u}^*,\ t) + \langle \underline{p}(x,t),\ \underline{f}\ (\underline{x},\ \underline{u}^*,\ t) \rangle \tag{129}$$

In the absence of control constraints, we may therefore equate

$$p_i = - \frac{\partial V}{\partial x_i} \tag{130}$$

to obtain

$$H = - H' \tag{131}$$

Thus the Hamiltonian (127) derived from dynamic programming is just the negative of Pontryagin's Hamiltonian; with the former, the maximum principle becomes the minimum principle, i.e., V is minimized by minimizing H' at all times. This is only a trivial variation on the maximum principle, but it must be borne in mind to avoid confusion. Generally speaking, Russian literature deals with the maximum principle and western literature with the minimum principle (although it is frequently called the maximum principle anyway in the latter case).

We may now re-write Bellman's equation, substituting (127), to obtain

$$- \frac{\partial V}{\partial t} = \min_{\underline{u}} \left[H' \, (\underline{x}, V_x, t) \right] \tag{132}$$

Both in this form and that of (122) this is also referred to as the Hamilton-Jacobi equation.

In summary, dynamic programming offers the possibility of computing an optimal feedback policy for a given dynamic system. Even when this is not computationally feasible, it constitutes an intuitively appealing method of deriving the maximum principle and (not shown here) the Euler-Lagrange equation. Moreover it is easily adapted to handle discrete time and stochastic systems.

5.8 Linear Systems with Quadratic Cost Functions

Having covered, or at any rate, skimmed, the field of optimization, we are legitimately entitled to ask whether this theory can really be used; i.e., can we do better with it what we now do with classical design methods such as root locus and compensation techniques? For linear systems the answer is yes, providing a digital computer is available. In this section we shall consider the application of dynamic programming to linear systems of the form

$$\dot{\underline{x}} = A\,\underline{x} + B\,\underline{u} \tag{133}$$

with quadratic cost functions

$$V = \int_0^T (\underline{x}^T Q\,\underline{x} + \underline{u}^T R\,\underline{u})\,dt \tag{134}$$

Note that the integrand in (134) is simply the multivariable form of a scalar function

$$q\,x^2 + r\,u^2 \tag{135}$$

Thus in a second order system the term $\underline{x}^T Q\,\underline{x}$ is

$$[x_1 \quad x_2] \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= q_{11}\,x_1^2 + (q_{12} + q_{21})\,x_1\,x_2 + q_{22}\,x_2^2$$

Without loss of generality, we may consider Q and R to be symmetric matrices. Also, it is assumed that Q is positive semidefinite and R is positive definite, which is to say that, for any vectors $\underline{x} \neq 0$, $\underline{u} \neq 0$,

$$\underline{x}^T Q\,\underline{x} \geq 0$$

$$\underline{u}^T R\,\underline{u} > 0$$

This is equivalent to specifying that q is nonnegative and r is positive in (135).

As previously discussed, all we need to do now is find a form of $V(\underline{x},t)$ which satisfies Bellman's equation (122), and derive an optimal feedback policy. For the system (133) with no constraints, and cost function (134) (by great good fortune a system which is at once useful and mathematically well-behaved) it can be shown that $V(\underline{x},t)$ has the form

$$V(\underline{x},t) = \underline{x}^T K(t)\, \underline{x} \tag{136}$$

where K(t) is a time-varying symmetric matrix whose elements are <u>not</u> functions of $\underline{x}$.

Recall Bellman's equation

$$-\frac{\partial V}{\partial t} = \underset{\underline{u}}{\text{Min}} \left[L + \langle V_x, \underline{f} \rangle \right]$$

We now substitute V, L and $\underline{f}$ from (136), (134) and (133) respectively:

$$-\underline{x}^T \dot{K}\, \underline{x} = \underset{\underline{u}}{\text{Min}} \Big[\underline{x}^T Q\, \underline{x} + \underline{u}^T R\, \underline{u}$$

$$+ (A\underline{x} + B\underline{u})^T K\underline{x} + \underline{x}^T K\, (A\underline{x} + B\underline{u}) \Big] \tag{137}$$

Setting $\frac{d}{d\underline{u}}$ (r.h.s.) = 0 gives

$$2\, R\, \underline{u} + 2\, B^T K\, \underline{x} = 0$$

$$\boxed{\underline{u}^*(t) = -\, R^{-1} B^T K(t)\, \underline{x}(t)} \tag{138}$$

Note from (138) that the optimal system is one using <u>linear feedback</u>. This is a very important result, since it specifies the structure of the controller without any a priori assumptions. All we need do now is find the unknown matrix K(t). To do this we substitute (138) into (137) to obtain

$$-\underline{x}^T \dot{K} \underline{x} = \underline{x}^T Q \underline{x} + \underline{x}^T K B R^{-1} B^T K \underline{x} + \underline{x}^T A^T K \underline{x}$$
$$- \underline{x}^T K B R^{-1} B^T K \underline{x} + \underline{x}^T K A \underline{x} - \underline{x}^T K B R^{-1} B^T K \underline{x}$$

$$\underline{x}^T \dot{K} \underline{x} = \underline{x}^T (K B R^{-1} B^T K - Q - A^T K - KA) \underline{x}$$

i.e. $$\boxed{\dot{K}(t) = K(t) B R^{-1} B^T K(t) - Q - A^T K(t) - K(t) A} \qquad (139)$$

Equation (139) is a form of the <u>matrix Riccati equation;</u> since the first term on the r.h.s. is quadratic in K, the equation is nonlinear. If, as in (134) there are no terminal costs, then the end point boundary value of K is K = 0 (null matrix). Equation (139) may then be integrated backwards in time so that a matrix function $K(t), 0 \leq t < T$ becomes available. Application of (138) yields the optimal feedback control policy.

Note that the development in this section closely parallels that for the scalar discrete time system in section 5.6. An important feature of the Riccati equation is that, going backwards in time K(t) reaches a steady state value, so that, provided the system operation is not near its endpoint, a stationary linear feedback mechanism is optimum. This stationary matrix is that which satisfies (139) when the l.h.s. is zero.

Example: We shall consider the design of an optimum position control system using an armature controlled d.c. motor. The motor transfer function is

$$\frac{C}{U}(s) = \frac{G}{s(1+\tau s)} \tag{140}$$

where $c(t)$ = shaft position (radians)

$u(t)$ = armature voltage

G = motor "gain" (rad. per sec. per volt)

τ = motor time constant (sec.)

The object of design is to minimize the tracking error without incurring either an excessive control input (causing saturation) or an excessive error rate. Suppose that the required output angle is zero but that the actual output angle is $c(t)$ and the shaft speed is $\dot{c}(t)$. What feedback policy will cause $c(t)$ to approach zero in an optimal fashion? We assume a cost function of the form

$$V = \int_0^\infty (\alpha\, x_1^2 + \beta\, x_2^2 + \gamma\, u^2)\, dt \tag{141}$$

where

$$\begin{aligned} x_1 &= c \\ x_2 &= \dot{c} \end{aligned} \tag{142}$$

Clearly the designer's choice of α, β, and γ will affect the parameters of the solution. For instance if γ is small relative to α and β, then large input voltages may be called for because the cost of control effort is relatively small. The choice of suitable values for the parameters of a cost function is not always simple, and remains in many cases an engineering art. Let us assume that we have chosen α, β, and γ and proceed.

From (140) and (142) we see that the state equation is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -\frac{1}{\tau} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{G}{\tau} \end{bmatrix} u \tag{143}$$

Inspection of (141) and (143) shows that

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -1/\tau \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ G/\tau \end{bmatrix}$$

$$Q = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \qquad R = \gamma$$

The Ricatti equation (139) is thus

$$\begin{bmatrix} \dot{k}_{11} & \dot{k}_{12} \\ \dot{k}_{12} & \dot{k}_{22} \end{bmatrix} = \frac{1}{\gamma} \begin{bmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{bmatrix} \begin{bmatrix} 0 \\ G/\tau \end{bmatrix} [0 \quad G/\tau] \begin{bmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{bmatrix}$$

$$- \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 1 & -1/\tau \end{bmatrix} \begin{bmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{bmatrix} - \begin{bmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & -1/\tau \end{bmatrix} \quad \ldots \tag{144}$$

Note that since K(t) is symmetric, $k_{12} = k_{21}$. Multiplication of the matrices in (144) yields three simultaneous nonlinear ordinary differential equations

$$\dot{k}_{11} = \frac{1}{\gamma}\left(\frac{k_{12}G}{\tau}\right)^2 - \alpha \tag{145}$$

$$\dot{k}_{12} = \frac{k_{11}k_{12}}{\gamma}\left(\frac{G}{\tau}\right)^2 - k_{11} + \frac{k_{12}}{\tau} \tag{146}$$

$$\dot{k}_{22} = \frac{1}{\gamma}\left(\frac{k_{22}G}{\tau}\right)^2 - \beta - 2k_{12} + 2\frac{k_{22}}{\tau} \tag{147}$$

Boundary conditions are $k_{11}(T) = k_{12}(T) = k_{22}(T) = 0$. The fact that $T = \infty$ in (141) need not bother us; this merely means that, as explained previously, the system is not near its endpoint in time, and we are looking for stationary solutions to (145) - (147). There are two ways of obtaining them:

i) begin at the boundary conditions and integrate backwards in time numerically until k_{11}, k_{12} and k_{13} are stationary;

ii) set the left hand sides equal to zero and solve the resultant simultaneous algebraic equations by successive approximation.

Once this solution is obtained, we take the final step of applying (138)

$$u^* = - R^{-1} B^T K \underline{x}$$

$$u^* = - \frac{1}{\gamma} \quad [0 \; G/\tau] \begin{bmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= - \frac{1}{\gamma} \left[\frac{k_{12}G}{\tau} \quad \frac{k_{22}G}{\tau} \right] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

to obtain the optimal feedback policy

$$u^* = - \frac{k_{12}G}{\gamma\tau} \; x_1 \; - \frac{k_{22}G}{\gamma\tau} \; x_2 \tag{148}$$

The resulting system configuration is shown in fig. 4.

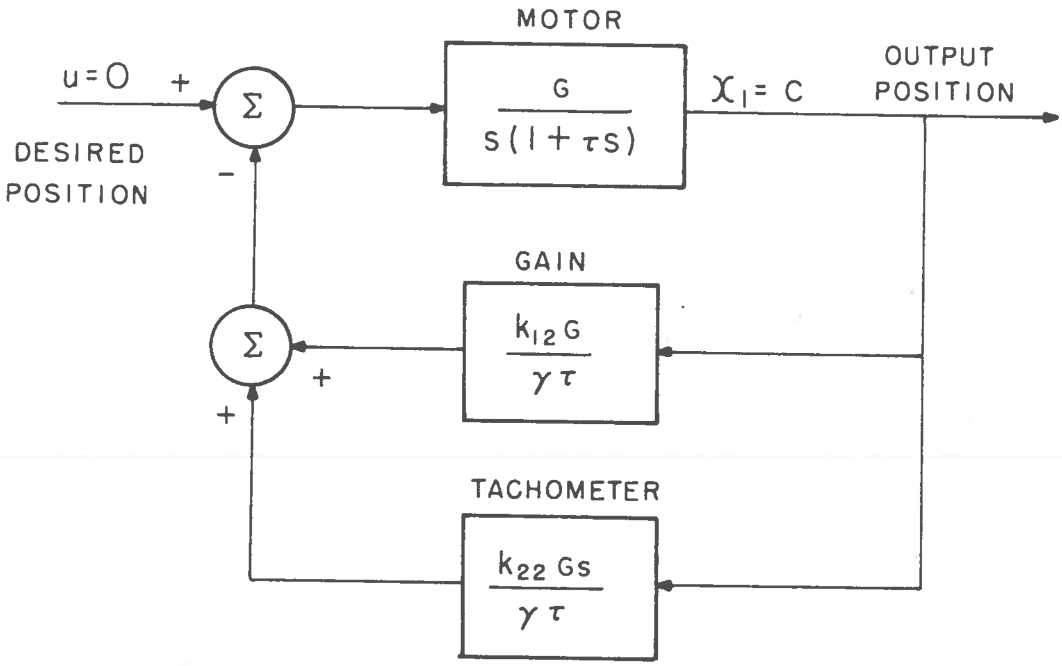

Figure 4 Optimal system configuration

The form of the optimal solution leads us to the observation that, in general, an optimal position control system employs both position feedback and rate feedback. Surprising? Hardly; this is just proportional plus derivative feedback, or equivalently, the use of lead network compensation. Nonetheless this example does indicate that the theory of optimal control is not wholly divorced from the real world; it is, in fact, a useful design tool. Generalizing further on equation (138), we see that an optimal control system employs feedback of every system state. We note that some of the feedback of coefficients may be small or zero, though.

The full advantages of optimal design are not apparent in this second order example. When the system is of higher order, and especially when it is a multi-input system, the advantages of this straightforward computer-oriented design method over traditional compensation techniques become quite clear. Many such examples exist in the literature.

6. CONCLUSION

The purpose of this report has been to introduce a number of concepts which constitute the foundations of automatic control theory. Obviously, mastery of these methods requires a very much more detailed study than is presented here; the interested reader is referred to the bibliography. To conclude the survey, we shall emerge from the foundations and cast a brief glance at a few items in the superstructure; i.e., topics which are of current research interest.

1. Stochastic Systems

When disturbances are present it is always desirable to exercise some form of closed loop control. The computation of suitable control policies for stochastic systems is ususally very difficult, however. The only case which has been treated extensively is the linear-quadratic system of section 5.8 in which additive gaussian noise is present.

2. Filtering and Estimation

An important aspect of the control of stochastic systems is the estimation of state variables since these are frequently either not accessible, or can be measured only with noisy transducers.

3. Identification

This area includes the problems of modelling and parameter estimation, that is: a) the choice of a suitable model structure which is neither too complicated for computer manipulation nor too gross an approximation of the real system, and b) the estimation of the model parameters from system operating records.

4. Adaptive Control

The combination of optimal control theory with areas 2) and 3) above allows us in principle to design an adaptive controller; i.e., one which attempts to optimize system performance in the face of uncertain and/or time-varying parameters. The advantage of this approach is that a fairly simple adaptive controller may give results which are equal or even superior to those obtained with a complex non-adaptive controller. On the other hand, the combination of control, estimation, and identification usually yields a problem which is, theoretically at least, greater than the sum of the three sub-problems because of interaction between the system cost functions and the estimation strategy.

5. Learning Control

The concept of "learning", as applied to control systems, comprises an ill-defined area usually taken to be more general in nature than adaptive control. Here the emphasis is on the evolution of a desirable pattern of behaviour (on the part of the controller) through generalization of the results of experiments on the system. Mathematical techniques used include stochastic approximation, decision theory, automata theory, and pattern recognition methods.

7. BIBLIOGRAPHY

This bibliography is selective and subjective, rather than exhaustive.

1. BOOKS

1.1 Classical Linear Control Systems

There exist scores of books dealing with the classical methods of Bode diagrams and root locus. Some recent readable ones are:

J. J. D'Azzo, C. H. Houpis, "Feedback Control Systems Analysis and Synthesis" 2nd ed., McGraw-Hill, 1966.

R. C. Dorf, "Modern Control Systems", Addison-Wesley, 1967.

B. C. Kuo, "Automatic Control Systems", 2nd. ed., Prentice-Hall, 1967.

1.2 Sampled Data Control Systems

E. J. Jury, "Theory and Application of the Z-Transform Method", Wiley, 1964.

B. C. Kuo, "Analysis and Synthesis of Sampled Data Control Systems", Prentice-Hall, 1963.

D. P. Lindorff, "Theory of Sampled-Data Control Systems", Wiley, 1965.

1.3 Nonlinear Systems

J. E. Gibson, "Nonlinear Automatic Control", McGraw-Hill, 1963.

D. Graham, D. McRuer, "Analysis of Nonlinear Control Systems", Wiley, 1961.

1.4 State Space Concepts

The best introductory text in this field is probably:

P. M. DeRusso, R. J. Roy, C. M. Close, "State Variables for Engineers", Wiley, 1965.

SEE ALSO:

L. K. Timothy, B. E. Bona, "State Space Analysis: An Introduction", McGraw-Hill, 1968.

Useful as a reference is:

Z. A. Zadeh, C. A. Desoer, "Linear System Theory: the State Space Approach", McGraw-Hill, 1963.

1.5 Static Optimization

Two well written books on hill-climbing methods are:

D. J. Wilde, "Optimum Seeking Methods", Prentice-Hall, 1964.

D. J. Wilde, C. S. Beightler, "Foundations of Optimization", Prentice Hall, 1967.

A simple introduction to linear programming is given in a soft cover publication

W. A. Spivey, "Linear Programming: An Introduction", Macmillan, 1963.

The theory of linear and nonlinear programming is treated by:

G. Hadley, "Linear Programming", Addison-Wesley, 1962.

G. Hadley, "Nonlinear and Dynamic Programming",Addison-Wesley, 1964.

1.6 Optimal Control

No completely satisfactory introductory text exists in this field. The following are useful, however:

V. W. Eveleigh, "Adaptive Control and Optimization Techniques", McGraw-Hill, 1967.

A. R. M. Noton, "Introduction to Variational Methods in Control Engineering", Pergamon, 1965.

D. G. Schultz, J. L. Melsa, "State Functions and Linear Control Systems", McGraw-Hill, 1967.

J. T. Tou, "Modern Control Theory", McGraw-Hill, 1964.

Several books which are oriented towards applications are:

L. Lapidus, R. Luus, "Optimal Control of Engineering Processes", Blaisdell, 1967.

E. S. Savas, "Computer Control of Industrial Processes", McGraw-Hill, 1965.

S. M. Shinners, "Techniques of System Engineering", McGraw-Hill, 1967.

More advanced books on the theory of optimal control are:

M. Athans, P. L. Falb, "Optimal Control", McGraw-Hill, 1966.

E. B. Lee, L. Markus, "Foundations of Optimal Control Theory", Wiley, 1967.

C. W. Merriam, "Optimization Theory and the Design of Feedback Control Systems", McGraw-Hill, 1964.

L. S. Pontryagin et al,, "The Mathematical Theory of Optimal Processes", Interscience, 1962.

The methods of functional analysis, increasingly used in the control literature, are introduced in:

W. A. Porter, "Modern Foundations of Systems Engineering", Macmillan, 1966.

2. PERIODICALS

2.1 Theory

Automatica; (Pergamon) this is the official journal of the International Federation of Automatic Control (IFAC)

Automation and Remote Control; translation of Automatika i Telemekhanika (Russian).

Automatisme (French)

Engineering Cybernetics; translation of Izvestiia. Tekhnicheskaia Kibernetika (Russian)

Information and Control

International Journal of Control

IEEE Transactions on Automatic Control

IEEE Transactions on Systems Science and Cybernetics

Proc. IEE (Control and Science Section)

Society of Industrial and Applied Mathematics (SIAM) Journal of Control

2.2 Applications

Canadian Controls and Instrumentation

Control (U.K)

Control Engineering (U.S.A.)

An annual publication of interest is:

C. F. Leondes, ed., "Advances in Control Systems", Academic Press.

3. ABSTRACTS

Automation Express (International Physical Index, Inc.) a digest of current Soviet control literature.

Control Abstracts (IEE)

Current Papers on Control (IEE); titles only.

4. CONFERENCES

Allerton Conference on Circuit and Systems Theory (Annual, USA)

Asilomar Conference on Circuit and Systems Theory (Annual, USA)

Canadian National Conference on Automatic Control (aperiodic)

The International Federation of Automatic Control, of which Canada is a member, holds triennial congresses whose proceedings cover all aspects of control theory and application. In addition several symposia are held each year on more specialized topics, e.g., sensitivity, identification, digital control, multivariable systems.

Joint Automatic Control Conference (Annual, USA)

National Electronics Conference (Annual, USA)

United Kingdom Automation Council Conference (Annual, U.K.)

5. SUMMER COURSES

Each year a number of summer courses on Automatic Control, usually of one or two weeks' duration, are offered by various universities in Canada and the U. S. A. These are advertised in journals such as IEEE"Spectrum".