Optimal & Adaptive Control

Texas Tech University (1976)

R.J. Marks II Class Notes

BOB MARKS
2232 AUBURN #6
LUBBOCK, TEXAS
79415

763-6288

OPTIMAL & ADAPTIVE
CONTROL SYSTEMS
EE5312

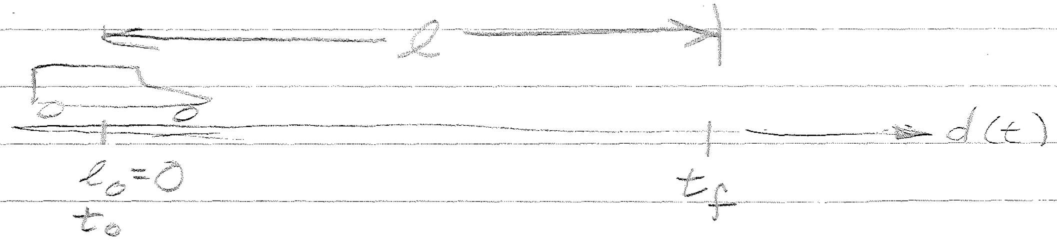
MON	TUES	WED	THURS	FRI
	MA 5383 FL&M 22 DURAN 9-10:30		MA 5383 FL&M 22 DUAN 9-10:30	
EE 5312 CHAO 9:30-10:30	EE 5360 RM 251 WALKUP 10:30-12	EE 5312 CHAO 9:30-10:30	EE 5360 RM 251 WALKUP 10:30-12	EE 5312 CHAO 9:30-10:30
10:30 11:30		10:30 12:30		10:30 11:30
EE 4362 RM 156 WALKUP 12:30-1:30		EE 4362 RM 156 WALKUP 12:30 -1:30		EE 4362 RM 156 WALKUP 12:30 -1:30
EE 233 SEC 2 RM 156 MARKS 11:30-12:30				EE 233 SEC 2 RM 156 MARKS 11:30-12:30
GERM 533 119 FL&M JARDINE 1:30-2:30		GERM 533 119 FL&M JARDINE 1:30-2:30		GERM 533 119 FL&M JARDINE 1:30-2:30

(iii) DESIRED OUTPUT (STATES)

(iv) PHYSICAL CONSTRAINTS

(V) PERFORMANCE MEASURE

EXAMPLE:



$$\alpha(t) + \beta(t) = \frac{d^2 d(t)}{dt^2}$$

ACCELERATION

BREAKING

WITH TO USE "NORMAL FORM EQUATION"

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

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad U = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}$$

$$x_1 = d \qquad \Longrightarrow \qquad \dot{x}_1 = x_2$$

$$x_2 = d\,d(t)/dt \qquad \dot{x}_2 = \alpha(t) + \beta(t)$$

STATE EQ.:

$$\dot{x} = Ax + BU$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha(t) \\ \beta(t) \end{bmatrix}$$

WHEN "A" MATRIX

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \end{bmatrix} \Longleftarrow \begin{array}{l} \text{PHASE} \\ \text{VARIABLE} \\ \text{FORM} \end{array}$$

CONSTRAINTS ON INPUT

$$\left.\begin{array}{l} 0 \leq \alpha(t) \leq M \\ -N \leq \beta(t) \leq 0 \end{array}\right\} \longrightarrow \begin{array}{l} \text{ADMISSABLE CONTROL} \\ \text{ie } U \in \underline{U} \end{array}$$

CONSTRAINTS ON X (STATES)

$$\left\{\begin{array}{l} x_1(t_0) = l_0 = 0 \\ x_1(t_f) = l \end{array}\right.$$

X_2 = VELOCITY

$$\left\{\begin{array}{l} X_2(t_0) = 0 \\ X_2(t_f) = 0 \end{array}\right.$$

IF CAR ONLY GOES FORWARD

$$0 \leq X(t) \leq l$$

$$0 \leq X_2(t)$$

SOLUTIONS SATISFYING THESE ARE "ADMISSABLE STATE TRAJECTORIES"

ASSUME FUEL CONSUMPTION IS PROP. TO VELOCITY. (RATE)

FUEL CONSUMPTION

$$= \int_{t_0}^{t} \left[k_1 d(t) + k_2 x_2(t) \right] dt \leq G =$$

PERFORMANCE MEASURE:

$$J = t_f - t_0$$

IN GENERAL

$$J = \underbrace{h(t_f, x(t_f))}_{\text{FINAL PENALTY FUNCTION}} + \int_{t_0}^{t_f} g\left[x(t), u(t), t\right] dt$$

SAY YOU FIND SOME OPTIMAL CONTROL $u^*(t) \ni$

$$u^*(t) = U\left[x(t), t\right]$$

← FEEDBACK OR

⇐ OPEN LOOP CONTROL

9-1-76 (WED)

STATE VARIABLE REPRESENTATION
OF SYSTEMS

NONLINEAR

$$\begin{cases} \dot{x} = f[\overset{\text{STATE}}{x}, \overset{\text{CONTROL}}{u}, \overset{\text{TIME}}{t}] \Leftarrow \text{STATE EQ. IN NORMAL FORM} \\ y = g[x, u, t] \Leftarrow \text{OUTPUT} \end{cases}$$

FOR A LINEAR SYSTEM

$$\begin{cases} \dot{x} = Ax + BU \quad ; x(t_0) = x_0 \\ y = Cx + DU \end{cases}$$

SOLUTION IS

$$x(t) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, \tau)\, B(\tau)\, u(\tau)\, d\tau$$

$\Phi(t, t_0)$ = STATE XSITION MATRIX

IT SATISFIES:

$\dot{\Phi}(t) = A\,\Phi(t)$ (NOT UNIQUE)

$\Phi(t)$ IS THE "FUNDEMENTAL MATRIX)

(i) $\Phi(t_0, t_0) = I$

(ii) $\Phi(t_1, t_2)\,\Phi(t_2, t_3) = \Phi(t_1, t_3)$

(iii) $\Phi(t_1, t_2) = \Phi^{-1}(t_2, t_1)$

CONSIDER, THEN

$\dot{x} = Ax \Longrightarrow x(t) = \Phi(t, t_0)\,x(t_0)$

IT FOLLOWS THAT

$x(t_1) = \Phi(t_1, t_0)\,x(t_0)$

$x(t_1) = \Phi(t_1, t_2)\,x(t_2)$

$x(t_2) = \Phi(t_2, t_3)\,x(t_3)$

COMBING ABOVE TWO GIVES

$x(t_1) = \Phi(t_1, t_2)\,\Phi(t_2, t_3)\,x(t_3)$

$x(t_1) = \Phi(t_1, t_3)\,x(t_3)$

OR, WE HAVE SHOWN

$$\Phi(t_1, t_3) = \Phi(t_1, t_2)\,\Phi(t_2, t_3)$$

PROPERTY iii IMMEDIATELY FOLLOWS

FOR LINEAR TIME-INVARIANT SYSTEMS:

$$x(t) = \Phi(t-t_0)x_0 + \int_{t_0}^{t} \Phi(t-\tau)B(\tau)u(\tau)d\tau$$

FOR $t_0 = 0$

$$x(t) = \Phi(t)x_0 + \int_{t_0}^{t} \Phi(t-\tau)B(\tau)u(\tau)d\tau$$

RECALL

$$\dot{\Phi}(t) = A\Phi(t) \Rightarrow \Phi(t) = e^{At}$$

$$= I + At + \frac{A^2}{2!}t^2 + \frac{1}{3!}A^3t^3 + \ldots$$

ANOTHER WAY USES LAPLASE XFORM:

$$sX(s) - x(0) = AX(s) + BU(s)$$

$$(sI-A)X(s) = x_0 + BU(s)$$

$$\Rightarrow X(s) = (sI-A)^{-1}[x(0) + BU(s)]$$

OR

$$x(t) = \mathcal{L}^{-1}[(sI-A)^{-1}x(0)] + \mathcal{L}^{-1}[(sI-A)BU(s)]$$

CONSIDER

$$\mathcal{L}^{-1}(sI-A)^{-1} = \mathcal{L}^{-1}\left[\frac{1}{s}\left(I-\frac{A}{s}\right)^{-1}\right]$$

$$= \mathcal{L}^{-1}\left[\frac{1}{s}\left(I + \frac{A}{s} + \frac{A^2}{s^2} + \ldots\right)\right]$$

$$= \mathcal{L}^{-1}\left[I/s + A/s^2 + A^2/s^3 + \ldots\right]$$

$$= \left(I + At + A^2\frac{t^2}{2} + \frac{1}{3!}A^3t^3 + \ldots\right)\cancel{u(t)}$$

THUS

$$x(t) = (I + At + \ldots)x_0 + \int_{t_0}^{t} \Phi(t-\tau)B \quad u(\tau)d\tau$$

NOW $\mathcal{L}\int_0^t f_1(t-\tau)f_2(\tau)d\tau = F_1(s)F_2(s)$

THUS

$$\mathcal{L}[\Phi(t)] = (sI-A)^{-1} = \mathcal{L}[e^{At}]$$

OR

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

AN ALTERNATE METHOD

(ii) $\underline{\Phi}(t) = e^{At}$

MAY DIAGONALIZE A MATRIX. i.e, FIND J

$\ni AP = PJ$

$$P = [\overset{\lambda_1}{P_1}\ \overset{\lambda_2}{P_2} \ldots \overset{\lambda_n}{P_n}] = \text{ASSOCIATED EIGEN VECTOR}$$

HERE, $A = PJP^{-1}$

$$\Rightarrow \underline{\Phi}(t) = e^{PJP^{-1}t}$$

$$= I + PJP^{-1}t + \frac{1}{2}(PJP^{-1})(PJP^{-1})t^2 + \frac{1}{3!}(PJP^{-1})^3 t^3 + \ldots$$

$$= P\left[I + Jt + \frac{1}{2} J^2 t^2 + \frac{1}{3!} J^3 t^3 \ldots \right] P^{-1}$$

$$= P e^{Jt} P^{-1}$$

J IS A DIAGONALIZED MATRIX.

THEN

$$\underline{\Phi}(t) = P \begin{bmatrix} e^{\lambda_1 t} & & 0 \\ & e^{\lambda_2 t} & \\ & & \ddots & \\ 0 & & & e^{\lambda_n t} \end{bmatrix} P^{-1}$$

WHICH IS CLOSED FORM

STATE VARIABLE REPRESENTATION
OF SYSTEMS

CONTROLLABILITY

CONSIDER

$$\dot{x} = f(x, u, t), \quad x(t_0) = x_0$$

IF ONE CAN CHOOSE A u TO MOVE x_0 TO A DESIRED FINAL STATE, THEN THE SYSTEM IN CONT.

COMPLETELY STATE CONTROLLABLE:

CONTROLLABLE $\forall\ x_0$. THAT IS, A SYSTEM IS SAID TO BE COMPLETELY STATE-CONTROLLABLE IF $\forall\ t_0$, EACH INITIAL STATE $x(t_0)$ CAN BE XFERRED TO ANY FINAL STATE $x(t_f)$ IN ANY FINITE TIME $(t_f - t_0)$ WITH A BOUNDED $u(t)$. i.e., $\|u(t)\| < \infty$

9-3-76 (FRI)

CONTROLLABILITY IN LINEAR SYSTEMS

IN LINEAR SYSTEM:

$$\dot{x}(t) = A(t)x(t) + B(t)u(\tau)$$

SOLUTION: $x(t) = \Phi(t,t_0)x(t_0) + \int_{t_0}^{t} \Phi(t,\tau)B(\tau)u(\tau)d\tau$

@ $x(t_f) = \underbrace{\Phi(t_f,t_0)x(t_0)}_{x^0(t_f)} + \int_{t_0}^{t_f} \Phi(t_f,\tau)B(\tau)u(\tau)d\tau$

(LOOKING AT STATE CONTROL)

$$\Rightarrow x(t_f) - x^0(t_f) = \int_{t_0}^{t_f} \Phi(t_f,\tau)B(\tau)u(\tau)d\tau$$

$$= x^d(t_f) = \int_{t_0}^{t_f} \mathcal{H}_x(t_f,\tau)u(\tau)d\tau$$

(WE ARE XLATING ORIGIN TO FINAL STATE)

THM: THE SYSTEM

$$\dot{x} = Ax + Bu$$
$$y = Cx$$

IS COMPLETELY STATE CONTROLLABLE ON $[t_0, t_f]$ IFF:

$$M_0(t_0,t_f) = \int_{t_0}^{t_f} H_x(t_f,\tau)H_x^T(t_f,\tau)d\tau$$

IS NON-SINGULAR.

(IS COMPLETELY STATE CONTROLLABLE IF THIS IS TRUE ∀ TIME INTERVAL $[t_0, t_f]$)

PROOF: (SUFFICIENCY)

LET: $u(\tau) = H_x^T(t_f,\tau)\lambda \neq 0$ ← COLUMN VECTOR

$$x^d(t_f) = \int_{t_0}^{t_f} \mathcal{H}(t_f,\tau)\,u(\tau)d\tau$$

THUS

$$x^d(t_f) = \int_{t_0}^{t_f} \mathcal{H}(t_f,\tau)\mathcal{H}^T(t_f,\tau)\lambda\, d\tau$$

$$= \int_{t_0}^{t_f} \mathcal{H}(t_f,\tau)\mathcal{H}^T(t_f,\tau)\,d\tau\,\lambda$$

IF $M_0(t_0,t_f)$ IS NON-SINGULAR,

THEN ∃

$$\lambda = M_0^{-1}(t_0,t_f)\,x^d(t_f)$$

QED FOR SUFFICIENCY

(NECESSITY)

PROOF BY COUNTERDICTION

ASSUME $M_o(t_o, t_f)$ IS SINGULAR

AND THE SYSTEM IS STATE CONTROLLABLE

THEN $\exists\ C \ni$

$$M_o C = 0 \ni C \neq 0$$

TAKING TRANSPOSES

$$C^T M_o^T = 0$$

$M_o = M_o^T \Rightarrow$ SYMMETRIC

SINCE M_o IS SYMMETRIC WE HAVE

$$C^T M_o = 0$$

OR, $\exists\ C \neq 0 \ni$

$$C^T \int_{t_o}^{t_f} H_x(t_f, \tau) H_x^T(t_f, \tau)\, d\tau = 0$$

FROM

$$M_o = \int_{t_o}^{t_f} H_x(t_f, \tau)\, H_x^T(t_f, \tau)\, d\tau$$

OR

$$C^T M_o C = \int_{t_o}^{t_f} C^T H_x(t_f, \tau) H_x^T(t_f, \tau)\, C\, d\tau$$

$$= \int_{t_o}^{t_f} \left[H_x^T(t_f, \tau) C\right]^T \underbrace{\left[\overset{m \times n}{H^T}(t_f, \tau)\overset{n \cdot 1}{C}\right]}_{m \times 1\ (\text{COLUMN})} d\tau$$

$$\triangleq \int_{t_o}^{t_f} \eta^T \eta\, d\tau$$

$$\ni \eta = H^T(t_f, \tau)\, C$$

$$= \int_{t_o}^{t_f} (\eta_1^2 + \eta_2^2 + \dots + \eta_m^2)\, d\tau$$

$$\eta = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_m \end{bmatrix}$$

THUS $C^T M_0 C \Rightarrow$ POS DEF. MATRIX

$$C^T M_0 C \geq 0$$

$$= 0 \text{ IFF } \eta_i = 0 \ \forall i$$

<u>BUT</u> $C^T M_0 = 0$

$$\Rightarrow \eta = H_x^T(t_f, \tau) C = 0$$

FROM THE STATE CONTROLLABILITY, WE HAVE

$$x(t_f) - x^0(t_0) = \int_{t_0}^{t_f} H_x(t_f, \tau) u(\tau) d\tau$$

MUST BE TRUE $\forall \ x(t_f)$

THUS, LET $x(t_f) = C$.

ALSO, LET $x(t_0) = 0$. THEN

$$C = \int_{t_0}^{t_f} H_x(t_f, \tau) u(\tau) d\tau$$

$$C^T C = \int_{t_0}^{t_f} \underbrace{C^T H_x(t_f, \tau)}_{\overset{\parallel}{0}} u(\tau) d\tau$$

THUS

$$C^T C = 0$$

$$\Rightarrow C_1 = C_2 = C_3 = C_4 = \ldots = 0$$

THUS, WE HAVE A COUNTERDICTION

OUTPUT CONTROLLABILITY

$$Y = CX$$
$$= C\Big[\underbrace{\Phi(t,t_0)x(t_0)}_{Y_0} + \underbrace{\int_{t_0}^{t_f} \Phi(t,\tau)B(\tau)u(\tau)d\tau}_{C^{-1}H_Y(t,\tau)}\Big]$$

$$\Rightarrow Y(t) = Y_0 + \int_{t_0}^{t_f} H_Y(t,\tau)u(\tau)d\tau$$

THUS, THE SYSTEM IS OUTPUT CONTCLLABLE IN THE SENSE

$\exists \; N_0(t_f, t_0) = \int_{t_0}^{t_f} H_Y(t_f,\tau)H_Y^T(t_f,\tau)d\tau$

THEN THE SYSTEM IS COMPLETELY OUTPUT CONTROLLABLE ON $[t_0, t_f]$ IF

$$N_0(t_f, t_0) = \int_{t_0}^{t_f} H_Y(t_f,\tau)H_Y^T(t_f,\tau)d\tau$$

PROOF FOLLOWS AS PREVIOUSLY

NOW

$M_o(t_f, t_o)$ IS NON-SINGULAR THEN THE COLUMN VECTORS OF $H_x^T(t,\tau)C$ ARE ALL LINEARLY INDEPENDENT

$$\eta = H_x^T(t_f, \tau)C \neq 0$$

OR $C^T H_x^T(t_f, \tau) \neq 0$ ON $t_0 < \tau < t_f \; \forall C \neq 0$

NOW

$$\frac{d}{d\tau}\Phi(t, \tau) = -\Phi(t, \tau)A(\tau)$$

9/8/76 (WED)

HOMEWORK #1

1-1, -3, -6, -9, -12 (a)(i)(f), -14 (a)(c)(e)

DUE WEDNESDAY (9/15)

CONTROLLABILITY

$\dot{x} = Ax + Bx \qquad y = Cx$

$M_o(t_o, t_f) = \int_{t_o}^{t_f} H_x(t_f, \tau) H_x^T(t_f, \tau)\, d\tau$

$H_x(t_f, \tau) = \Phi(t_f, \tau)B(\tau)$

CONTROL. IF M_o IS NON-SINGULAR

(COLUMNS OF $H_x^T(t_f, \tau)$ LINEARLY IND)

$\Rightarrow H_x^T(t_f, \tau)\eta \neq 0 \quad \forall \; \eta \neq 0$

$\eta^T H_x(t_f, \tau) \neq 0$

CONSIDER

$\eta^T H_x(t_f, \tau) = 0$

$= \eta^T \Phi(t_f, \tau) B(\tau) = 0$

~~DIFFERENTIATING~~ $\Longrightarrow$

WE KNOW

$$\frac{d\,\Phi(t_f,\tau)}{d\tau} = -\,\Phi(t_f,\tau)\,A(\tau)$$

$$\frac{d\,\Phi(t,t_0)}{dt} = A(t)\,\Phi(t,t_0)$$

$\Phi(t,\tau)$ IS ALWAYS NONSINGULAR

$$\Phi(t_0,t_0) = \Phi(\tau,\tau) = I$$

$$\Phi(t_1,t_2)\,\Phi(t_2,t_3) = \Phi(t_1,t_3)$$

THEN $\Phi(\tau,t)\,\Phi(t,\tau) = I$

DIFFERENTIATING

$$\frac{d\,\Phi(\tau,t)}{d\tau}\,\Phi(t,\tau) + \Phi(\tau,t)\frac{d\Phi}{d\tau} = \frac{d}{d\tau}I = 0$$

$$A(\tau)\,\Phi(\tau,t)\,\Phi(t,\tau) + \Phi(\tau,t)\,\frac{d\,\Phi(t,\tau)}{d\tau} = 0$$

$$A(\tau) + \Phi(\tau,t)\,\frac{d\,\Phi(t,\tau)}{d\tau} = 0$$

OR

$$\Phi(\tau,t)\frac{d}{d\tau}\Phi(t,\tau) = -A(\tau)$$

$$\frac{d}{d\tau}\Phi(t,\tau) = -\,\Phi^{-1}(\tau,t)\,A(\tau)$$

$$= -\,\Phi(t,\tau)\,A(\tau)$$

LET'S NOW DIFFERENTIATE

$$\eta^T \Phi(t_f, \tau) B(\tau) = 0 \quad (\text{ACTUALLY} \neq 0)$$

$$\eta^T \left[\dot{\Phi}(t_f, \tau) B(\tau) + \eta^T \Phi(t_f, \tau) \dot{B}(\tau) \right] = 0$$

$$\eta^T \left[-\Phi(t_f, \tau) A(\tau) B(\tau) + \Phi(t_f, \tau) \dot{B}(\tau) \right] = 0$$

$$\eta^T \Phi(t_f, \tau) \left[\dot{B}(\tau) - A(\tau) B(\tau) \right] = 0$$

LET $\Gamma_1(\tau) = B(\tau)$

THEN $\eta^T \Phi(t_f, \tau) B(\tau) = 0$

$\Gamma_2(\tau) = \dot{\Gamma}_1(\tau) - A(\tau) \Gamma_1(\tau)$

THEN $\eta^T \Phi(t_f, \tau) \Gamma_2(\tau) = 0$

DIFFERENTIATING AGAIN

$$\eta^T \Phi(t_f, \tau) \Gamma_3(\tau) = \eta^T \Phi(t_f, \tau) \left[\dot{\Gamma}_2(\tau) - A(\tau) \Gamma_2(\tau) \right] = 0$$

DIFFERENTIATING $n-1$ TIMES GIVES

$$\eta^T \Phi(t_f, \tau) \Gamma_n(\tau) = 0$$

$$\Rightarrow \Gamma_n(\tau) = \dot{\Gamma}_{n-1}(\tau) - A(\tau) \Gamma_{n-1}(\tau)$$

OR, IN GENERAL, WE HAVE RECURSION

$$\Gamma_{K+1} = \dot{\Gamma}_K(\tau) - A(\tau) \Gamma_K \qquad K = 1, \dots, n-1$$

$$\Gamma_1(\tau) = B(\tau)$$

COMBINING THE DERIVATIVES

$$\lambda^T \Phi(t_f, \tau) \left[\Gamma_1 \vdots \Gamma_2 \vdots \cdots \vdots \Gamma_n \right] = [0]$$

ACTUALLY, IF WE CARRIED $\neq$, WE'D HAVE

$$\lambda^T \Phi(t_f, \tau) \left[\Gamma_1 \vdots \Gamma_2 \vdots \cdots \vdots \Gamma_n \right] \neq [0]$$

$\therefore$ THE COLUMNS OF THE $[\Gamma_1 \ldots \Gamma_n]$ MUST BE LINEARLY INDEPENDENT. (ie, OF RANK n). $[\Gamma_1 \ldots \Gamma_n]$ IS A $n \times nm$ MATRIX AND IS CALLED THE CONTROLLABILITY MATRIX. BY FULL RANK, WE NEED n COLUMNS TO BE LINEAR INDEPEND. FOR 1 INPUT $(m=1)$, Γ IS A SQUARE MATRIX

FOR LINEAR TIME INVARIANT SYSTEM, $\Gamma_1 = B$, $\Gamma_2 = -AB$, $\Gamma_3 = A^2B$, $\Gamma_4 = -A^3$ ETC.

$$\Rightarrow \Gamma = \left[B \vdots -AB \vdots A^2B \cdots (-A)^{n-1} B \right]$$

OBSERVABILITY

(DUAL OF CONTROLABILITY)

$$\dot{x} = Ax + Bu \qquad (*)$$
$$y = Cx$$

DEF: THE SYSTEM (*) IS SAID TO BE COMPLETELY OBSERVABLE ON $[t_0, t_f]$ IF, FOR SPECIFIED $t_0 \neq t_f$, THE INITIAL STATE $x(t_0) = x_0$ OF THE UNFORCED SYSTEM CAN BE DETERMINED FROM THE KNOWLEDGE OF $y(t)$ ON $[t_0, t_f]$

THEM: THE SYSTEM (*) IS COMPLETELY OBSERVABLE ON INTERVAL $[t_0, t_f]$ IFF THE COLUMBS OF $C(t)\Phi(t, t_0)$ ARE LINEARLY INDEPENDENT ON $[t_0, t_f]$

(NOTE: FOR CONT, WE WANT $H_x(,)$)

PROOF:

(i) SUFF

$\dot{x} = A\underline{x} \Leftarrow$ UNFORCED SYSTEM

$x(t) = \Phi(t,t_0)x_0$

$y(t) = Cx = C\Phi(t,t_0)x_0$

$[C\Phi(t,t_0)]^T y(t) = [C\Phi(t,t_0)]^T C\Phi(t,t_0)x_0$

$$\int_{t_0}^{t_f}[C\Phi(t,t_0)]^T y(t)\,dt = \underbrace{\int_{t_0}^{t_f}[C\Phi(t,t_0)]^T C\Phi(t,t_0)\,dt}_{\text{SIMILAR TO } M_0}\,x_0$$

RECALL

$$M_0(t_0,t_f) = \int_{t_0}^{t_f} H_x(t_f,\tau)H_x^T(t_f,\tau)\,d\tau$$

THUS, SYSTEM IS OBSERVABLE IF COLUMBS OF $[C\Phi(t,t_0)]^T$ ARE LINEARLY IND.

(ii) NECESSITY: ESTABLISHED AS BEFORE, BY COUNTERDICTION

THE CASE FOR OB. (DERIVED SAME AS FOR CONT) IS, FOR TIME INV. IS

$$\Gamma = [C^T \mid AC^T \mid A^2C^T \mid \ldots \mid A^{n-1}C^T]$$

$$\Gamma_1 = C^T$$

$$\Gamma_{K+1} = \dot{\Gamma}_K(\tau) + A\Gamma_K(\tau)$$

(ALL A'S SHOULD BE TRANSPOSED. SOMEWONE GOOFED)

(ie $\Gamma_{K+1} = \Gamma_K(\tau) + A^T\Gamma_K(\tau)$)

9-10-76 (FRI)

EXAMPLE:

$$\begin{bmatrix}\dot{X}_1\\ \dot{X}_2\end{bmatrix} = \overset{(A)}{\begin{bmatrix}0 & 1\\ -1 & -2\end{bmatrix}}\begin{bmatrix}X_1\\ X_2\end{bmatrix} + \overset{(B)}{\begin{bmatrix}1\\ -1\end{bmatrix}}u(t)$$

$$\begin{bmatrix}Y_1\\ Y_2\end{bmatrix} = \overset{(C)}{\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}}\begin{bmatrix}X_1\\ X_2\end{bmatrix}$$

LOOK @ CONTROLLABILITY (STATE)

$$\Gamma_c = \left[B \,\vdots\, -AB \,\vdots\, A^2B \,\vdots\, \ldots\, (A)^{n-1}B\right]$$

$$= \begin{bmatrix}\underbrace{\begin{matrix}-1\\ 1\end{matrix}}_{B} & \vdots & \underbrace{\begin{matrix}+1\\ -1\end{matrix}}_{-AB}\end{bmatrix} \Leftarrow \text{SINGULAR!}$$

SYSTEM IS <u>NOT</u> CONTROLLABLE

LOOK @ OBSERVABILITY (STATE)

$$\Gamma_0 = \left[C^T \; A^TC^T \; A^{T2}C^T \ldots (A^{n-1})C^T\right]$$

$$= \begin{bmatrix}1 & 1 & \vdots & 0 & -1\\ 0 & 1 & \vdots & 1 & -1\end{bmatrix} \Leftarrow \text{RANK OF TWO.}$$

(i.e. 2 LINEARLY IND. COLUMNS)

SYSTEM <u>IS</u> OBSEVABLE

CONTROLABILITY (OUTPUT)

$$\Gamma_Y = \left[CB \,\vdots\, -CAB \,\vdots\, CA^2B \,\vdots\, \ldots\right]$$

$$= \begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}1\\ -1\end{bmatrix} - \begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}1\\ -1\end{bmatrix} = \begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}$$

∴ <u>NOT</u> OUTPUT CONTROLLABLE

LET'S LOOK @ ORIGINAL EQUATION:

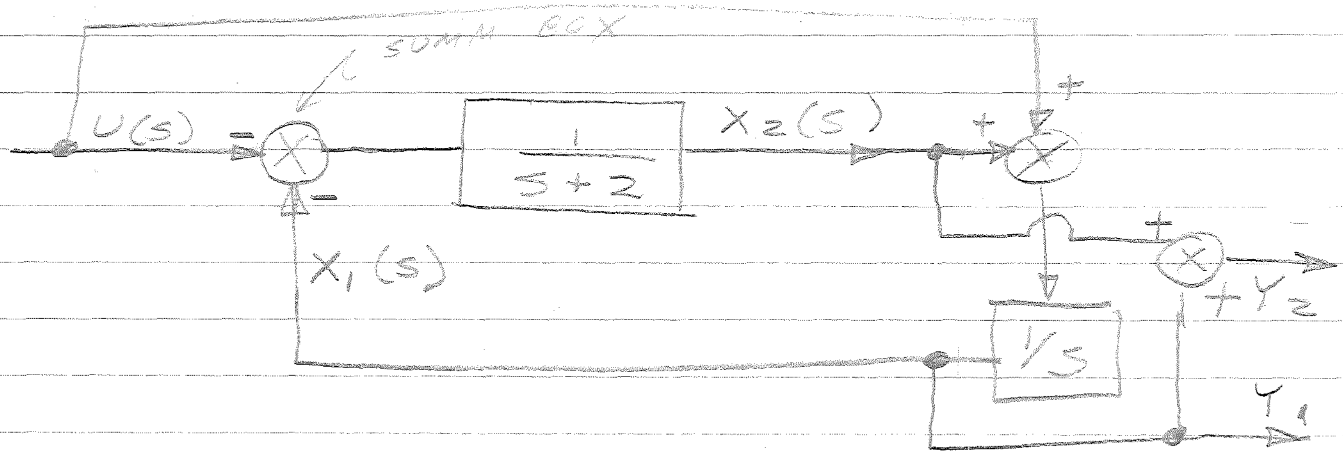
$$\begin{cases} \dot{x}_1 = x_2 + u(t) \\ \dot{x}_2 = -x_1 - 2x_2 - u(t) \end{cases}$$

OR

$$\begin{cases} sX_1(s) = X_2(s) + U(s) \\ sX_1(s) = -X_1(s) - 2X_2(s) - U(s) \end{cases}$$

$$X_1(s) = \frac{1}{s}\left[X_2(s) + U(s)\right]$$
$$X_2(s) = \frac{1}{s+2}\left[-X_1(s) - U(s)\right]$$

LOOK @ BLOCK DIAGRAM



WHERE $Y_1 = X_1$
$Y_2 = X_1 + X_2$

CONSIDER

$$\frac{X_1(s)}{U(s)} = \frac{1}{s}\left[\frac{X_2}{U(s)} + 1\right]$$

$$\frac{X_2(s)}{U(s)} = \frac{1}{s+2}\left[-\frac{X_1(s)}{U(s)} - 1\right]$$

SUBSTITUTING

$$\frac{X_1(s)}{U(s)} = +\frac{1}{s}\left[\frac{-1}{s+2}\,\frac{X_1(s)}{U(s)} - \frac{1}{s+2} + 1\right]$$

OR

$$\frac{X_1(s)}{U(s)}\left[1 + \frac{1}{s(s+2)}\right] = \frac{1}{s}\left(\frac{-1}{s+2} + 1\right)$$

$$\frac{X_1(s)}{U(s)} \quad \frac{s(s+2)+1}{s(s+2)} = \frac{s+1}{s(s+2)}$$

$$\frac{X_1(s)}{U(s)} = \frac{s+1}{(s+1)^2} = \frac{1}{s+1} \leftarrow \text{POLE ZERO CANCEL,}$$

$$(s+1)\,X_1(s) = U(s)$$

$$\dot{X}_1 + X_1 = U(t)$$

X_2 DOES NOT APPEAR.

SUBSTITUTING THIS RESULT BACK:

$$\frac{X_2(s)}{U(s)} = \frac{1}{s+2}\left(\frac{-1}{s+1} - 1\right) = \frac{-1}{s+1}$$

OR $\dot{X}_2 + X_2 = -U(t)$

TWO DECOUPLED SYSTEMS:

$$\frac{X_2(s)}{U(s)} = -\frac{X_1}{U(s)} \Rightarrow X_1 + X_2 = 0$$

ON $X_1 = X_2 = 0$, LINE ONLY IS THE SYSTEM CONTROLLABLE

LET'S LOOK @ THE OUTPUT

$$\frac{Y_1(s)}{U(s)} = \frac{X_1(s)}{U(s)} = \frac{1}{s+1}$$

CONTROLLABLE

$$\frac{Y_2(s)}{U(s)} = \frac{X_1(s)}{U(s)} + \frac{X_2(s)}{U(s)} = 0$$

<u>NOT</u> CONTROLLABLE

<u>END OF CHAPT 1</u>

<u>PERFORMANCE MEASURE</u>

CONTROL PROBLEM

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

FIND AN OPTIMAL ADMISSIBLE CONTROL u^*
TO FIND ADMISSIBLE TRAJECTORY
x^* ∋ THE PERFORMANCE
MEASURE J, IS MINIMIZED:

$$J = h[x(t_f), t] + \int_{t_0}^{t_f} g[x(t), u(t), t]\,dt$$

= THE PERFORMANCE MEASURE

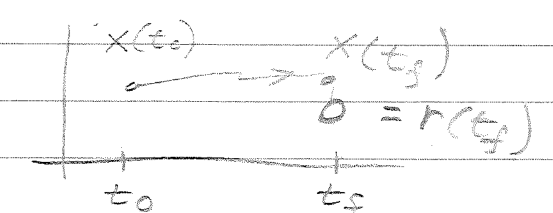
SOME SPECIAL CASES

(a) THE MINIMUM TIME PROBLEM

GO FROM $X(t_0)$ TO $x(t_f)$ IN MIN. TIME

$$J = t_f - t_0 = \int_{t_0}^{t_f} dt$$

(b) TERMINAL CONTROL PROBLEM



CONCERNED
BY FINAL STATE

$$J = \sum_{i=1}^{n} \left[x_i(t_f) - r_i(t_f) \right]^2$$

$$= \left[x(t_f) - r(t_f) \right]^T \left[x(t_f) - r(t_f) \right]$$

$$\triangleq \| x(t_f) - r(t_f) \|$$

OR, WEIGHING:

$$J = \left[x(t_f) - r(t_f) \right]^T H \left[x(t_f) - r(t_f) \right]$$

$$\triangleq \| x(t_f) - r(t_f) \|_H^2$$

H: SYMM. POS. SEMIDEFINATE MATRIX (NORMALLY DIAGONAL)

H WEIGHTS COMPONENTS

(c) MINIMUM CONTROL EFFORT PROBLEM

(i) MINIMUM FUEL PROBLEM

CONSIDER ROCKET

$u(t) =$ THRUST $\propto \frac{d\,\text{FUEL}}{dt}$

$$\Rightarrow |u(t)| = k \frac{dF}{dt}$$

$$\Rightarrow F = \frac{1}{k} \int_{t_0}^{tf} |u(t)| dt$$

THAT IS

$$J = \int_{t_0}^{t_f} |u(t)| dt$$

(ii) MINIMUM CONTROL ENERGY

$$J = \int_{t_0}^{t_f} \|u\|_R^2 dt$$

$$\|u\|_R^2 = u^T R u$$

R ≡ SYMMETRIC POSITIVE DEF

9/13/76 (MON)

(d) TRACKING PROBLEM

KEEP $r(t)$ = DESIRED STATE, CLOSE AS POSSIBLE

$$J = \int_{t_0}^{t_f} \|x(t) - r(t)\|_{Q(t)}^2 dt$$

$$= \int_{t_0}^{t_f} \underbrace{e^T}_{\text{ERROR}} Q(t) e \, dt$$

$\Rightarrow \int_{t_0}^{t_f} \Sigma_{i=1}^{n} q_i(t) e_i(t)$ IS NO GOOD (e_i CAN BE + OR NEG)

SPECIAL CASE: REGULATOR PROBLEM $\Rightarrow r(t) = 0$

WE ALSO HAVE CONSTRAINTS.

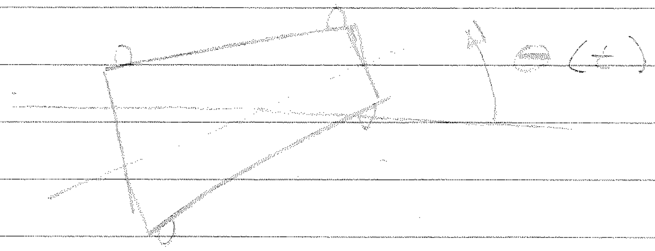
SAY $\|u(t)\| \leq 1$. MAY ALSO PUT IN ENERGY CONSTRAINT:

$$J = \int_{t_0}^{t_f} \|u\|_{R(t)}^2 dt + \int_{t_0}^{t_f} \|x(t) - r(t)\|_Q^2 dt$$

$$+ \underbrace{\|x(t_f) - r(t_f)\|_{H_f}^2}_{\text{WEIGHTED PENALTY TERM}}$$

THIS IS MOST GENERAL TRACKING PROB. STATEMENT.

EXAMPLE:



TURNS OUT

τ - TORQUE = $I \frac{d^2\theta}{dt^2}$

I - MOMENT OF INERTIA

LET $x = \theta$, $x_2 = d\theta/dt$

THEN $\dot{x}_1 = x_2$

$\dot{x}_2 = \tau / I = u$

PICK A PERFORMANCE MEASURE

USE REGULATOR J -

$$J = \int_0^\infty \|x(t)\|_Q^2 \, dt + \int_0^\infty R u^2(t) \, dt$$

$$= \int_0^\infty \underline{x_1 x_2} \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} dt + \int_0^\infty R u^2 dt$$

LET $q_2 = 0$;

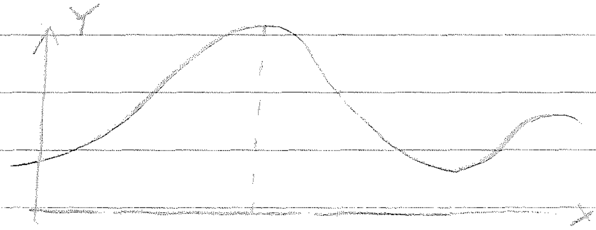
$$J = \int_0^\infty \underline{x_1 x_2} \begin{bmatrix} q_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} dt + \int_0^\infty R u^2 dt$$

SEE AROUND PGS 37 ON

FOR SOME SOLUTIONS

CALCULUS OF EXTREMA

$Y = f(x)$



$f(\alpha + \Delta x) - f(\alpha) < 0 \Rightarrow$ REL MAX

$f(\alpha + \Delta x) - f(\alpha) > 0 \Rightarrow$ REL MIN

(i) $\frac{dy}{dx} = 0$

(ii) $\left.\frac{d^2y}{dx^2}\right|_{x^*} \overset{>}{\underset{<}{=}} 0$

$> 0 \Rightarrow$ MIN

$= 0 \Rightarrow$ STATIONARY PTS

$< 0 \Rightarrow$ MAX

WHAT ABOUT A FUNCTION OF SEVERAL VERIABLES, SET n FIRST DERIVATIVES TO ZERO.

Ex: (i) $f(x) = \frac{1}{(x_1 - 1)^2 + (x_2 - 1)^2 + 1}$

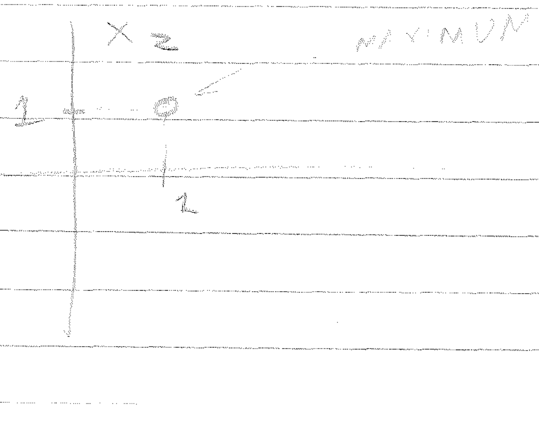
$\frac{\delta Y}{\delta x_1} = 0 \qquad \frac{\delta Y}{\delta x_2} = 0$

$\frac{\delta Y}{\delta x_1} = \frac{-\frac{\delta}{\delta x_1}(x_1 - 1)^2}{[(x_1 - 1)^2 + (x_2 - 1)^2 + 1]^2}$

$= \frac{2(x_1 - 1)}{[(x_1 - 1)^2 + (x_2 - 1)^2 + 1]^2} = 0 \Rightarrow x_1 = \underline{1}$

SIMILARLY

$\frac{\delta Y}{\delta x_2} = 0 \Rightarrow x_2 = \underline{1}$



CONSIDER

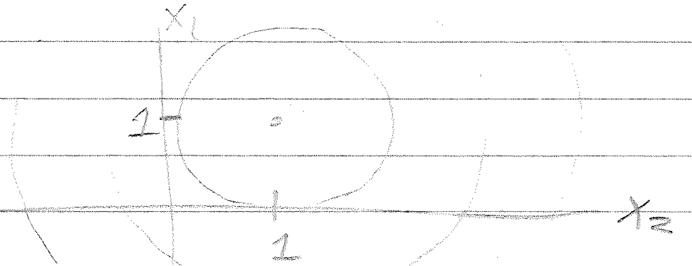
$$Y(X) = C = \text{CONST}$$

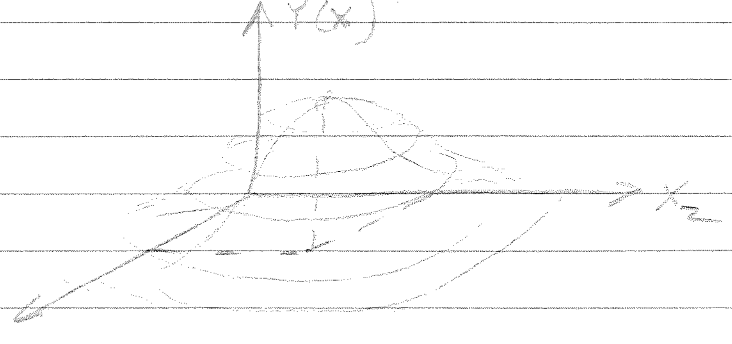
$$= \frac{1}{(X_1-1)^2 + (X_2-1)^2 + 1}$$

OR

$$(X_1-1)^2 + (X_2-1)^2 = \frac{1}{C} - 1$$

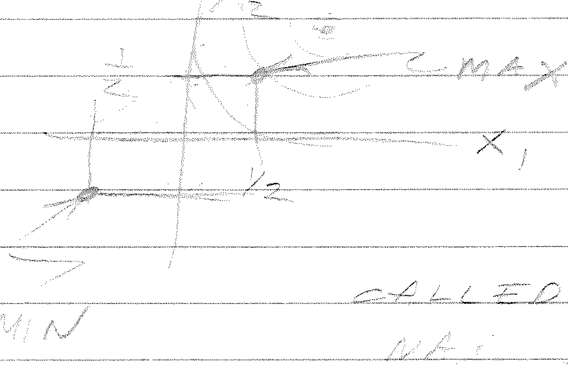
A FAMILY OF CIRCLES





(ii) EXTREMIZE $Y(X) \ni$

$$|X_1| \leq 1/2 \qquad |X_2| \leq \frac{1}{2}$$



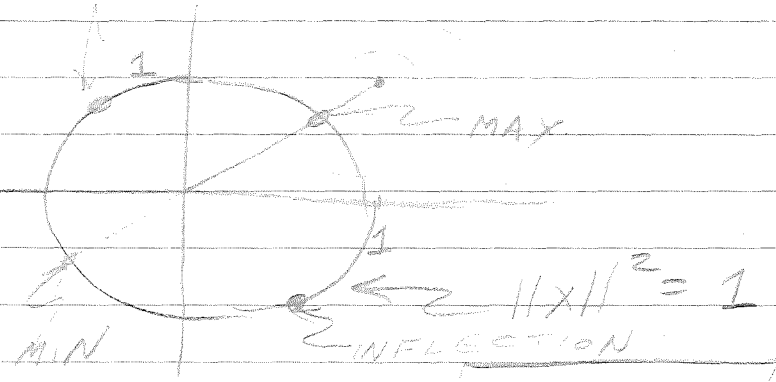
CALLED EXTERIOR

MAX OR MIN

$$MAX = \begin{bmatrix} -1/2 \\ 1/2 \end{bmatrix} \qquad MIN = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$$

(iii) EXTREMIZE $Y(x) \Rightarrow$

wrt $\|x\|^2 = 1$ $= x_1^2 + x_2^2$



NEW $x_2 = \pm\sqrt{1 - x_1^2}$

$$1$$

$$\Rightarrow Y(x) = (x_1 - 1)^2 + (\pm\sqrt{1 - x_1^2} - 1)^2 + 1$$

$$dY/dx = 0 \Rightarrow x_1^2 = 1 - x_1^2$$

OR $x_1 = \pm \frac{1}{\sqrt{2}}$

$$\Rightarrow x_2 = \pm 1/\sqrt{2}$$

9-12-76 (WED)

CONSTRAINED EXTREMA PROBLEMS

— LAGRANGE MULTIPLIER'S THEORY

MAXIMIZE

$$Y(x) = \frac{1}{(x_1-1)^2+(x_2-1)^2+1} \quad \text{SUBJECT TO } x_1^2+x_2^2=1$$

IN GENERAL

EXTREMEMIZE $Y(x)$ SUBJECT TO

$$\left.\begin{aligned} f_1(x) &= 0 \\ f_2(x) &= 0 \\ &\vdots \\ f_m(x) &= 0 \end{aligned}\right\} m \text{ CONSTRAINTS} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

DIFFERENTIATE

$$dY(x) = 0$$

$$= \frac{\delta Y}{\delta x_1} dx_1 + \frac{\delta Y}{\delta x_2} dx_2 + \cdots + \frac{\delta Y}{\delta x_n} dx_n = 0$$

ALSO:

$$\left.\begin{aligned} df_1 &= 0 \\ df_2 &= 0 \\ &\vdots \\ df_m &= 0 \end{aligned}\right\} \Rightarrow \left\{\begin{aligned} &\frac{\delta f_i}{\delta x_1} dx_1 + \frac{\delta f_i}{\delta x_2} dx_2 + \cdots + \frac{\delta f_i}{\delta x_n} dx_n = 0 \\ &i = 1, 2, \ldots, m \end{aligned}\right.$$

INTRODUCE LAGRANGIAN MULTIPLIERS

$$\begin{aligned} \lambda_1 df_1 &= 0 \\ \lambda_2 df_2 &= 0 \\ &\vdots \\ \lambda_m df_m &= 0 \end{aligned} \quad \Rightarrow \lambda_i \left[\frac{\delta f_i}{\delta x_1} dx_1 \ldots \right] = 0$$

ADD THE EQUATIONS:

$$dY + \lambda_1 df_1 + \lambda_2 df_2 + \cdots + \lambda_m df_m = 0$$

$$\triangleq dH$$

$$dH = d\left(Y + \underbrace{\lambda_1 \cdots \lambda_m}_{} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix}\right)$$

$$= d(Y + \lambda^T f)$$

OR

$$H \stackrel{\Delta}{=} Y + \lambda^T f$$

$$dH = 0 \Rightarrow$$

$$\frac{\delta Y}{\delta x_1} dx_1 + \frac{\delta Y}{\delta x_2} dx_2 + \cdots + \frac{\delta Y}{\delta x_n} dx_n$$

$$+ \sum_{i=1}^{m} \lambda_i \left(\frac{\delta f_i}{\delta x_1} dx_1 + \frac{\delta f_i}{\delta x_2} dx_2 + \ldots + \frac{\delta f_i}{\delta x_n} dx_n \right)$$

$$= 0$$

COMBINING VARIABLES

$$\left(\frac{\delta Y}{\delta x_1} + \sum_{i=1}^{m} \lambda_i \frac{\delta f_i}{\delta x_1} \right) dx_1$$

$$+ \ldots + \left(\frac{\delta Y}{\delta x_n} + \sum_{i=1}^{m} \lambda_i \frac{\delta f_i}{\delta x_n} \right) dx_n$$

$$= 0$$

ALL TERMS ARE LINEARLY INDEPENDENT, THUS, THEY <u>ALL</u> MUST BE = 0:

$$\frac{\delta Y}{\delta x_K} + \sum_{i=1}^{m} \lambda_i \frac{\delta f_i}{\delta x_K} = 0 \quad ; K = 1, 2, \ldots n$$

OR

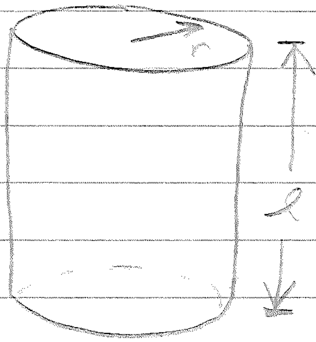
$$\frac{\delta H}{\delta x_K} = 0 \quad \ni \quad H = H(\underset{\overset{\uparrow}{n}}{x}, \underset{\overset{\uparrow}{m}}{\lambda})$$

$$\text{OR} \quad \frac{\delta H}{\delta x_K} = \frac{\delta Y}{\delta x_K} + \lambda^T \frac{\delta f}{\delta x_K} = 0 \Leftarrow n \text{ EQ}$$

$$\text{RECALL} \quad f(x) = 0 \Leftarrow m \text{ EQ}$$

$\therefore n + m$ EQUATIONS & AS MANY UNKNOWNS

EXAMPLE



WANNA MAXIMIZE VOLUME FOR A GIVEN SURFACE AREA.

$V(r,\ell) = \pi r^2 \ell$ ① ⟸ WANNA MAXIMIZE

SUBJECT TO

$A = 2\pi r^2 + 2\pi r \ell = A_0 = \text{CONSTANT}$

OR

$f(x) = 2\pi r^2 + 2\pi r \ell - A_0 = 0$ ②

① METHOD 1 OF DIRECT SUBSTITUTION

FROM ②:

$$\ell = \frac{A_0 - 2\pi r^2}{2\pi r} \quad ③$$

PLUG INTO ①:

$$V(r) = \pi r^2 \frac{A_0 - 2\pi r^2}{2\pi r}$$

$\frac{dV(r)}{dr} = 0$ GIVES EXTREMA

$$\frac{d}{dr} V_r = \frac{A_0}{2} - 3\pi r^2 = 0$$

$$r = \pm\sqrt{A_0/6\pi} \quad \text{OR} \quad r_{opt} = \sqrt{\frac{A}{6\pi}}$$

PLUG r INTO 2 GIVES

$$\ell_{opt} = \sqrt{\frac{2A_0}{3\pi}}$$

$$\frac{\ell_{opt}}{r_{opt}} = 2 \Longrightarrow \ell_{opt} = 2r_{opt}$$

USING LAGRANGE MULTIPLIERS

$H(x, \lambda) = H(r, l, \lambda)$

$\mathcal{H} = \gamma + \lambda^T f$

$= V(r, l) + \lambda(2\pi r^2 + 2\pi r l - A_0)$

$= \pi r^2 l + \lambda(2\pi r^2 + 2\pi r l - A_0)$

① $\frac{\delta \mathcal{H}}{\delta r} = 0 = 2\pi r l + \lambda(4\pi r + 2\pi l)$

② $\frac{\delta \mathcal{H}}{\delta l} = 0 = \pi r^2 + \lambda(2\pi r)$

TWO EQ & 3 UNKNOWNS (l, r, λ)

BUT $f(x) = 0$

③ $\Rightarrow 2\pi r^2 + 2\pi r l - A_0 = 0$

FROM ② ; $\lambda = -r/2$

GIVES US (WHEN ALL IS SAID & DONE)

$r(l - 2r) = 0 \Rightarrow r = 0, \ l/2$

THUS

$r_{opt} = \frac{l_{opt}}{2}$

FROM ③ $2\pi r^2 - 2\pi r l - A_0 = 0$

PLUG IN $\lambda = -r/2$ & $r = l/2$

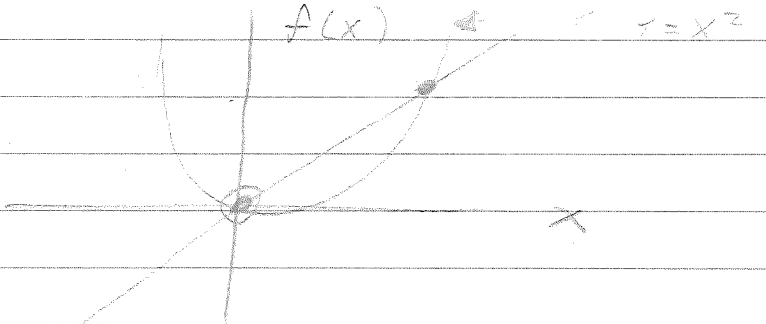
TO GET λ

$\Rightarrow \lambda = \pm\sqrt{\frac{A_0}{24\pi}}$

$r = \mp 2\sqrt{\frac{A_0}{24\pi}} = \sqrt{\frac{A_0}{6\pi}} = r_{opt}$

EXAMPLE:

$H = Y + \lambda^T f_{(x)}$ Min $Y(x) = x$ $\ni$ $x^2 = 0 = f(x)$



$\mathcal{H} = x + \lambda x^2$

$\frac{\delta H}{\delta x} = 1 + 2\lambda x = 0$

THIS DON WORK!

(CALLED A "SINGULAR" PROBLEM)

NEED MORE MULTIPLIERS

$\mathcal{H} = \lambda_0 Y + \lambda^T f$

9-16-76 (FRI)

EXTREMIZE $Y(X)$ SUBJECT TO

$f(x) = 0$ M EQUASIONS

$H = Y(x) + \lambda^T f(x)$

$\delta H / \delta x = 0$ (r EQUA)

$\delta Y = 0$ & δH ARE EQUIVALENT

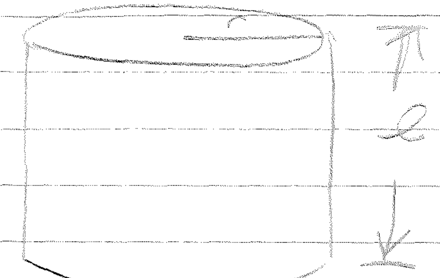
FOR EXTREMA

NEC. CONDITION $\delta Y = 0$

SUFF. CONDITION $\delta^2 Y \gtreqless 0$

(OR NOT USE $\delta^2 H$)

BACK TO TIN CAN PROBLEM



MAXIMIZE

$V = \pi r^2 l$

SUBJECT TO

$2\pi r l + 2\pi r^2 = A$

NEC. CONDITION IS

$$dV = 0$$
$$= \frac{\delta V}{\delta r} dr + \frac{\delta V}{\delta l} dl$$

ALSO, WE WANT

$$df = d(2\pi r l + 2\pi r^2 - A) = 0$$
$$= \frac{\delta f}{\delta r} dr + \frac{\delta f}{\delta l} dl$$

THEN

$$dV = 2\pi r l \, dr + \pi r^2 dl$$

$$df = (2\pi l + \pi 4 r) dr + (2\pi r) dl = 0$$

IN MATRIX FORM

$$\begin{bmatrix} 2\pi r & \pi r^2 \\ 2\pi \ell + 4\pi r & 2\pi r \end{bmatrix} \begin{bmatrix} dr \\ d\ell \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$\Rightarrow$ NEC CONDITION IS

$$\det \begin{vmatrix} 2\pi r & \pi r^2 \\ 2\pi \ell + 4\pi r & 2\pi r \end{vmatrix} = 0$$

$$= 4\pi^2 r^2 \ell - 2\pi^2 r^2 \ell - 4\pi^2 r^3 = 0$$

OR

$$2\cancel{\pi}^2 \cancel{r}^2 \ell = 4\cancel{\pi}^2 r^3$$

$$\Rightarrow 2\ell = r$$

SAME RELATION WE GOT LAST TIME. WE MAY USE THIS PROCEDURE (AND BYPASS LAGRANGE MULTIPLIERS) WHEN $m = n - 1$

NONLINEAR EQUATION

$$f(x) = 0$$

$$f: R^n \rightarrow R^n \longrightarrow \begin{cases} f_1(x_1 \ldots x_n) = 0 \\ \vdots \\ f_n(x_1 \ldots x_n) = 0 \end{cases}$$

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

CONSIDER

$$dy/dt = -y \Rightarrow y = y(0)e^{-t}$$

FOR $y(0) = 0 \quad y = 0$

CONSIDER

$$\left.\begin{aligned} df/dt = \dot{f}_1 &= -f_1 \\ \dot{f}_2 &= -f_2 \\ &\vdots \\ \dot{f}_{n-1} &= -f_{n-1} \\ \dot{f}_n &= \pm f_n \end{aligned}\right\} \qquad \begin{aligned} f_1 &= f_1(0)e^{-t} \\ f_2 &= f_2(0)e^{-t} \\ &\vdots \\ f_{n-1} &= f_{n-1}(0)e^{-t} \\ f_n &= f_n(0)e^{\pm t} \end{aligned}$$

INITIAL CONDITIONS

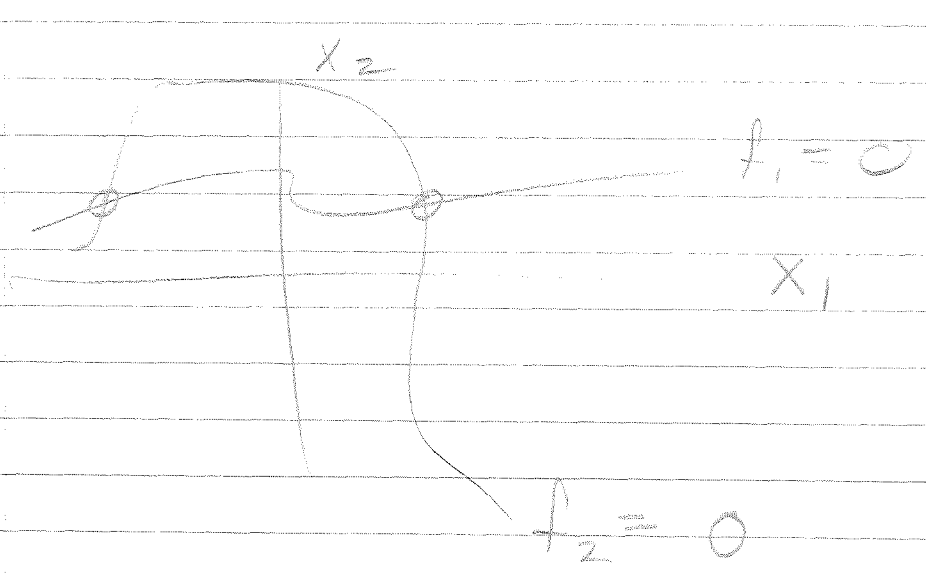
$$f[x(0)] = 0$$

$$f_2[x(0)] = 0$$

$$\vdots$$

$$f_{n-1}[x(0)] = 0$$

$$f_n[x(0)]$$

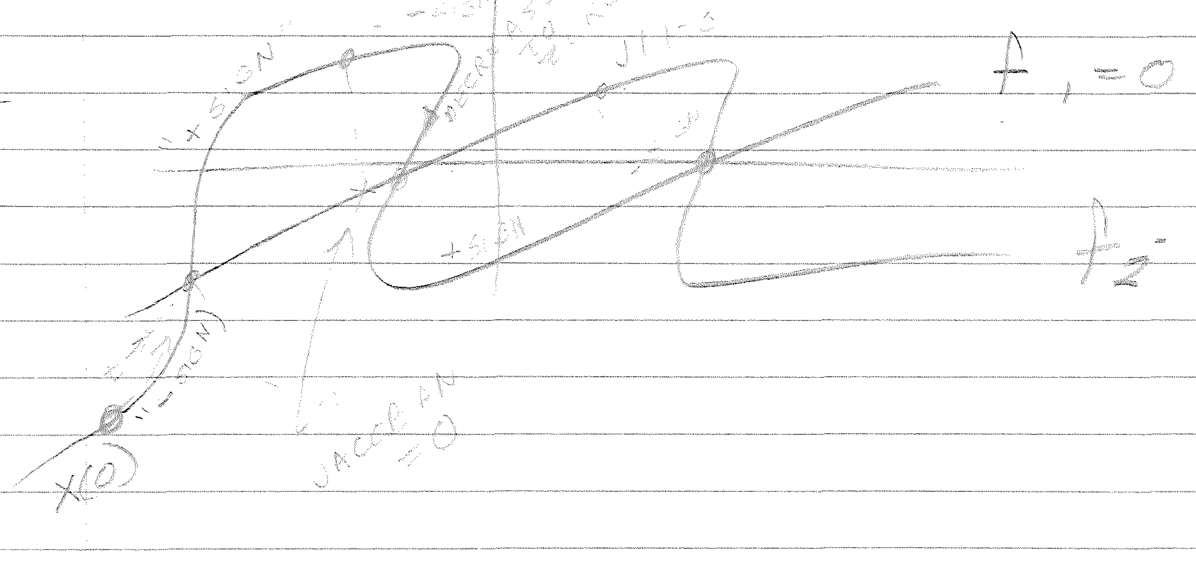


IN GENERAL SOLUTION OF n-1 EQUATIONS & n UNKNOWNS GIVES US A TWO DIMENSIONAL LINE IN AN n-DIMENSIONAL SPACE. NOTE FOR GIVEN INITIAL CONDITIONS

$$f_K = 0 \quad \forall \; K = 1, 2, \ldots, n-1$$

BUT,
NOW WE GOTTA FIND $f_n(0)$
TO FIND $f_n = f_n(0) e^{\pm t}$

$$f_1(t) = 0 e^{-t} = 0$$
$$f_2(t) = f_2(0) e^{\pm t}$$



THIS GIVES US

$$\frac{\delta f_1}{\delta x_1}\frac{dx_1}{dt} + \frac{\delta f_1}{\delta t}\frac{\delta x_2}{\delta t} + \ldots + \frac{\delta f_1}{\delta x_n}\frac{dx_n}{dt} = -f_1$$
$$\vdots$$
$$\frac{\delta f_n}{\delta x_1}\frac{dx_1}{dt} + \frac{\delta f_n}{\delta t}\frac{\delta x_2}{\delta t} + \ldots + \frac{\delta f_n}{\delta x_n}\frac{\delta x_n}{\delta t} = \pm f_n$$

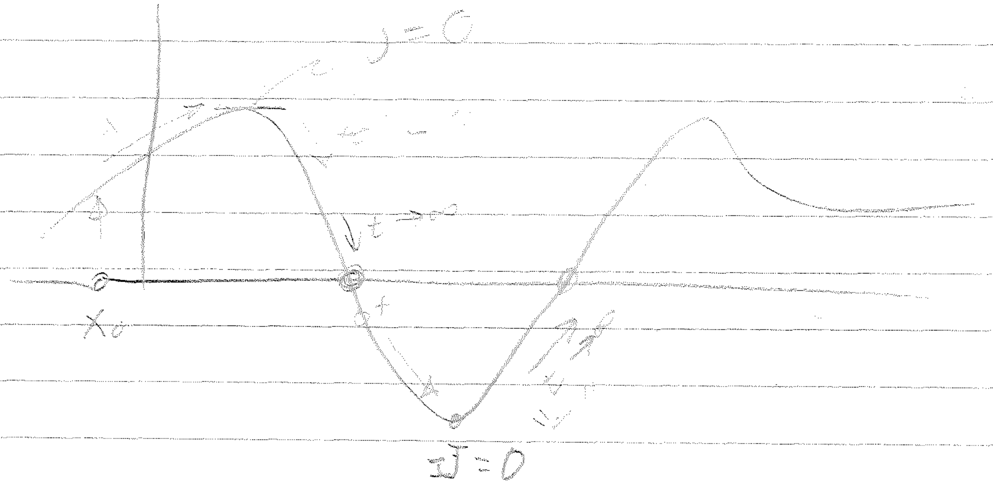
GIVES US MATRIX

$$\underbrace{\begin{bmatrix} \delta f_1/\delta x_1 & \frac{\delta f_1}{\delta x_2} & \cdots & \frac{\delta f_1}{\delta x_n} \\ \delta f_2/\delta x_1 & & & \\ \vdots & & & \\ \delta f_n/\delta x_1 & \frac{\delta f_n}{\delta x_2} & \cdots & \frac{\delta f_n}{\delta x_n} \end{bmatrix}}_{\text{JACOBIAN}} \begin{bmatrix} dx_1/dt \\ dx_2/dt \\ \vdots \\ \frac{dx_n}{dt} \end{bmatrix} = \begin{bmatrix} -f_1 \\ \vdots \\ -f_{n-1} \\ \pm f_n \end{bmatrix}$$

OR $J\dot{X} = F$

$$\Rightarrow \dot{X} = J^{-1} \begin{bmatrix} -f_1 \\ \vdots \\ \pm f_n \end{bmatrix} \quad x(0) \in \ell$$

EX



9/21/76 (MON)

HOMEWORK CHAPT 2-1, -3, -4, -5, -6 DUE MON

VECTOR FORMULATION
(OF EXTREMA PROBLEM)

$J = J(x, u)$ SUBJECT $f(x, u) = 0$

x : n VECTOR u : m VECTOR

ASSUME f : n-VECTOR

ON TO LAGRANGE MULTIPLIER THEORY

$$H = J + \lambda^T f \overset{= f^T \lambda}{} = H(x, u, \lambda) \quad \lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}$$

$$\frac{\delta H}{\delta u} = \frac{\delta J}{\delta u} + \frac{\delta}{\delta u}\left(f^T \lambda\right)$$

$$= \delta J / \delta u + \left(\delta f^T / \delta u\right) \lambda$$

$$\frac{\delta f^T}{\delta u} = \frac{\delta (f_1 \cdots f_n)}{\delta u} = \begin{bmatrix} \frac{\delta f_1}{\delta u_1} & \frac{\delta f_2}{\delta u_1} & \cdots & \frac{\delta f_n}{\delta u_1} \\ \frac{\delta f_1}{\delta u_2} & \vdots & & \\ \vdots & & & \\ \frac{\delta f_1}{\delta u_m} & \frac{\delta f_2}{\delta u_m} & \cdots & \frac{\delta f_n}{\delta u_m} \end{bmatrix}$$

TRANSPOSE OF JACOBIAN

ALSO, WE GOTTA HAVE

$$\frac{\delta H}{\delta x} = \frac{\delta J}{\delta x} + \frac{\delta}{\delta x}\left(f^T \lambda\right) = 0$$

$$\frac{\delta H}{\delta x} = \frac{\delta J}{\delta x} + \frac{\delta f^T}{\delta x} \lambda = 0$$

SUFFICIENCY CONDITION FOR SPECIFYING EXTREMA $\longrightarrow$

$$\delta J \triangleq dJ = \frac{\delta J}{\delta x_1} dx_1 + \frac{\delta J}{\delta x_2} dx_2 + \ldots$$

$$+ \frac{\delta J_n}{\delta x_n} dx_n + \frac{\delta J}{\delta u_1} du_1$$

$$+ \frac{\delta J}{\delta u_2} du_2 + \ldots + \frac{\delta J}{\delta u_n} du_n$$

$$= \left[\frac{\delta J}{\delta x_1} \; \frac{\delta J}{\delta x_2} \; --- \; \frac{\delta J}{\delta x_n} \right] \begin{bmatrix} dx_1 \\ dx_2 \\ \vdots \\ dx_n \end{bmatrix}$$

$$+ \left[\frac{\delta J}{\delta u_1} \; \frac{\delta J}{\delta u_2} \; --- \; \frac{\delta J}{\delta u_m} \right] \begin{bmatrix} \delta u_1 \\ \delta u_2 \\ \vdots \\ \delta u_m \end{bmatrix}$$

OR

$$\delta J = \left(\frac{\delta J}{\delta x} \right)^T \delta x + \left(\frac{\delta J}{\delta u} \right)^T \delta u$$

DIFFERENTIATE AGAIN
(FIND SECOND VARIATION)

$$\delta^2 J = d(dJ)$$

$$= \left[\frac{\delta}{\delta x} (\delta J) \right]^T \delta x$$

$$+ \left[\frac{\delta}{\delta u} (\delta J) \right]^T \delta u$$

$$= \delta x^T \left[\frac{\delta}{\delta x} \delta J \right]$$

$$+ \delta u^T \left[\frac{\delta}{\delta u} \delta J \right] \Longrightarrow$$

$$\doteq \delta x^T \left[\frac{\delta}{\delta x} \left\{ \left(\frac{\delta J}{\delta x} \right)^T \delta x + \left(\frac{\delta J}{\delta u} \right)^T \delta u \right\} \right]$$

$$+ \delta u^T \left[\frac{\delta}{\delta u} \left\{ \left(\frac{\delta J}{\delta x} \right)^T \delta x + \left(\frac{\delta J}{\delta u} \right)^T \delta u \right\} \right]$$

$$= \delta^2 J \Big|_{\substack{x = x^* \\ u = u^*}} = \delta x^T \left[\left\{ \frac{\delta}{\delta x} \left(\frac{\delta J}{\delta x} \right)^T \right\} \delta x + 0 + \left\{ \frac{\delta}{\delta x} \left(\frac{\delta J}{\delta u} \right)^T \right\} \delta u + 0 \right]$$

$$+ \delta u^T \left[\frac{\delta}{\delta u} \left(\frac{\delta J}{\delta x} \right)^T \delta x + \frac{\delta}{\delta x} \left(\frac{\delta J}{\delta u} \right)^T \delta u \right]$$

$$\Rightarrow \delta J (x^*, u^*) = 0$$

$$\delta^2 J \Big|_{\substack{x^* \\ u^*}} = \underline{(\delta x^T \; \delta u^T)} \begin{bmatrix} \frac{\delta}{\delta x} \left(\frac{\delta J}{\delta x} \right)^T & \frac{\delta}{\delta x} \left(\frac{\delta J}{\delta u} \right)^T \\ \frac{\delta}{\delta u} \left(\frac{\delta J}{\delta x} \right)^T & \frac{\delta}{\delta u} \left(\frac{\delta J}{\delta u} \right)^T \end{bmatrix} \begin{bmatrix} \delta x \\ \delta u \end{bmatrix}$$

$$= \left[\delta x^T \; \delta u^T \right] \begin{bmatrix} \frac{\delta}{\delta x} \left(\frac{\delta J}{\delta x} \right)^T & \frac{\delta}{\delta x} \left(\frac{\delta J}{\delta u} \right)^T \\ \left[\frac{\delta}{\delta x} \left(\frac{\delta J}{\delta u} \right)^T \right]^T & \frac{\delta}{\delta u} \left(\frac{\delta J}{\delta u} \right)^T \end{bmatrix} \begin{matrix} \delta x \\ \delta u \end{matrix}$$

POS DEF $\Longrightarrow$ MIN

NEG " $\Longrightarrow$ MAX

ZERO $\Longrightarrow$ SADDLE

EXAMPLE ; EXTREMIZE $J = J(x, u)$
SUBJECT TO $f(x; u) = 0$.
FOR LINEAR SYSTEMS
$f(x, u) = Ax + Bu + C = 0$
$J = \frac{1}{2}\|u\|_R^2 + \frac{1}{2}\|x\|_Q^2$

$= \frac{1}{2} u^T R u + \frac{1}{2} x^T Q x$

$x - n$ $B : n \times m$
$u - m$ $C : n$
$A - n \times n$ $R : m \times m$ POS. DEF., SYMMETRIC
$Q : n \times n$ POS " "

NOW
$$\mathcal{H} = \frac{1}{2} u^T R u + \frac{1}{2} x^T Q x$$
$$+ \lambda^T (Ax + Bu + C)$$

$$\frac{\delta \mathcal{H}}{\delta x} = 0 = \frac{1}{2} \frac{\delta x^T}{\delta x} Q x + \frac{1}{2} \frac{\delta x^T}{\delta x} Q^T x$$
$$+ \frac{\delta x^T}{\delta x} A^T \lambda \quad = Qx + A^T \lambda = 0$$

$$\frac{\delta \mathcal{H}}{\delta u} = 0 = Ru + B^T \lambda = 0$$

THUS, WE GOTTA SOLVE

$$\begin{cases} Qx + A^T \lambda = 0 & ① \\ Ru + B^T \lambda = 0 & ② \\ Ax + Bu + C = 0 & ③ \end{cases}$$

$\frac{\delta x^T}{\delta x} = I$ $\quad \frac{\delta x}{\delta x^T}$ NOT [illegible]

LET'S SOLVE

FROM ①:

$$QX = -A^T\lambda \Rightarrow X = -Q^{-1}A^T\lambda$$

FROM ②

$$U = -R^{-1}B^T\lambda$$

SUBSTITUTING THESE INTO ③:

$$-AQ^{-1}A^T\lambda - BR^{-1}B^T\lambda + C = 0$$

$$= (AQ^{-1}A^T + B^{+1}R^{-1}B^T\lambda) + C$$

OR

$$\lambda = (AQ^{-1}A^T + BR^{-1}B^T)^{-1}(C)$$

IS IT MAXIMAM OR MINIMUM

AGAIN

$$\delta^2 J = \underline{\delta X^T \delta U^T} \begin{bmatrix} \frac{\delta}{\delta X}\left(\frac{\delta J}{\delta X}\right)^T & \frac{\delta}{\delta X}\left(\frac{\delta J}{\delta U}\right)^T \\ \left[\frac{\delta}{\delta X}\left(\frac{\delta J}{\delta U}\right)\right]^T & \frac{\delta}{\delta U}\left(\frac{\delta J}{\delta U}\right)^T \end{bmatrix} \begin{matrix} \delta U \\ \delta \lambda \end{matrix}$$

$$\left(\frac{\delta J}{\delta X}\right)^T = (QX)^T = X^T Q^T$$

$$\frac{\delta}{\delta X}\left(\frac{\delta J}{\delta X}\right) = Q^T = Q$$

$$\left(\frac{\delta J}{\delta U}\right)^T = (RU)^T = U^T R^T$$

$$\frac{\delta}{\delta U}\left(\frac{\delta J}{\delta U}\right)^T = R^T = R$$

ALSO

$$\begin{matrix} \frac{\delta}{\delta X} & \frac{\delta}{\delta U} \end{matrix}$$

$$\Rightarrow \bar{J} = \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} = \text{POS. DEF.}$$

$\bar{J}$ IS ALSO POS- DEFINATE

$\Rightarrow$ A MINIMUM!

9/22/76 (WED)

EXAMPLE: $Z = \mathcal{H}X + \underset{\text{NOISE}}{\underset{\uparrow}{V}}$

X: n - VECTOR

Z: m - VECTOR

V = m - VECTOR

$\mathcal{H}$: $m \times n$ MATRIX

FIND BEST ESTIMATE OF X ∋

$$J = \frac{1}{2}\| Z - \mathcal{H}X \|_R^2 \quad \text{IS MINIMUM}$$

∋ R IS P.D. SYM. MATRIX

$$J = \frac{1}{2}(Z - HX)^T R (Z - HX)$$

DIFFERENTIATE:

$$\frac{\delta}{\delta X} J = \frac{1}{2} \frac{\delta}{\delta X}(Z - \mathcal{H}X)^T R (Z - \mathcal{H}X)$$

$$+ \frac{1}{2} \frac{\delta}{\delta X}\left[(Z - HX)^T\right] R^T (Z - HX)$$

$$= \frac{\delta}{\delta X}\left[(Z - HX)\right] R (Z - HX)$$

$$= -\mathcal{H}^T R (Z - HX) = 0$$

$$\Rightarrow +H^T R Z = H^T R H X$$

$$X^* = (H^T R H)^{-1} H^T R Z \leftarrow \overset{\text{LEAST}}{\text{MSE}}$$

MEAN SQUARE ERROR

THIS IS MINIMUM X

SPECIAL CASES

(i) m ESTIMATES OF A SCALAR.

(n = 1)

$$\Rightarrow \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix} = \begin{bmatrix} h_{11} \\ h_{21} \\ \vdots \\ h_{m1} \end{bmatrix} X + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}$$

LET h's = 1. Z's ARE THEN A BUNCH OF X ESTIMATES.

ASSUME R = I

$$X = \left[\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right]^{-1}$$

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_m \end{bmatrix}$$

$$= [m]^{-1} (z_1 + z_2 + \cdots + z_m)$$

$$= \bar{X} = \text{AVERAGE OF } Z\text{'s!}$$

$$\begin{bmatrix} z_1 \\ \cdot \\ \cdot \\ z_m \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \end{bmatrix} x + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}$$

$$z_{m+1} \quad 1 \quad v_{m+1} \leftarrow \text{NEW MEAS}$$

X_m = BEST ESTIMATE OF X FROM m MEASUREMENTS

$$= \frac{1}{m}\left(z_1 + z_2 \ldots + z_m\right)$$

$$X_{m+1} = \frac{1}{m+1}\left(z_1 + \ldots + z_m + z_{m+1}\right)$$

(ii) SEQUENTIAL ESTIMATION

FOR m MEASURES OUR BEST ESTIM

$$\hat{X}_m = (H^T R H)^{-1} (H^T R Z)$$

WHAT IS $\hat{X}_{m+1}$.

OR

$$\Rightarrow \begin{bmatrix} Z \\ \hline z_{m+1} \end{bmatrix} = \begin{bmatrix} H \\ \hline h \end{bmatrix} X + \begin{bmatrix} V \\ \hline v_{m+1} \end{bmatrix}$$

OR

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ \cdot \\ z_m \\ \hline z_{m+1} \end{bmatrix} = \begin{bmatrix} h_{11} & \cdots & h_{1n} \\ h_{21} & \cdots & h_{2n} \\ \vdots & & \\ h_{m1} & \cdots & h_{mn} \\ \hline h_{m+1,1} & \cdots & h_{m+1,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \\ \hline v_{m+1} \end{bmatrix}$$

RESOLVE:

$$J = \frac{1}{2} \left\| \begin{bmatrix} z \\ \hline z_{m+1} \end{bmatrix} - \begin{bmatrix} H \\ \hline h^T \end{bmatrix} x \right\|^2_{R=I}$$

WE SHOWED

$$X_m = (H^T R H)^{-1} (H R z)$$

$$R = I$$

$$\Rightarrow X_m = (H^T H)^{-1} H^T z$$

$$\frac{\delta J}{\delta X} = \frac{\delta}{\delta X} \left\{ \left(\begin{bmatrix} z \\ \hline z_{m+1} \end{bmatrix} - \begin{bmatrix} H \\ \hline h' \end{bmatrix} x \right)^T \left(\begin{bmatrix} z \\ \hline z_{m+1} \end{bmatrix} - \begin{bmatrix} H \\ \hline h' \end{bmatrix} x \right) \right\}$$

$$- \begin{bmatrix} H \\ \hline h' \end{bmatrix} \left(\begin{bmatrix} z \\ \hline z_{m+1} \end{bmatrix} - \begin{bmatrix} H \\ \hline h' \end{bmatrix} x \right) = 0$$

OR

$$\begin{bmatrix} H \\ \hline h^T \end{bmatrix}^T \begin{bmatrix} H \\ \hline h^T \end{bmatrix} X = \begin{bmatrix} H \\ \hline h^T \end{bmatrix} \begin{bmatrix} z \\ \hline z_{m+1} \end{bmatrix}$$

$$\Rightarrow \hat{X}_{m+1} = \left(\begin{bmatrix} H \\ \hline h^T \end{bmatrix}^T \begin{bmatrix} H \\ \hline h^T \end{bmatrix} \right)^{-1} \begin{bmatrix} H \\ \hline h' \end{bmatrix}^T \begin{bmatrix} z \\ \hline z_{m+1} \end{bmatrix}$$

WE WANNA WRITE THIS AS

$$X_{m+1} = X_m + \Delta X \Longrightarrow$$

NOW

$$\hat{X}_{m+1} = \left\{ \begin{bmatrix} H^T & h \end{bmatrix} \begin{bmatrix} H \\ h^T \end{bmatrix} \right\}^{-1} \begin{bmatrix} H \\ h^T \end{bmatrix}^T \begin{bmatrix} Z \\ Z_{m+1} \end{bmatrix}$$

$$= \left\{ H^T H + h h^T \right\}^{-1} \begin{bmatrix} H \\ h^T \end{bmatrix}^T \begin{bmatrix} Z \\ Z_{m+1} \end{bmatrix}$$

WE KNOW (FROM X_m)

$$P_m^{-1} = H^T H$$

DENOTE

$$P_{m+1}^{-1} = H^T H + h h^T$$

MATRIX INVERSION LEMMA (IN SAGE)

$$P_{m+1} = P_m - P_m h \underbrace{\left(h^T P_m h + 1 \right)^{-1}}_{\text{SCALAR}} \times h^T P_m$$

$$\hat{X}_{m+1} = \left\{ P_m - P_m h \left[h^T P_m h + 1 \right]^{-1} h^T P_m \right\} \left(H^T Z + h Z_{m+1} \right)$$

$$= P_m H^T Z + P_m h Z_{m+1} - P_m h \left[h^T P_m h + 1 \right] h^T P_m$$

$$= \hat{X}_m + \Delta X$$

$$= \hat{X}_m + P_m h \left[h^T P_m h + 1 \right]^{-1} \left(Z_{m+1} - h^T \hat{X}_m \right)$$

9-23-76 FRF

THE CALCULUS OF VARIATIONS

$$J = \int_{x_1}^{x_2} F\left[\bar{Y}(x), \bar{Y}'(x), x\right] dx$$

BOUNDRY CONDITIONS $Y(x_1) = Y_1$ $Y(x_2) = Y_2$

CONSIDER, FIRST, THE UNCONSTRAINED PROBLEM (DYNAMIC OPTIMIZATION W/O CONSTRAINTS)

APPROACH OF THE VARIED PATH.

ASSUME $Y(x)$ IS THE OPTIMAL SOLUTION - LET

$$\bar{Y}(x) = Y(x) + \varepsilon \eta(x)$$

NOTE; WE HAVE OPTIMAL SOLUTION WHEN $\varepsilon = 0$

$$\frac{d\bar{Y}}{dx} = \frac{dY(x)}{dx} + \varepsilon \frac{d\eta(x)}{\delta x}$$

$$\bar{Y}'_{(x)} = Y'(x) + \varepsilon \eta'(x)$$

THEN

$$J = \int_{x_1}^{x_2} F(Y + \varepsilon \eta, Y' + \varepsilon \eta', X)\, dx$$

$$\frac{dJ}{d\varepsilon} = \int_{x_1}^{x_2} \frac{dF}{d\varepsilon}\, dx$$

$$+ F\Big|_{x=x_2} \cancel{\frac{dx_2}{d\varepsilon}}^{\to 0} - F\Big|_{x=x_1} \cancel{\frac{dx}{d\varepsilon}}^{\to 0}$$

$$= \int_{x_1}^{x_2} \left[\frac{\delta F}{\delta \bar{Y}} \frac{d\bar{Y}}{d\varepsilon} + \frac{\delta F}{\delta \bar{Y}'} \frac{d\bar{Y}'}{d\varepsilon} + \frac{\delta F}{\delta X} \cdot \cancel{\frac{\delta x}{\delta \varepsilon}}^{0} \right] dx$$

$$\bar{Y} = Y + \epsilon \eta \qquad \bar{Y}' = Y' + \epsilon \eta'$$

$$d\bar{Y}/d\varepsilon = \eta \qquad d\bar{Y}'/d\varepsilon = \eta'$$

$$\Rightarrow \frac{dJ}{d\varepsilon} = \int_{x_1}^{x_2} \left[\frac{\delta F}{\delta \bar{Y}} \eta(x) + \frac{\delta F}{\delta \eta'(x)} \right] dx$$

$$\left.\frac{dJ}{d\epsilon}\right|_{\epsilon=0} = 0$$

$$= \int_{x_1}^{x_2} \left[\frac{\delta F}{\delta Y} \eta(x) + \frac{\delta F}{\delta Y'} \cdot \eta'(x) \right] dx = 0$$

$$\epsilon = 0 \implies Y = \overline{Y}$$

$$\left.\frac{dJ}{d\epsilon}\right|_{\epsilon=0} = \int_{x_1}^{x_2} \eta(x) \frac{dF(Y,Y',x)}{dY} \, dx$$

$$+ \int_{x_1}^{x_2} \eta'(x) \frac{\delta F(Y,Y',x)}{\delta Y'} \, dx$$

CONSIDER SECOND TERM
(INTEGRATION BY PARTS)

$$U = \frac{dF}{dY'} \qquad dV = \eta' dx$$

$$dU = \frac{d}{dx}\left(\frac{\delta F}{\delta Y'}\right) dx \qquad V = \eta$$

THEN

$$\left.\frac{dJ}{d\epsilon}\right|_{\epsilon=0} = \int_{x_1}^{x_2} \eta(x) \frac{dF}{\delta Y} dx$$

$$+ \eta(x) \frac{dF}{dY'} \Big|_{x_1}^{x_2}$$

$$- \int_{x_1}^{x_2} \eta(x) \frac{d}{dx}\left(\frac{\delta F}{\delta Y'}\right) dx = 0$$

$$\therefore \eta(x) \frac{\delta F}{\delta Y'} \Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} \eta(x) \left[\frac{\delta F}{\delta Y} - \frac{d}{dx}\left(\frac{\delta F}{\delta Y'}\right) \right] dx = 0$$

FROM THIS, WE GET TWO CONDITIONS

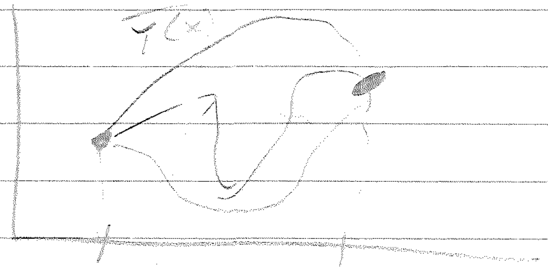
$$\eta(x) \frac{dF}{dY'}\Big|_{x_1}^{x_2} = 0 \qquad \left\{ \begin{array}{l} \eta(x) \frac{\delta F}{\delta Y'}\Big|_{x_2} = 0 \\ \eta() \frac{\delta F}{\delta Y'}\Big|_{x_1} = 0 \end{array} \right. \quad ①$$

$$\frac{\delta F}{\delta Y} - \frac{d}{dX}\left(\frac{\delta F}{\delta Y'}\right) = 0 \qquad ②$$

① IS
TRANSVERSALITY CONDITION
② IS
EULER LA GRANGE EQU,

THESE ARE NECESSARY CONDITIONS FOR THE EXISTANCE OF AN EXTREMUM ON J.

(i) FIXED END POINTS



$$\bar{Y} = Y + \varepsilon\eta$$

$$\bar{Y}(x_1) = Y(x_1) + \varepsilon\eta(x_1)$$

IF IT'S FIXED @ X_1, THEN

$$Y(x_1) = \bar{Y}(x_1)$$

OR

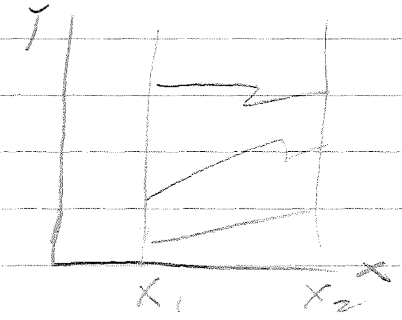
$$\eta(x_1) = 0$$

ALSO $\eta(x_2) = 0$

THUS

$$\eta(x) \frac{dF}{dY'}\Big|_{x_1}^{x_2} = 0$$

(ii) VARIABLE ENDPOINTS



TO SATISFY TRANSVERSATALITY

$$\eta(x_2) \frac{\delta F}{\delta Y'}\Big|_{x_2} - \eta(x_1) \frac{\delta F}{\delta Y'}\Big|_{x_1} = 0$$

SO, HERE OUR CONDITION IS
(NATURAL BOUNDRY CONDITION)

$$\frac{\delta F}{\delta Y'}\Big|_{x_2} = \frac{\delta F}{\delta Y'}\Big|_{x_1} = 0$$

(iii) $Y(x_1)$ FIXED
$Y(x_2)$ IS VARIABLE)

$Y(x_1)$ FIXED $\Longrightarrow \eta(x_1) = 0$
$(\eta(x_2) \neq 0)$

$Y(x_2)$ VARIABLE

$$\Longrightarrow \frac{\delta F}{\delta Y'}\Big|_{x_2} = 0$$

(iv) $Y(x_2)$ IS FIXED $\Longrightarrow \eta(Y_2) = 0$
$Y(X_1)$ IS VARIABLE

$$\Longrightarrow \frac{\delta F}{\delta Y'}\Big|_{X_1} = 0$$

F MUST BE TWICE DIFFERENTIABLE

EULER - LAGRANGE EQ. IS

$$\frac{\delta F}{\delta Y} - \frac{\delta}{\delta X}\left(\frac{\delta F}{\delta Y'}\right) = 0$$

EXPANDING

$$\frac{\delta F}{\delta Y} - \frac{\delta}{\delta Y}\left(\frac{\delta F}{\delta Y'}\right)\frac{\delta Y}{\delta X} + \frac{\delta}{\delta Y'}\left(\frac{\delta F}{\delta Y'}\right)\frac{\delta Y'}{\delta X} + \frac{\delta}{\delta X}\left(\frac{\delta F}{\delta Y'}\right)$$

$$= \frac{\delta F}{\delta Y} - \frac{\delta^2}{\delta Y\, d Y'} Y' + \frac{\delta^2 F}{\delta Y'^2} Y'' + \frac{\delta^2 F}{\delta X\, \delta Y'} = 0 \quad = 0$$

SUFFICIENT CONDITIONS

$$\left.\frac{d^2 J}{d\epsilon^2}\right|_{\epsilon=0} = \left.\frac{d}{d\epsilon}\left(\frac{dJ}{d\epsilon}\right)\right|_{\epsilon=0}$$

$$= \left.\frac{d}{d\epsilon}\int_{x_1}^{x_2}\left[\eta(x)\frac{dF}{dY} + \eta'\frac{dF}{dY'}\right] dx\right|_{\epsilon=0}$$

$$= \int_{x_1}^{x_2}\frac{d}{d\epsilon}\left[\eta(x)\frac{dF}{\delta Y} + \eta'\frac{\delta F}{\delta Y'}\right] dx$$

9/27/76 (MON)

CALCULUS OF VARIATIONS

$$J = \int_{x_1}^{x_2} F[\bar{Y}, \bar{Y}', x]\, dx$$

$$\bar{Y} = Y + \varepsilon \eta(x)$$

$$\bar{Y}' = Y' + \varepsilon \eta'(x)$$

$$\left.\frac{dJ}{d\varepsilon}\right|_{\varepsilon=0} = \int_{x_1}^{x_2} \left[\eta(x) \frac{dF}{d\bar{Y}} + \eta'(x) \frac{dF}{d\bar{Y}'}\right] dx = 0$$

$$\eta(x) \left.\frac{\delta F}{\delta Y'}\right|_{x_1}^{x_2} = 0 \leftarrow \text{TRANSVERSALITY}$$

$$\frac{\delta F}{\delta Y} - \frac{d}{dx}\left(\frac{\delta F}{\delta Y'}\right) = 0 \leftarrow \text{EULER LAGRANGE EQ.}$$

SECOND VARIATION

$$\left.\frac{d^2 J}{d\varepsilon^2}\right|_{\varepsilon=0} = \frac{d}{d\varepsilon}\left(\frac{dJ}{d\varepsilon}\right)$$

$$= \frac{d}{d\varepsilon} \int_{x_1}^{x_2} \left[\eta(x) \frac{\delta F}{\delta \bar{Y}} + \eta'(x) \frac{\delta F}{\delta \bar{Y}'}\right] dx$$

$$= \int_{x_1}^{x_2} \frac{d}{d\varepsilon}\left[\eta \frac{\delta F}{d\bar{Y}} + \eta' \frac{\delta F}{\delta \bar{Y}'}\right] dx$$

INT. BY PART

$$= \int_{x_1}^{x_2} \left[\frac{\delta}{\delta \bar{Y}}\left[\eta \frac{\delta F}{\delta \bar{Y}} + \eta' \frac{\delta F}{\delta \bar{Y}'}\right] \frac{d\bar{Y}}{d\varepsilon} + \frac{\delta}{\delta \bar{Y}'}\left[\eta \frac{\delta F}{\delta \bar{Y}} + \eta' \frac{\delta F}{\delta \bar{Y}'}\right] \frac{d\bar{Y}'}{\delta \varepsilon}\right] dx$$

$$= \int_{x_1}^{x_2} \left\{\frac{\delta}{\delta \bar{Y}}\left[\eta \frac{\delta F}{\delta \bar{Y}} + \eta' \frac{\delta F}{\delta \bar{Y}'}\right] \eta + \frac{\delta}{\delta \bar{Y}'}\left(\eta \frac{\delta F}{\delta \bar{Y}} + \eta' \frac{\delta F}{\delta \bar{Y}'}\right) \eta' \right\} dx$$

η IS NOT A FUNCTION OF $Y \Longrightarrow$

$$\Rightarrow \frac{d^2J}{d\epsilon^2}\bigg|_{\epsilon=0} = \int_{x_1}^{x_2} \eta^2 \frac{\delta^2 F}{\delta Y^2} + \eta\eta' \frac{\delta^2 F}{\delta Y \delta Y'}$$

① $$+ \eta\eta' \frac{\delta^2 F}{\delta Y' \delta Y} + \eta'^2 \frac{\delta^2 F}{\delta Y'^2} \Big) dx$$

$$= \int_{x_1}^{x_2} \lfloor \eta \; \eta' \rfloor \begin{bmatrix} \frac{\delta^2 F}{\delta Y^2} & \frac{\delta^2 F}{\delta Y \delta Y'} \\ \frac{\delta^2 F}{\delta Y' \delta Y} & \frac{\delta^2 F}{\delta Y'^2} \end{bmatrix} \begin{bmatrix} \eta \\ \eta' \end{bmatrix} dx$$

FOR $\epsilon = 0$, $Y = \overline{Y}$

$$= \int_{x_1}^{x_2} \lfloor \eta \; \eta' \rfloor \begin{bmatrix} \frac{\delta^2 F}{\delta Y^2} & \frac{\delta^2 F}{\delta Y \delta Y'} \\ \frac{\delta^2}{\delta Y \delta Y'} & \frac{\delta^2 F}{\delta Y'^2} \end{bmatrix} \begin{bmatrix} \eta \\ \eta' \end{bmatrix} dx$$

FOR A MINIMUM, > 0

" " MAX , < 0

JUST CHECK DEFINITNESS OF MATRIX, ANOTHER WAY START FROM EQ. 1

$$\frac{\delta^2 J}{\delta \epsilon^2}\bigg|_{\epsilon=0} =$$

CONSIDER $\int_{x_1}^{x_2} \eta\eta' \left(\frac{\delta^2 F}{\delta Y \delta Y'}\right) dx$

$$U = \frac{\delta^2 F}{\delta Y \delta Y'} \qquad dv = \eta \eta' dx$$

$$du = \frac{d}{dx}\left(\frac{d^2 F}{\delta Y \delta Y'}\right) dx \qquad v = \int \eta \eta' dx$$

$$\left(\begin{matrix} u = \eta & dv = \eta' dx \\ d\eta = \eta' dx & v = \eta \end{matrix} \right.$$

$$\int \eta\eta' dx = v = \eta^2 - \int \eta\eta' dx$$

$$\Rightarrow v = \tfrac{1}{2} \eta^2$$

THUS

$$2\int_{x_1}^{x_2} \eta\eta' \frac{\delta^2 F}{\delta y \delta y'} dx$$

$$= 2\frac{\eta^2}{2}\frac{\delta^2 F}{\delta y \delta y'}\Bigg|_{x_1}^{x_2} - 2\int_{x_1}^{x_2}\frac{\eta^2}{2}\frac{d}{dx}\left(\frac{\delta^2 F}{\delta y \delta y'}\right)dx$$

ASSUME = 0 FOR FIXED ENDS ($\eta(x_1) = \eta(x_2) = 0$)

PLUGGING BACK INTO 1

$$\frac{\delta^2 J}{\delta \epsilon^2}\Bigg|_{\epsilon=0} = \int_{x_1}^{x_2}\left\{\eta^2\left[\frac{\delta^2 F}{\delta y^2} - \frac{\delta}{\delta x}\left(\frac{\delta^2 F}{\delta y \delta y'}\right)\right] + \eta'^2\frac{\delta^2 F}{\delta y'^2}\right\}dx$$

$$> 0 \quad \text{MIN}$$

$$< 0 \quad \text{MAX}$$

EX(i) BOTH ENDS ARE FIXED
FIND AN EXTREMUM OF
THE FUNCTIONAL

$$J(x) = \int_0^{\pi/2}\left[\dot{x}^2(t) - x^2(t)\right]dt$$

WHICH SATISFIES $x(0) = 0$
$x(\pi/2) = 1$

NOW

$$\frac{\delta F}{\delta x} - \frac{\delta}{\delta t}\left(\frac{\delta F}{\delta x'}\right) = 0$$

$$\frac{\delta}{\delta x}\left[\dot{x}^2 - x^2\right] - \frac{\delta}{\delta t}\left[\frac{\delta}{\delta \dot{x}}(\dot{x}^2 - x^2)\right] = 0$$

$$-2x - \frac{d}{dt}(2\dot{x}) = 0$$

$$\Rightarrow \ddot{x} + x = 0$$

x IS NOT A FUNC OF $\dot{x}$
$\dot{x}$ " " " " x

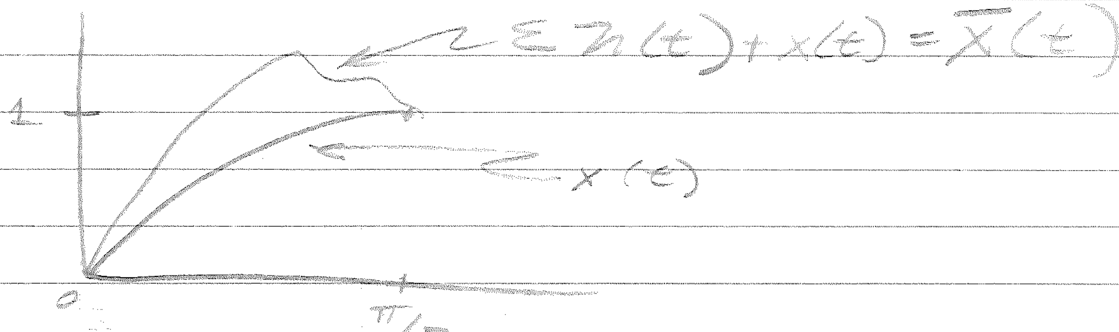
SOLUTION IS

$x(t) = A \cos t + B \sin(t)$

$x(0) = 0 \Longrightarrow A = 0$

$x(\frac{\pi}{2}) = 1 \Longrightarrow B = 1$ AND

$x(t) = \sin(t)$



LOOK @ VARIED PATH

$\bar{x}(t) = x(t) + \epsilon \eta(t)$

LET $\eta(t) = \sin 2t$

$\bar{X}(t) = \sin t + \epsilon \sin 2t$

(NOTE: BOUNDARY CONDITIONS CHECK)

$$J(x) = \int_0^{\pi/2} [\cos^2(t) - \sin^2 t)] dt$$

$$= \int_0^{\pi/2} \cos 2t \, dt = 0$$

$$J(x+\delta x) = J(\bar{X}) = \int_0^{\pi/2} [\cos(t) + 2\epsilon \cos 2t]^2 - [\sin t + \epsilon \sin 2t]^2 \; dt$$

$$= \frac{3\pi}{4} \epsilon^2 > 0$$

$\therefore X = \sin t$ IS PROBABLY A MINIMUM. LET'S CHECK

RECAL

$$\left.\frac{\delta^2 J}{\delta \epsilon^2}\right|_{\epsilon=0} = \int_{x_1}^{x_2} \underline{\left[\eta \;\; \eta'\right]} \begin{bmatrix} \frac{\delta^2 F}{\delta Y^2} & \frac{\delta^2 F}{\delta Y \delta Y'} \\ \frac{\delta^2 F}{\delta Y \delta Y'} & \frac{\delta^2 F}{\delta Y'^2} \end{bmatrix} \begin{bmatrix} \eta \\ \eta' \end{bmatrix} dx$$

$$\begin{bmatrix} \delta^2 F & \delta^2 F \\ \delta Y^2 & \delta Y \delta Y' \\ \delta^2 F / \delta Y' & \delta^2 F / \delta Y'^2 \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}$$

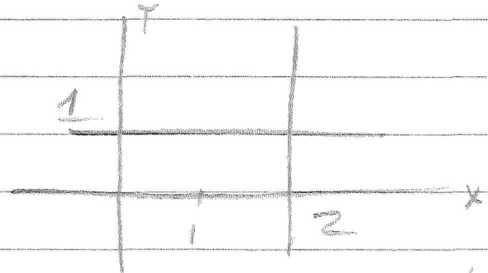
INDEFINATE $\Rightarrow$ NO CONCLUSION

SINCE $\left.\eta^2 \delta^2 F / \delta Y \delta Y'\right|_{x_1}^{x_2} = 0$

$$\frac{\delta^2 J}{\delta \epsilon^2} = \int_{x_1}^{x_2} \left\{ \eta^2 \left[\frac{\delta^2 F}{\delta Y^2} - \frac{\delta}{\delta x}\left(\frac{\delta^2 F}{\delta Y \delta Y'}\right)\right] + \eta'^2 \frac{\delta^2 F}{\delta Y'^2} \right\} dx$$

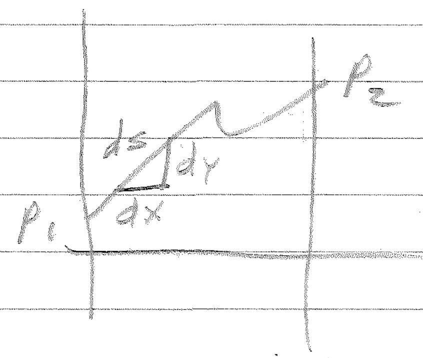
(?)

EX WE WANT TO FIND THE CURVE WITH MINIMUM ARC LENGTH BETWEEN THE POINT $Y(0) = 1$ AND THE LINE $X_2 = 2$



SOLUTION IS OBVIOUSLY $Y = 1$. LET'S SOLVE

SUPPOSE WE DON'T KNOW



S = TOTAL LENGTH $= \int_{P_1}^{P_2} ds$

$$ds = \sqrt{dx^2 + dy^2}$$

S = TOTAL LENGTH

$$= \int_{P_1}^{P_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

$$= \int_{P_1}^{P_2} \sqrt{1 - y'^2}\, dx$$

⟹

EULER LAGRANGE

$$\frac{\delta F}{\delta y} - \frac{\delta}{\delta x}\frac{\delta F}{\delta y'} = 0$$

$$0 - \frac{\delta}{\delta x}\left(\frac{\delta}{\delta y'}\sqrt{1+y'^2}\right) = 0$$

THUS

$$\frac{\delta}{\delta y'}\sqrt{1+y'^2} = C = \text{CONSTANT}$$

$$\frac{1}{2}\frac{2y'}{\sqrt{1+y'^2}} = C$$

$$y'^2 = C^2(1+y'^2)$$
$$= C'(1+y'^2)$$

$$y'^2 - C'y'^2 = C'$$
$$y'^2 = \pm\sqrt{\frac{C'}{1-C'}} = A$$

$$y = ax + b$$

BOUNDRY CONDITIONS

$$y(0) = 1 = b$$

NOW FROM TRANS:

$$\eta(x)\frac{\delta F}{\delta y'}\Bigg|_{x_1}^{x_2} = \eta(x_2)\frac{\delta F}{\delta y'} - 0$$

$$\Rightarrow \frac{\delta F}{\delta y'}\Bigg|_{x_2} = 0 = \text{CONST.}$$

FOLLOWING STEPS THROUGH

$$\Longrightarrow a = 0$$

$\therefore y = 1$

2nd DER MATRIX IS $\begin{bmatrix} 0 & 0 \\ 0 & 1/(1+y'^2)^{3/2} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

OR $\frac{\delta^2 J}{\delta \varepsilon^2} = \int_{x_1}^{x_2} \eta'^2(x)\,dx > 0$

(WED) 9-29-76

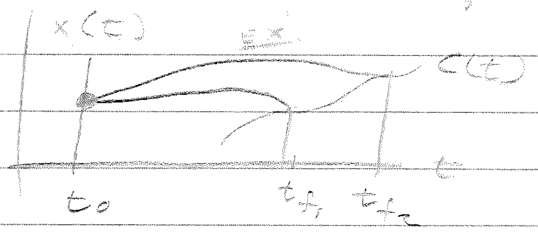
UNSPECIFIED TERMINAL TIME PROBLEM

$$J = \int_{x_1}^{x_2} F[y, y', x]\,dx$$

OR

$$J = \int_{t_0}^{t_f} \phi(x, \dot{x}, t)\,dt$$

NEED TO FIND t_f AND $x(t)$



ASSUME

$$x(t) = \hat{x}(t) + \epsilon\,\eta_x(t)$$

$\ni \hat{x}$ IS OPTIMUM SOLUTION

$$t_f = \hat{t}_f + \epsilon\,\eta_t(t_f)$$

$\ni \hat{t}_f$ IS OPTIMUM

$$\Rightarrow J = \int_{t_0}^{t_f + \epsilon\eta_t} \phi[\hat{x} + \epsilon\eta_x, \dot{\hat{x}} + \xi\dot{\eta}_x, t]\,dt$$

$$\frac{dJ}{d\epsilon} = \int_{t_0}^{\hat{t}_f + \epsilon\eta_t} \frac{d\phi}{d\epsilon}\,dt + \phi\Big|_{\hat{t}_f + \epsilon\eta} \times \frac{d(t_f + \epsilon\eta_t)}{d\epsilon}$$

$$= \int_{t_0}^{t_f + \epsilon\eta_t}\left(\frac{\delta\phi}{\delta x}\,\frac{dx}{d\epsilon} + \frac{\delta\phi}{\delta\dot{x}}\,\frac{d\dot{x}}{d\epsilon}\right)dt$$

$$+ \phi\Big|_{\hat{t}_f + \xi\eta_t} \eta_t$$

OPTIMAL OCCURS @ $\left.\frac{dJ}{d\epsilon}\right|_{\epsilon=0} = 0$

$$\left.\frac{dJ}{d\epsilon}\right|_{\epsilon=0} = 0 = \int_{t_0}^{\hat{t}_f}\left[\frac{\delta\phi[\hat{x}, \dot{\hat{x}}, t]}{\delta\hat{x}}\cdot\eta_x + \frac{\delta\phi(\hat{x}, \dot{\hat{x}}, t)}{\delta\dot{\hat{x}}}\,\dot{\eta}_x\right]dt$$

$$+ \phi(\hat{x}, \dot{\hat{x}}, t)\Big|_{\hat{t}_f} \eta_t(t_f)$$

(i) CONSIDER THE CASE WHERE t_f & $x(t_f)$ ARE NOT RELATED $\Rightarrow$ η_x & η_t ARE NOT RELATED. THEN WE NEED TO SET BOTH TERMS $= 0$

BEFORE PROCEEDING LET'S SIMPLIFY PARTS INTEGRATION

$$u = \frac{d\phi}{d\dot{x}} \qquad dv = \dot{\eta}_x \, dt$$

$$du = \frac{d}{d}\left(\frac{d\phi}{d\dot{x}}\right) d \qquad v = \eta_x$$

$$\Rightarrow \frac{dJ}{d\epsilon}\bigg|_{\epsilon=0} = \int_{t_0}^{\hat{t}_f} \left[\frac{d\phi}{dx} \eta_x \, dt \right.$$

$$+ \eta_x \frac{d\phi}{d\dot{x}}\bigg|_{t_0}^{\hat{t}_f} - \int_{t_0}^{\hat{t}_f} \eta_x \frac{d}{dt}\left(\frac{d\phi}{d\dot{x}}\right) dt$$

$$+ \phi\Big|_{\hat{t}_f} \eta_t(\hat{t}_f)$$

$$\frac{dJ}{d\epsilon}\bigg|_{\epsilon=0} = \int_{t_0}^{t_f} \eta_x \left[\frac{d\phi}{dx} - \frac{d}{dt}\left(\frac{d\phi}{d\dot{x}}\right) \right] dt$$

$$+ \eta_x \frac{d\phi}{d\dot{x}}\bigg|_{t_0}^{\hat{t}_f} + \phi\Big|_{\hat{t}_f} \eta_t(\hat{t}_f) = 0$$

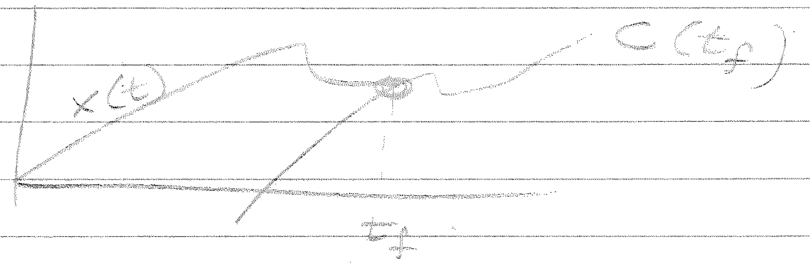
GIVES EULER LAGRANGE EQ'S:

$$\frac{\delta \phi}{\delta \hat{x}} - \frac{\delta}{\delta t}\left(\frac{d\phi}{d\hat{\dot{x}}}\right) = 0$$

$$\eta_x \frac{d\phi}{\delta \hat{\dot{x}}}\Big|_{t_0}^{\hat{t}_f} = 0$$

THIS IS SOLUTION WHEN η_t & η_x ARE NOT RELATED.

(ii) FOR THE CASE WHERE t_f AND $x(t_f)$ ARE RELATED BY $x(t_f) = C(t_f)$



@ $t = t_f$

$$\hat{x}(t_f) + \varepsilon \eta_x(t_f) = C(t_f) = x(t_f)$$

$$\hat{x}[\hat{t}_f + \varepsilon \eta_t(t_f)] + \varepsilon \eta_x[\hat{t}_f + \varepsilon \eta_t(t_f)] = C(t_f)$$

DIFFERENTIATE W.R.T. ε & EVALUATING RESULT @ $\varepsilon = 0$ YIELDS

$$\frac{d\hat{x}}{dt_f}\frac{dt_f}{d\varepsilon} + \eta_x(t_f) + \varepsilon \frac{\delta \eta_x(t_f)}{\delta t_f}\frac{dt_f}{d\varepsilon}$$

$$= \frac{\delta C(t_f)}{dt_f}\frac{dt_f}{d\varepsilon}$$

EVALUATE @ $\varepsilon = 0$ ⟶

$$\frac{\delta \hat{x}}{\delta \hat{t}_f} \eta_t(\hat{t}_f) + \eta_x(\hat{t}_f) + 0$$

$$= \frac{\delta c(\hat{t}_f)}{\delta \hat{t}_f} \eta_t(\hat{t}_f)$$

OR

$$\eta_x(\hat{t}_f) = \eta_t(\hat{t}_f) \left[\frac{\delta c}{\delta \hat{t}_f} - \frac{\delta \hat{x}}{\delta \hat{t}_f} \right]$$

RECALL

$$\left.\frac{dJ}{d\epsilon}\right|_{\epsilon=0} = \int_{t_0}^{t} \eta_x \left[\frac{d\phi}{\delta \hat{x}} - \frac{d}{dt}\left(\frac{d\phi}{d\dot{\hat{x}}}\right) \right] dt$$

$$+ \eta_x \frac{d\phi}{d\dot{\hat{x}}} \Big|_{t_0}^{t_f} + \phi \Big|_{\hat{t}_f} \eta_t(t_f) = 0$$

PLUG IN

$$\left.\frac{dJ}{d\epsilon}\right|_{\epsilon=0} = \int_{t_0}^{t_f} \eta_x \left[\frac{d\phi}{\delta \hat{x}} - \frac{d}{dx}\left(\frac{d\phi}{d\dot{\hat{x}}}\right) \right] dt$$

$$+ \eta_t(\hat{t}_f) \left(\frac{dc}{d\hat{t}_f} - \frac{\delta \hat{x}}{\delta \hat{t}_f} \right) \frac{d\phi}{d\dot{\hat{x}}} \Big|_{\hat{t}_f}$$

$$- \eta_t(t_0) \frac{\delta \phi}{\delta \dot{\hat{x}}} \Big|_{t=t_0} + \eta_t(\hat{t}_f) \phi \Big|_{\hat{t}_f}$$

$$= 0$$

REARRANGE TERMS:

$$\frac{\delta J}{\delta \epsilon}\bigg|_{\epsilon=0} = \int_{t_0}^{t_f} \eta_x \left[\frac{d\phi}{dx} - \frac{\delta}{\delta x}\left(\frac{d\phi}{d\dot{x}}\right)\right] dt \quad ①$$

$$+ \eta_t(\hat{t}_f)\left[\left\{\dot{c}(\hat{t}_f) - \dot{\hat{x}}(\hat{t}_f)\right\}\frac{d\phi}{d\dot{x}}\bigg|_{\hat{t}_f} + \phi\bigg|_{\hat{t}_f}\right] \quad ②$$

$$- \eta_x(t_0)\frac{d\phi}{d\dot{x}}\bigg|_{t_0} = 0$$

③

NECESSARY CONDITIONS

① = 0 (EULER LAGRANGE

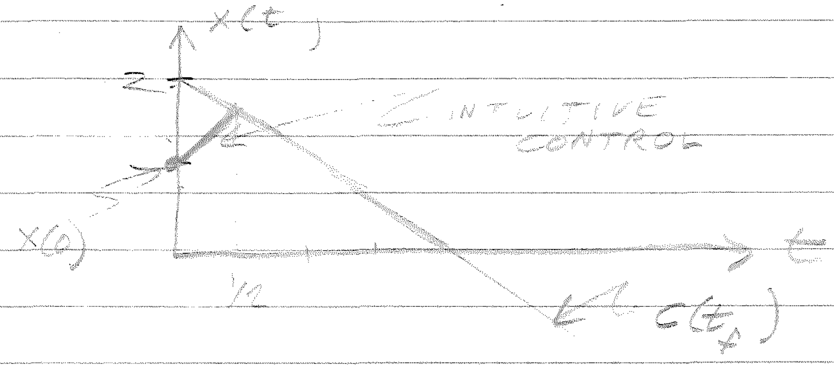
② = 0

③ = 0

EXAMPLE: WE WANNA MINIMIZE

$$J = \int_0^{t_f} (1 + \dot{x}^2)^{\frac{1}{2}} dt$$

$$\ni x(0) = 1 \quad \ni x(t_f) = c(t_f) = 2 - t_f$$



J WILL MINIMIZE ARC LENGTH OF x(t)

$$ds = \sqrt{dx^2 + dt^2} = \sqrt{\left(\frac{dx}{dt}\right)^2 + 1}\, dt$$

① WILL GIVE (AS BEFORE) FAMILY OF LINES

$$x(t) = at + b$$

③ $\eta_x(t_0) = 0$ (ALREADY SATISFIED)

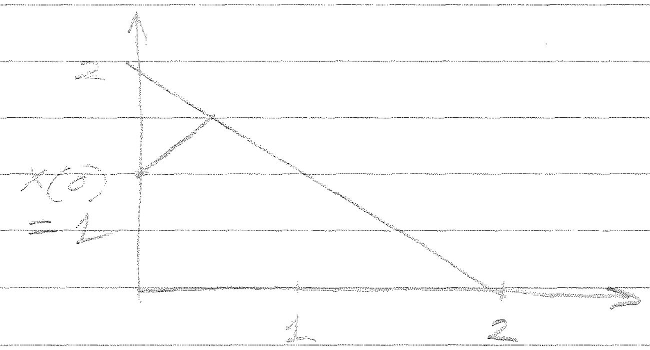
② (WILL DO NEXT LECTURE)

10-1-76 (ERI)

EXAMPLE:

$$J = \int_0^{t_f} (1 + \dot{x}^2)^{\frac{1}{2}} dt = \int \phi \, dt$$

$$x(0) \Big|_{x(t)} \quad ; x(t_f) = c(t_f) = 2 - t_f$$



INITIAL TIME & STATE FIXED

FINAL TIME ISN'T

EULER LAGRANGE CONDITIONS

$$\frac{\delta \phi}{\delta x} - \frac{\delta}{\delta t}\left(\frac{\delta \phi}{\delta \dot{x}}\right) = 0$$

AND

$$\eta_t \left[\{ \dot{c} - \dot{x}) \frac{\delta \phi}{\delta \dot{x}} + \phi \right] \Big|_{t = t_f} = 0$$

$$\eta_x \frac{\delta \phi}{\delta \dot{x}} \Big|_{t = t_0} = 0$$

$$\frac{\delta \phi}{\delta x} = 0$$

$$\frac{\delta \phi}{\delta \dot{x}} = \frac{\dot{\hat{x}}}{\sqrt{1 + \dot{\hat{x}}^2}}$$

$$(0) \frac{d}{dt}\left[\frac{\dot{\hat{x}}}{\sqrt{1 + \dot{\hat{x}}^2}}\right] = 0 \Rightarrow \frac{\dot{\hat{x}}}{\sqrt{1 + \dot{\hat{x}}^2}} = c$$

$$\dot{\hat{x}}^2 = c^2 (1 + \dot{\hat{x}}^2)$$

$$\dot{\hat{x}}^2 = \frac{c^2}{1 - c^2} = a^2$$

$$\Rightarrow \dot{\hat{x}} = a$$

$$\hat{x} = at + b$$

FINDING CONSTANTS

TRANS. COND:

$$\left(\dot{c} - \dot{\hat{x}}\right)\frac{\delta \phi}{\delta \dot{\hat{x}}} + \phi \Big|_{t = \hat{t}_f} = 0$$

$$(-1 - \dot{\hat{x}}) \frac{\dot{\hat{x}}}{\sqrt{1 + \dot{\hat{x}}^2}} + \sqrt{1 + \dot{\hat{x}}^2} \Big|_{t = \hat{t}_f} = 0$$

$$= \frac{(-1 - \dot{\hat{x}})\dot{\hat{x}} + 1 + \dot{\hat{x}}^2}{\sqrt{1 + \dot{\hat{x}}^2}}$$

$$= 1 - \dot{\hat{x}} \Big|_{t = 0}$$

$$\dot{\hat{x}} \Big|_{t = \hat{t}_f} = 1 = a$$

$$\Rightarrow \hat{x} = t + b$$

ALSO, SINCE $x(0) = 1 \Rightarrow b = 1$

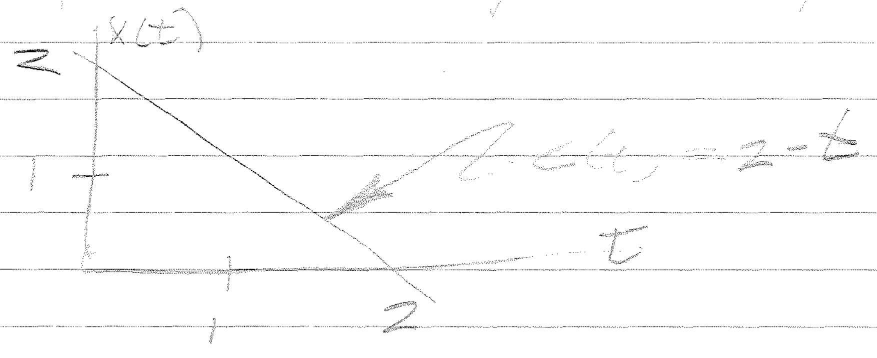
AND

$$\hat{x} = t + 1$$

NOW

$x(t_f) = C(t_f)$

$t_f + 1 = 2 - t_f \Longrightarrow t_f = 1/2$



[illegible]

CONSIDER

$$J = \int_{t_0}^{t_f} \phi\left[x_1 x_2, \ldots \dot{x}_1 \dot{x}_2, \ldots x_n, t\right] dt$$

$\left.\begin{array}{l} x(t_0) \\ x(t_1) \end{array}\right\}$ MAY OR MAY NOT BE SPECIFIED

SOLVING IN GENERAL

$x(t) = \hat{x}(t) + \epsilon \eta_x(t)$

OR

$x_1(t) = \hat{x}_1(t) + \epsilon \eta_{1x}(t)$

$\vdots$

$x_n(t) = \hat{x}_n(t) + \epsilon \eta_{nx}$

PLUG THESE INTO THE EQ. FOR J, DIFFERENTIATE WRT ϵ, & SET TO 0

$\Longrightarrow$

$$\frac{dJ}{d\epsilon} = \int_{t_0}^{t_f} \frac{d\phi}{d\epsilon}\, dt$$

(t_0 AND t_f FIXED)

$$\frac{dJ}{d\epsilon} = \int_{t_0}^{t_f} \Big[\Big(\frac{d\phi}{dx_1}\frac{dx_1}{d\epsilon} + \frac{\delta\phi}{\delta\dot{x}_1}\frac{\delta\dot{x}_1}{\delta\epsilon}\Big)$$

$$+ \Big(\frac{\delta\phi}{\delta x_2}\frac{\delta x_2}{\delta\epsilon} + \frac{\delta\phi}{\delta\dot{x}_2}\frac{\delta\dot{x}_2}{\delta\epsilon}\Big)$$

$$+ \ldots + \Big(\frac{\delta\phi}{\delta x_n}\frac{\delta x_n}{\delta\epsilon} + \frac{\delta\phi}{\delta\dot{x}_n}\frac{\delta\dot{x}_n}{\delta\epsilon}\Big)\Big] dt$$

$$= \int_{t_0}^{t_f} \Big[\Big(\frac{\delta\phi}{\delta x_1}\eta_{1x} + \frac{\delta\phi}{\delta\dot{x}_1}\dot{\eta}_{1x}\Big) + \ldots\Big] dt$$

$$= \int_{t_0}^{t_f} \sum_{i=1}^{n} \Big(\frac{d\phi}{dx_i}\eta_{ix} + \frac{\delta\phi}{\delta\dot{x}_i}\dot{\eta}_{ix}\Big) dt$$

LET

$$u = \frac{d\phi}{d\dot{x}_i} \qquad dv = \dot{\eta}_{ix}\, dt$$

$$du = \frac{d}{dt}\Big(\frac{d\phi}{\delta\dot{x}_i}\Big) dt \qquad v = \eta_{ix}$$

$$\Rightarrow \frac{dJ}{d\epsilon} = \int_{t_0}^{t_f} \sum_{i=1}^{n} \frac{\delta\phi}{\delta x_i}\eta_{ix}\, dt + \frac{\delta\phi}{\delta\dot{x}_i}\eta_{ix}\Big|_{0}^{0}$$

$$- \int_{t_0}^{t_f} \sum_{i=1}^{n} \eta_{ix} \frac{\delta}{\delta t}\Big(\frac{\delta\phi}{\delta\dot{x}_i}\Big) dt$$

$$= \frac{d\phi}{\delta\dot{x}_i}\eta_{ix}\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \sum_{i=1}^{n} \eta_{ix}\Big(\frac{d\phi}{\delta x_i}$$

$$- \frac{d}{dt}\Big(\frac{d\phi}{d\dot{x}_i}\Big)\Big) dt.$$

@ $\epsilon = 0$, $x = \hat{x}$ ETC. THIS GIVES

1) EULER - LAGRANGE EQUATIONS

$$\frac{\delta\phi}{\delta x_i} - \frac{d}{dt}\left(\frac{d\phi}{\delta \dot{x}_i}\right) = 0$$

TRANSVERSALITY:

$$\frac{\delta\phi}{\delta \dot{x}_i}\eta_{ix}\Big|_{t_0}^{t_f} = 0$$

IF ENDPOINT IS FIXED

$\eta_{ix}(t_0)$ OR $\eta_{ix}(t_f) = 0$

IF t_f IS NOT FIXED

$$\frac{dJ}{d\epsilon} = \int_{t_0}^{t} \frac{d\phi}{d\epsilon}\,dt + \phi\Big|_{t=t_f}\frac{dt_f}{d\epsilon}$$

$$\Rightarrow t_f = \hat{t}_f + \epsilon\,\eta_t(t_f)$$

$$\frac{dJ}{d\epsilon} = \int_{t_0}^{t_f}\left[\sum_{i=1}^{n}\left(\frac{d\phi}{dx_i}\eta_{ix} + \frac{d\phi}{\delta\dot{x}_i}\dot{\eta}_{ix}\right)\right]dt$$

$$+ \phi\Big|_{t=t_f}\eta_t(t_f)$$

SO, WE GET A THIRD CONDITION:

$$\phi\Big|_{t=\hat{t}_f}\eta_t(t_f) = 0$$

'ANOTHER WOULD BE

$$x(t_f) = c(t_f) \Rightarrow \hat{x}(t_f) + \epsilon\eta_x = c(t$$

DIFFERENTIATE $\rightarrow$

$$\frac{d}{d\epsilon}(\quad) \Rightarrow \frac{\delta \hat{x}}{\delta \epsilon}$$

$$\frac{\delta \hat{x}}{\delta t_f}\underbrace{\frac{dt_f}{d\epsilon}}_{\overset{\uparrow}{\eta_t}} + \eta_x + \epsilon\,\frac{\delta \eta_x}{\delta t_f}\underbrace{\frac{dt_f}{\delta \epsilon}}_{\overset{\uparrow}{\eta_t}} = \frac{\delta c}{\delta t_f}\underbrace{\frac{dt_f}{\delta \epsilon}}_{\overset{\uparrow}{\eta_t}}$$

@ $\epsilon = 0$

$$\eta_t\left(\frac{\delta \hat{x}}{\delta t_f} - \frac{\delta c}{\delta t_f}\right) = 0 \Leftarrow \text{NEW CONSTRAINT}$$

EX. FIND AN EXTREMUM FOR THE
p. 149 FUNCTIONAL

$$J(x) = \int_0^{\pi/4} \left[x_1^2 + \dot{x}_1\dot{x}_2 + \dot{x}_2^2\right] dt$$

$x_1(0) = 1 \;;\; x_1\left(\frac{\pi}{4}\right) = 2$
$x_2(0) = 3/2 \;;\; x_2(\pi/4)$ IS FREE)

APPLY EULER LAGRANGE:

$$\frac{\delta\phi}{\delta x_1} - \frac{d}{dt}\left(\frac{d\phi}{d\dot{x}_1}\right) = 0 \Rightarrow 2x_1 - \frac{d}{dt}(\dot{x}_2) = 0$$

$$\frac{\delta\phi}{\delta x_2} - \frac{d}{dt}\left(\frac{d\phi}{\delta \dot{x}_2}\right) = 0 \Rightarrow 0 - \frac{d}{dt}(\dot{x}_1 + 2\dot{x}_2) = 0$$

OR $2x_1 - \ddot{x}_2 = 0$ (1)
$-\ddot{x}_1 - 2\ddot{x} = 0$ (2)

$2(1) - (2) \Longrightarrow 4x_1 + \ddot{x}_1 = 0 \Rightarrow \ddot{x}_1 = -4x_1$

$$x_1(t) = C_1 \cos 2t + C_2 \sin 2t$$

$\ddot{x}_2(t) = 2x_1$

$= 2c_1 \cos 2t + 2c_2 \sin 2t$

$\dot{x}_2(t) = c_1 \sin 2t - c_2 \cos 2t - c_3$

$= \frac{c_1}{2} \sin 2t - \frac{1}{2} c_2 \cos 2t$

$+ c_3 t + c_4$

GOTTA FIND FOUR UNKNOWN CONSTANTS

NOW

$\frac{\delta \phi}{\delta \dot{x}_1} \eta_{1x} \Big|_{t_0}^{t_f} = 0$

$\eta_{1x}(t_0) = 0$
$\eta_{1x}(t_f) = 0$
$\eta_{2x}(t_0) = 0$
$\eta_{2x}(t_f) \neq 0$

THUS $\frac{\delta \phi}{\delta \dot{x}_2} \Big|^{t_f} = 0$

$\Rightarrow \dot{x}_1 + 2\dot{x}_2 \Big|_{t_f} = 0$

$-2c_1 \sin 2t + 2 \cos 2t$
$+ 2c_1 \sin 2t - 2 \cos 2t + c_3 \Big|_{t_f} = 0$

$\Longrightarrow c_3 = 0$

$c_1, c_2 \,\&\, c_4$ CAN BE GOT BY GIVEN BOUNDARY CONDITION)

$x_1(0) = 1 = c_1$
$x_2(0) = 3/2 = -\frac{1}{2} c_1 + c_4 \Rightarrow c_4 = 2$

$x_1\left(\frac{\pi}{4}\right) = c_1 \cdot 0 + c_2 \cdot 1 = 2 \Rightarrow c_2 = 2$

10-4-76 (MON)

VARIATIONAL NOTATION

$$J = \int_{t_0}^{t_f} \phi(x, \dot{x}, t)\,dt$$

$$x = \hat{x} + \delta x$$

$$\phi(\hat{x} + \delta x, \dot{\hat{x}} + \delta\dot{\hat{x}}, t) = \phi(\hat{x}, \dot{\hat{x}}, t) + \left(\frac{\delta\phi}{\delta x}\delta x + \frac{\delta\phi}{\delta \dot{x}}\delta\dot{x}\right)_{\hat{x}, \dot{\hat{x}}} + H.O.T.$$

$$\Rightarrow J = \int_{t_0}^{t_f} \phi(\hat{x}, \dot{\hat{x}}, t)\,dt + \int_{t_0}^{t_f}\left(\frac{\delta\phi}{\delta \hat{x}}\delta x + \frac{\delta\phi}{\delta \dot{\hat{x}}}\right)dt + \int_{t_0}^{t_f}(H.O.T)\,dt$$

$$= \int_{t_0}^{t_f} \phi(x, \dot{x}, t)\,dt$$

$$\Delta J = \int_{t_0}^{t_f} \phi(x, \dot{x}, t)\,dt - \int_{t_0}^{t_f} \phi(\hat{x}, \dot{\hat{x}}, t)\,dt$$

$$= \underbrace{\int_{t_0}^{t_f}\left(\frac{\delta\phi}{\delta \hat{x}}\delta \hat{x} + \frac{\delta\phi}{\delta \dot{\hat{x}}}\delta \dot{\hat{x}}\right)dt}_{\text{VARIATION OF } J} + O\|\delta x\|^2$$

IF $\hat{x}$ IS AN EXTREMUM, THE VARIATION OF J MUST VANISH ON $\hat{x}$. ie

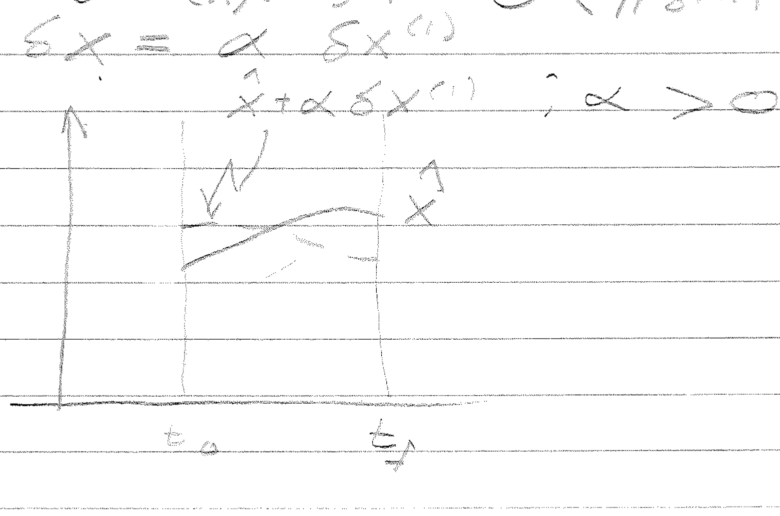
$$\delta J(\hat{x}, \delta x) = 0 \quad \forall \text{ ADMISSIBLE } \delta x$$

PROOF

$$\Delta J = J(\hat{x} + \delta x) - J(\hat{x})$$

$$= \delta J(x, \delta x) + O(\|\delta x\|^2)$$

(i) PICK $\delta x = \alpha\,\delta x^{(1)}$



SUPPOSE

$\delta J[\hat{x}, \alpha\,\delta x^{(1)}] < 0$

FURTHERMORE (INCORRECTLY)

ASSUME $\exists\ \delta x \ni$

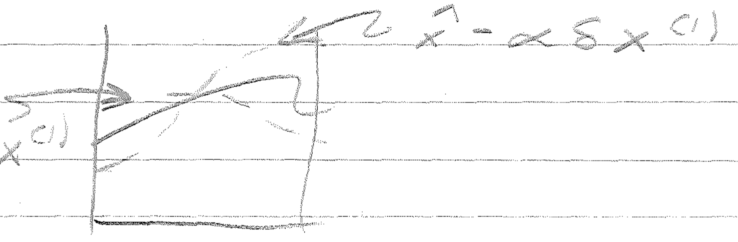
$\delta J(\hat{x}, \delta x) \neq 0$

AGAIN, SUPPOSE

$\delta J[\hat{x}, \alpha\,\delta x^{(1)}) < 0$

$\Delta J \cong \delta J = \alpha\,\delta J(\hat{x}, \delta x^{(1)}) < 0$

WE CAN DO THIS CAUSE δJ IS LINEAR (FIRST ORDER APPROX.)



(ii) CORNSIDER $\delta x = -\alpha\,\delta x^{(1)}$

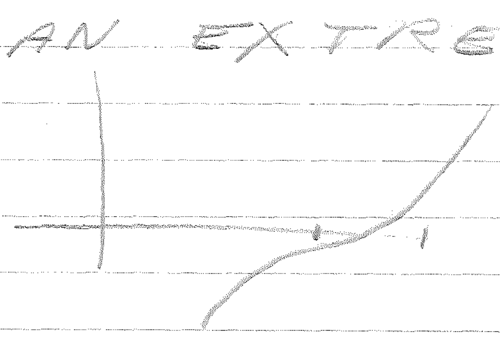
$\Delta J \cong \delta J(\hat{x}, -\alpha\,\delta x_1^{(1)})$

$= -\alpha\,\delta J(\hat{x}, \delta x^{(1)}) > 0$

NOW

$\delta J(\hat{x}, \delta x^{(1)}) < 0$

THUS, IT PASSES THRU AXIS AND IS (BY CONTRADICTION), NOT AN EXTREMA



CONSIDER

$$\delta J = \int_{t_0}^{t_f} \left(\frac{\delta \phi}{\delta \hat{x}} d\hat{x} + \frac{\delta \phi}{\delta \dot{\hat{x}}} \delta \dot{\hat{x}} \right) dt = 0$$

$$= \int_{t_0}^{t_f} \frac{\delta \phi}{\delta \hat{x}} \delta \hat{x} dt + \frac{\delta \phi}{\delta \dot{\hat{x}}} \delta \hat{x} \Big|_{t=t_0}^{t_f} - \int_{t_0}^{t_f} \left[\frac{d}{dt} \left(\frac{\delta \phi}{\delta \dot{\hat{x}}} \right) dt \right] = 0$$

GIVES THE CONDITIONS

$$\frac{\delta \phi}{\delta \dot{\hat{x}}} \delta x \Big|_{t=t_0}^{t_f} = 0$$

$$\frac{\delta \phi}{\delta \hat{x}} - \frac{d}{dt}\left(\frac{\delta \phi}{\delta \dot{\hat{x}}} \right) = 0$$

VARIATION OF FUNCTION WITH TERMINAL TIMES <u>NOT</u> FIXED

WE HAVE CONSIDERED $x(t_f) = c(t_f)$
LET'S LOOK @ ANOTHER CASE:

$$J = \int_{t_0}^{t_f} \phi(x, \dot{x}, t) dt$$

$$\Delta J = \delta J + O(\|\delta x\|^2)$$

$$= \left(\frac{\delta J}{\delta x}\right)^T (x - \hat{x}) + \left(\frac{\delta J}{\delta \dot{x}}\right)^T (\dot{x} - \dot{\hat{x}})$$

$$+ \frac{\delta J}{\delta t_f}(t_f - \hat{t}_f) + H.O.T.$$

$$\delta J = \int_{t_0}^{t_f} \left[\left(\frac{\delta \phi}{\delta x}\right)^T \delta x + \left(\frac{\delta \phi}{\delta \dot{x}}\right)^T \delta \dot{x} \right] dt$$

$$+ \phi(x, \dot{x}, t) \Big|_{t=\hat{t}_f} \delta t_f$$

OR, USING PART INTEGRATION -

$$\delta J = \int_{t_0}^{t_f} \delta x^T \left(\frac{\delta \phi}{\delta x}\right) + (\delta x)^T \frac{\delta \phi}{\delta \dot{x}}\Big|_{t_0}^{t_f}$$

$$+ \int_{t_0}^{t_f} (\delta x)^T \frac{\delta}{\delta t}\left(\frac{\delta \phi}{\delta \dot{x}}\right) dt$$

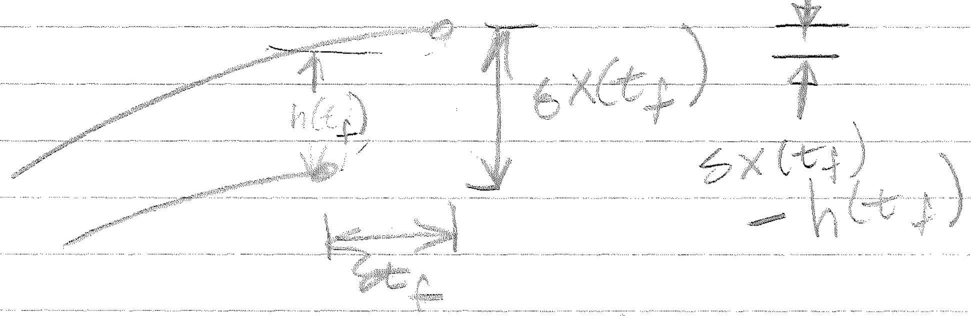
$$+ \phi\Big|_{t_f} dt_f - \cancel{\phi\Big|_{t_0} dt_0}$$

LET $\delta x = h$, $\delta \dot{x} = \dot{h}$

① $$\Rightarrow \delta J = \int_{t_0}^{t_f} h^T \left(\frac{\delta \phi}{\delta x}\right) dt$$

$$+ h^T \frac{\delta \phi}{\delta \dot{x}}\Big|_{t_0}^{t_f} - \int_{t_0}^{t_f} h^T \frac{d}{dt}\left(\frac{d\phi}{\delta \dot{x}}\right) dt$$

$$+ \phi\Big|_{t_f} \delta t_f$$

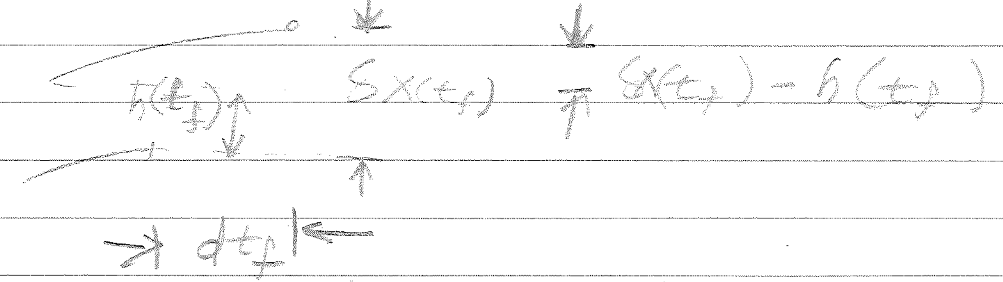


$$\dot{X}(t_f) \simeq \frac{\delta X(t_f) - h(t_f)}{\delta t_f}$$

$$\text{OR } h(t_f) = \delta X(t_f) - \dot{x}(t_f)\,\delta t_f$$

PLUG THIS IN TO ①

10-6-76 (WED)



$$\delta x(t_f) = h(t_f) + \dot{x}(t_f)\,\delta t_f$$

NOW

$$J = \int_{t_0}^{t_f} \Phi(x, \dot{x}, t)\,dt$$

$$\delta J = \int_{t_0}^{t_f} h^T \frac{\delta \phi}{\delta x}\,dt + h' \frac{\delta \phi}{d\dot{x}}\Big|_{t_0}^{t_f}$$

$$- \int_{t_0}^{t_f} h^T \frac{d}{dt}\left(\frac{d\phi}{\delta \dot{x}}\right) dx + \phi\Big|_{t_f} dt_f$$

$$= \int_{t_0}^{t_f} h^T \left[\frac{\delta \phi}{\delta x} - \frac{d}{dt}\left(\frac{d\phi}{d\dot{x}}\right)\right] dt$$

$$+ h^T d\phi/d\dot{x}\Big|_{t_0}^{t_f} + \phi\Big|_{t_f} \delta t_f$$

$$= \int_{t_0}^{t_f} h^T \left(\frac{\delta \phi}{\delta x} - \frac{d}{dt}\left(\frac{d\phi}{d\dot{x}}\right)\right\} dt \quad ①$$

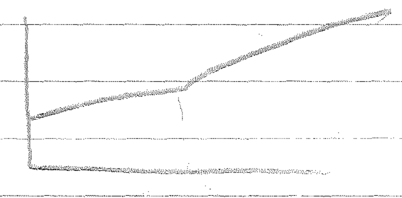
$$② \begin{cases} + \left(\delta x^T - \dot{x}^T(t_f)\,\delta t_f\right) \dfrac{\delta \phi(x(t_f), \dot{x}(t_f), t_f)}{\delta \dot{x}(t_f)} \\ - \overset{= 0 \text{ FOR } t_0 \text{ FIXED}}{h'(t_0)}\, d\Phi[x(t_0)\,\dot{x}(t_0)\,t_0]/(\delta \dot{x}(t_0)) \\ + \phi[x(t_f), \dot{x}(t_f), t_f]\,\delta t_f \end{cases}$$

$= 0$ (FOR EXTREMUM)

$\Rightarrow$ ① $= 0$

② $= 0$

PIECEWISE SMOOTH CURVE



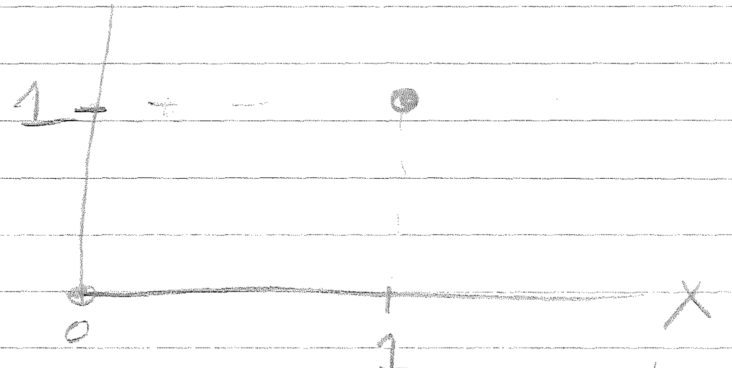
WORK IN 2 PROBLEMS

EXAMPLE

MINIMIZE

$$J = \int_0^1 x^2(2 - \dot{x})^2 dt$$

$x(0) = 0 \qquad x(1) = 1$



BY INSPECTION $J_{min} = 0$

LOOK @ EULER LAGRANGE:

$$\frac{\partial f}{\partial x} - \frac{\partial}{\partial t}\left(\frac{df}{d\dot{x}}\right)$$

$$= 2x(2 - \dot{x})^2 + \frac{d}{dt}\left[2x^2(2 - \dot{x})\right] = 0$$

OR

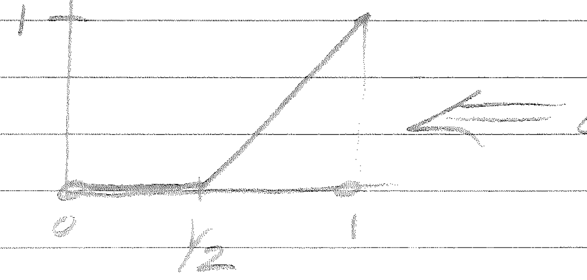
$$x^2 \ddot{x} + x \dot{x}^2 - 4x = 0$$

NOW $x^2(2 - \dot{x})^2 = 0$

$\Rightarrow x^2 = 0$

$\Rightarrow x = 2t + c$

$x(0) = 0 \Rightarrow c = -1$



$\Longleftarrow$ AN OPTIMAL SOLUTION

CONSIDER MINIMIZING

$$J = \int_{t_0}^{t_f} \phi(x, \dot{x}, t)\, dt$$

WE WISH TO ALLOW "CORNERS"

$$J = \int_{t_0}^{t_1^-} \phi(x, \dot{x}, t)\, dt + \int_{t_f}^{t_0} \phi(x, \dot{x}, t)\, dt$$

$$= J_1 + J_2$$

$$\delta J = \delta J_1 + \delta J_2 = \delta \int_{t_0}^{t^-} \phi(t) + \delta \int_{t_f}^{t_f} \phi(t)\, dt$$

WE DO <u>NOT</u> KNOW K_1

$$\delta J = \int_{t_0}^{t'^-} h^T \left[\frac{\delta\phi}{\delta x} - \frac{d}{dt}\left(\frac{\delta\phi}{\delta \dot{x}}\right)\right] dx$$

$$+ \left(\delta x' - \dot{x}'(t')\, \delta t'\right) \frac{\delta\phi}{\delta \dot{x}}\Big|_{t'}$$

$$+ \phi|_{t_1} \delta t$$

SAME AS BEFORE

$$+ \left\{ \begin{aligned} &\int_{t_1'}^{t_f} h' \left[\frac{\delta\phi}{\delta x} - \frac{d}{dt}\left(\frac{d\phi}{d\dot{x}}\right)\right] dt \\ &- \left(\delta x(t_f') - \dot{x}'(t_f')\, \delta t_f'\right) \frac{d\phi}{d\dot{x}} \\ &- \phi|_{t_f'}\ \delta t_{f'} \end{aligned} \right.$$

NOTE $t_+' = t_-'$

$x(t_+) = x(t_-)$

BUT $\dot{x}(t_+) \neq \dot{x}(t_-)$

$$\Rightarrow \frac{\delta\phi}{\delta x} - \frac{d}{dt}\left(\frac{d\phi}{d\dot{x}}\right) = 0$$

$$\left[\phi\Big|_{t'} - \dot{x}'(t')\frac{d\phi}{d\dot{x}}\Big|_{t_-} - \left(\phi\Big|_{t_+'} + \dot{x}'(t_+')\frac{d\phi}{d\dot{x}}\Big|_{t_f'}\right)\right]\delta t$$

$$\delta x'(t_-')\frac{\delta\phi}{\delta\dot{x}}\Big|_{t_-'} - \delta x'(t_f')\frac{d\phi}{d\dot{x}}\Big|_{t_f'}$$

$$= 0$$

OR

WEITRSTRASS-ERDMAN CONDITIONS

$$\left(\phi - \dot{x}'\frac{\delta\phi}{\delta\dot{x}}\right)\Big|_{t_-'} = \left(\phi - \dot{x}'\frac{d\phi}{d\dot{x}}\right)\Big|_{t_f'} \quad ①$$

$$\frac{\delta\phi}{\delta\dot{x}}\Big|_{t_-'} = \frac{\delta\phi}{\delta\dot{x}}\Big|_{t_+'} \quad ②$$

HOMEWORK:

4-6

-8

-9

-10 a, c

-11

10/8/76 (FRI)

WEIRSTRASS CORNER CONDITON

$$\phi - \dot{x}' \frac{\delta\phi}{\delta\dot{x}}\Big|_{t=t_1^-} = \phi - \dot{x}' \frac{\delta\phi}{\delta\dot{x}}\Big|_{t=t_1^+} \quad (1$$

AND ~~OR~~ $$\frac{\delta\phi}{\delta\dot{x}}\Big|_{t=t_1^-} = \frac{\delta\phi}{\delta\dot{x}}\Big|_{t=t_1^+}, \quad (2)$$

EX: FIND A PIECEWISE SMOOTH CURVE THAT MINIMIZED

$$J(x) = \int_0^2 \dot{x}^2(t)\,[1-\dot{x}(t)]^2\,dt$$

SUBJECT TO

$$x(0) = 0 \qquad x(2) = 1$$

$$\frac{\delta\phi}{\delta x} - \frac{d}{dt}\left(\frac{\delta\phi}{\delta\dot{x}}\right) = 0, \quad \phi = \dot{x}^2(1-\dot{x}^2)$$

$$\Longrightarrow 0 - \frac{d}{dt}\left[\frac{\delta\phi}{\delta\dot{x}}\right] = -\left[\frac{\delta}{\delta\dot{x}} \frac{\delta\phi}{\delta\dot{x}}\right]\frac{d\dot{x}}{dt} = 0$$

$$\left(\frac{d}{dt} f(\dot{x}) = \frac{\delta f}{\delta\dot{x}} \frac{\delta\dot{x}}{\delta t}\right)$$

EULER'S EQ. BECOMES, IN THIS CASE BECOMES

$$\ddot{x}\,\frac{\delta^2\phi}{\delta\dot{x}^2} = 0$$

OR:

$$\begin{cases} \ddot{x} = 0 \quad \Longrightarrow \quad x = at + b \\ \dfrac{\delta^2\phi}{\delta\dot{x}^2} = (2 - 12\dot{x} + 12\dot{x}^2) = 0 \end{cases}$$

SOLUTION OF THIS MUST BE OF FORM

$$(\dot{x} + C_1)(\dot{x} + C_2) \cdots = 0$$

THIS WOULD ALSO GIVE

$$X(t) = at + b$$

#2

$x = at + b$

LOOK @ CORNER CONDITION

$$\left.\frac{\delta\phi}{\delta\dot{x}}\right|_{t'_-} = \left.\frac{\delta\phi}{\delta\dot{x}}\right|_{t'_+}$$

$$\frac{\delta\phi}{\delta\dot{x}} = 2\dot{x} - 6\dot{x}^2 + 2\dot{x}^3$$

$$= 2\dot{x}(1 - 3\dot{x} + 2\dot{x}^2)$$

$$= 2\dot{x}(2\dot{x}-1)(\dot{x}-1)$$

THUS

$$\left.2\dot{x}(2\dot{x}-1)(\dot{x}-1)\right|_{t'_-} = \left.2\dot{x}(2\dot{x}-1)(\dot{x}-1)\right|_{t'_+}$$

$$\dot{x}(t'_-) = 0, \tfrac{1}{2}, 1$$

$$\dot{x}(t'_+) = 0, \tfrac{1}{2}, 1$$

ANY SET OF THESE #'S MUST HOLD. LET'S LOOK @ SECOND CORNER CONDITION:

$$\left.\phi - \dot{x}\frac{\delta\phi}{\delta\dot{x}}\right|_{t=t'_-} = \left.\phi - \dot{x}'\frac{\delta\phi}{\delta\dot{x}}\right|_{t=t'_+}$$

$$= \dot{x}^2(1-\dot{x})^2 - 2\dot{x}^2(2\dot{x}-1)$$

$$= \dot{x}^2(1-\dot{x})\left[(1-\dot{x})+2(2\dot{x}-1)\right.$$

$$= \dot{x}^2(1-\dot{x})(3\dot{x}-1)$$

THUS, WE WANT

$$\dot{x}^2(1-\dot{x})(3\dot{x}-1)\Big|_{t=t'} = \dot{x}^2(1-\dot{x})(3\dot{x}-1)\Big|_{t=t_+}$$

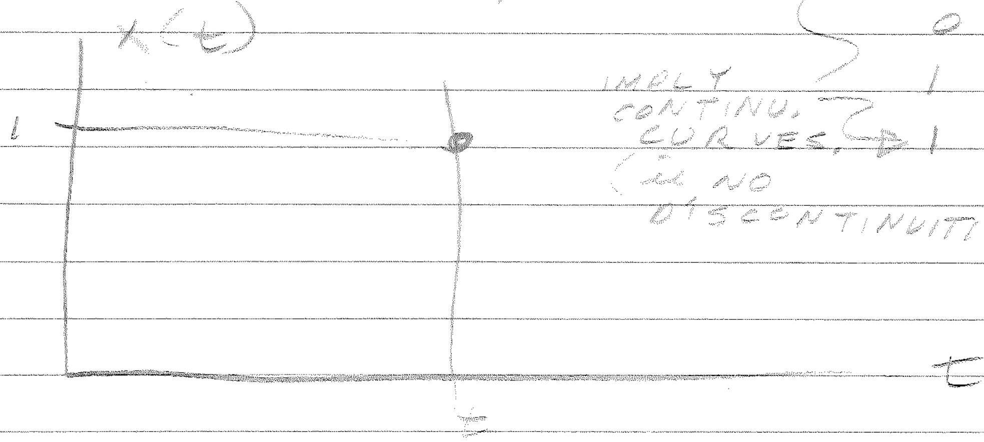
$$\Rightarrow \dot{x}(t') = 0, 1, \tfrac{1}{3}$$
$$\dot{x}(t_+) = 0, 1, \tfrac{1}{3}$$

COMPARING THIS WISH OTHER CORNER CONDITIONS GIVES

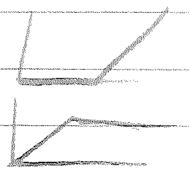
$x(t') = 0, 1$
$x(t_+) = 0, 1$

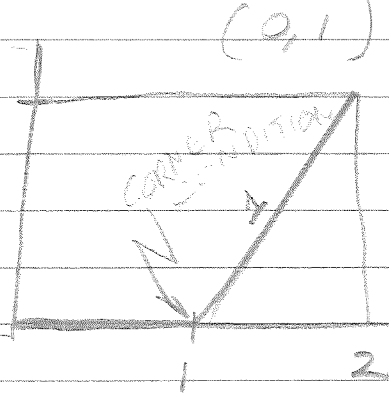
$\dot{x}(t')$	$\dot{x}(t_+)$
0	0
0	1
1	0
1	1

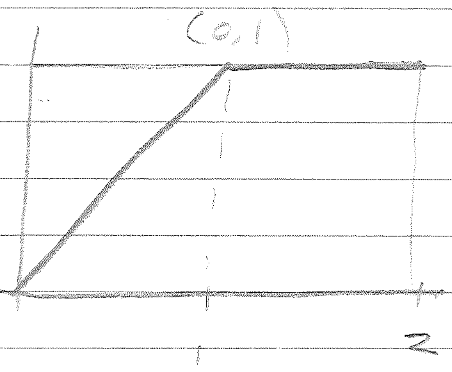
IMPLY CONTINU. CURVES (& NO DISCONTINUITIES)



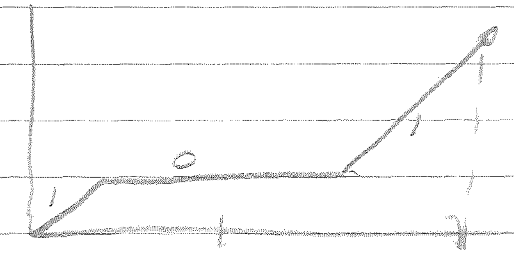
SO $\dot{x}(t')$ $\dot{x}(t_+)$ $\Rightarrow 0, 1$
$\Rightarrow 1, 0$

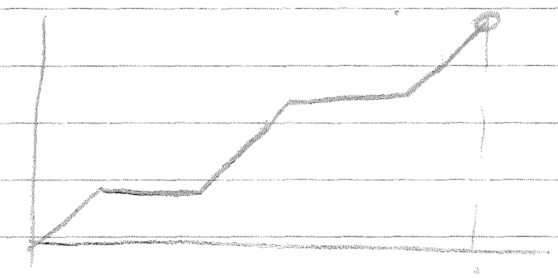






ANOTHER OPTIMAL IS





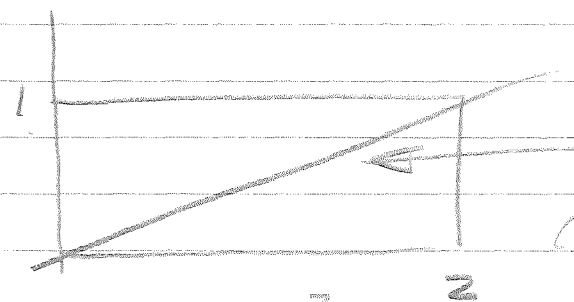
NOW

$$J = \int_0^2 \dot{x}^2 (1-\dot{x}^2)\, dt$$

OPTIMAL IS OBVIOUSLY @ J = 0

WHICH OCCURS @ $\dot{x} = 0$ & 1

IF WE WANNA SMOOTH CURVE OPTIMAL SOLUTION IS STRAIGHT LINE.



SLOPE = $\frac{1}{2}$

(SOL TO EULER EQ. WAS STRAIT LINE)

$$J_{min} = \int_0^2 \left(\frac{1}{2}\right)^2 \left(1-\frac{1}{2}\right)^2 dt = \frac{1}{8}$$

CONSTRAINED EXTREMA

TO OPTIMIZE J SUBJECT TO n NONLINEAR CONSTRAINTS

$$f_i(x) = 0 \quad , \quad i = 1, 2, \ldots, n \quad (*)$$

DEF: A POINT $\hat{x}$ SATISFYING ~~THESE~~ CONSTRAINTS (*) IS SAID TO BE A REGULAR POINT OF THESE CONSTRAINTS IF THE n LINEAR FUNCTIONALS $f_1'(\hat{x}), f_2'(\hat{x}), \ldots$ ARE LINEARLY INDEPENDENT. $\left(\frac{df}{dx}\right)$

RECALL LAGRANGE MULTIPLIERS

$$\phi + \lambda^T f$$

CONSIDER

$$J = \int_{t_0}^{t_f} g(w, \dot{w}, t)\, dt$$

$$w = \begin{bmatrix} x \\ u \end{bmatrix} \begin{matrix} n \\ m \end{matrix}$$

$$1$$

CONSTRAINTS: $f_i(w, t) = 0, \; i = 1, \ldots, n$

$$\left.\begin{matrix} f_1(x, u) = 0 \\ \vdots \\ f_2(x, u) = 0 \end{matrix}\right\} x = F(u, t)$$

PLUG INTO J + THIS BECOMES UNCONSTRAINED PROBLEM. FINDING F(u, t) IS USUALLY TO HARD.

BETTER TO USE LAGRANGE MULTIPLIER THEORY, USE

$$J_a = \int_{t_0}^{t_f} [g(w, \dot{w}, t + \lambda^T f] dt$$

10-11-76 (MON)

$$J = \int_{t_0}^{t_f} \phi(w, \dot{w}, t)\, dt$$

CONSTRAINT

$$f(w, t) = 0$$

$J_a = J$ AUGMENTED

$$= \int_{t_0}^{t_f} \left[\phi(w, \dot{w}, t)\, dt + \lambda^T f \right] dt$$

$$\delta J_a = \int_{t_0}^{t_f} \left[\left(\frac{\delta \phi}{\delta w}\right)^T \delta w + \left(\frac{\delta \phi}{\delta \dot{w}}\right)^T \delta \dot{w} + \lambda^T \left(\frac{\delta f}{\delta w}\right) \delta w + f^T (\delta \lambda) \right] dt$$

$$\boxed{\left[\frac{\delta \phi}{\delta w_1} \; \frac{\delta \phi}{\delta w_2} \cdots \right]} \begin{bmatrix} \delta w_1 \\ \delta w_2 \\ \vdots \end{bmatrix}$$

INTEGRATE SECOND TERM BY PARTS

$$\left(\frac{\delta \phi}{\delta \dot{w}}\right)^T \delta \dot{w} = \delta \dot{w}^T \left(\frac{\delta \phi}{\delta \dot{w}}\right)$$

$$U = \frac{d\phi}{\delta \dot{w}} \qquad \delta v = (\delta \dot{w})^T dt$$

$$dU = \frac{d}{dt}(\delta\phi / \delta \dot{w}) \qquad v = \delta w$$

$$\Rightarrow \delta J_a = \int_{t_0}^{t_f} \left[(\delta w)^T \left[\frac{\delta \phi}{\delta w}\right] + (\delta w)^T \left(\frac{\delta f}{\delta w}\right)^T \lambda + f^T \delta \lambda - \delta w^T \frac{d}{dt}\left[\frac{\delta \phi}{\delta \dot{w}}\right] \right] dt$$

$$+ (\delta w)^T \left.\frac{\delta \phi}{\delta \dot{w}}\right|_{t_0}^{t_f}$$

BUT, @ EXTREMUM $f = 0$

$$\Rightarrow f^T \delta \lambda = 0$$

$$\delta J_a = \int_{t_0}^{t_f} (\delta w)^T \Big[\frac{\delta \phi}{\delta w} - \frac{d}{dt}\Big(\frac{\delta \phi}{\delta \dot{w}}\Big) + \frac{\delta f^T}{\delta w}\Big) \lambda \Big] dt$$

$$+ (\delta w)^T \frac{\delta \phi}{\delta \dot{w}}\Big|_{t_0}^{t_f} = 0$$

NECESSARY CONDITIONS:

① $\frac{\delta \phi}{\delta w} - \frac{d}{dt}\Big(\frac{\delta \phi}{\delta \dot{w}}\Big) + \frac{\delta f^T}{\delta w} \lambda = 0$

② $(\delta w)^T \, \delta \phi / \delta \dot{w} \,\Big|_{t_0}^{t_f} = 0$

NOTE: THIS IS WHAT WE'DA GOT FOR

$$\phi_a = \phi(w, \dot{w}, t) + \lambda^T f(w, t)$$

$$\Rightarrow \left\{ \begin{array}{l} \frac{\delta \phi_a}{\delta w} - \frac{d}{dt}\Big(\frac{\delta \phi_a}{d \dot{w}}\Big) = 0 \\ (\delta w)^T \frac{\delta \phi_a}{\delta \dot{w}} \Big|_{t_0}^{t_f} = 0 \end{array} \right\} \text{POINT CONSTRAINT}$$

LOOK @ DIFFERENTIAL CONSTRAINT

$$f = f(w, \dot{w}, t) = 0$$

THEN SIMPLY

$$\phi_a = \phi + \lambda^T f(w, \dot{w}, t)$$

SAME EQUATIONS.

EXAMPLE (SIMPLE) ROCKET PROB

$$\ddot{\theta} = u(t)$$

$$J = \frac{1}{2}\int_0^2 (\ddot{\theta})^2 dt$$

$$\theta(0) = 1 \quad , \quad \theta(2) = 0$$

$$\dot{\theta}(0) = 1 \qquad \dot{\theta}(2) = 0$$

$$x_1 = \theta, \quad x_2 = \dot{\theta}$$

$$\dot{x}_1 = x_2 \qquad \dot{x}_2 = u$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

NOTE:

$$\dot{x}_1 - x_2 = 0$$

$$\dot{x}_2 - u = 0$$

$$J_a = \frac{1}{2}\int_0^2 \left[(\ddot{\theta})^2 + \lambda^T f \right] dt$$

$$= \int_0^2 \left(\frac{1}{2}(\dot{x}_2)^2 + \underline{\lambda_1 \lambda_2} \begin{bmatrix} \dot{x}_1 - x_2 \\ \dot{x}_2 - u \end{bmatrix} \right) dt$$

$$= \int_0^2 \Big[\frac{1}{2}(u)^2 + \lambda_1(\dot{x}_1 - x_2) + \lambda_2(\dot{x}_2 - u) \Big] dt$$

$$\phi_a = \frac{1}{2}u^2 + \lambda_1(\dot{x}_1 - x_2) + \lambda_2(\dot{x}_2 - u)$$

LET $w = \begin{bmatrix} x \\ u \end{bmatrix}$

$$\frac{\delta\phi_a}{\delta x_1} - \frac{\delta}{\delta t}\left(\frac{\delta\phi_a}{\delta \dot{x}_1}\right) = 0$$

$$\Rightarrow 0 - \frac{\delta}{\delta t}\lambda_1 = 0 \quad \Rightarrow \dot{\lambda}_1 = 0 \Rightarrow \lambda_1 = \text{CONST.}$$
$$= C_1$$

$$\frac{\delta\phi_a}{\delta x_2} - \frac{\delta}{\delta t}\left(\frac{\delta\phi_a}{\delta t}\right) = 0$$

$$\Rightarrow -\lambda_1 - \frac{d}{dt}\lambda_2 = 0 \Rightarrow \dot{\lambda}_2 = -\lambda_1 = -C_1$$
$$\Rightarrow \lambda_2 = -C_1 t + C_2$$

$$\frac{\delta\phi_a}{\delta u} - \frac{d}{dt}\frac{\delta\phi_a}{\delta \dot{u}} = 0$$
$$\Rightarrow = u - \lambda_2 = 0 \Rightarrow u = \lambda_2$$

THUS

$$\hat{x}_2 = u = -c_1 t + c_2$$
$$\Rightarrow x_2 = -\tfrac{1}{2}c_1 t^2 + c_2 t + c_3$$
$$x_2(0) = \dot{\theta}(0) = 1 \Rightarrow c_3 = 1$$
$$\Rightarrow x_2 = -\tfrac{1}{2}c_1 t^2 + c_1 t + 1$$
$$x_2(2) = -2c_1 + 2c_2 + 1 = 0 \qquad ①$$

$$\dot{x}_1 = x_2 = -\tfrac{1}{2}c_1 t^2 + c_2 t + 1$$
$$\Rightarrow x_1 = -\tfrac{1}{6}c_1 t^3 + \tfrac{1}{2}c_2 t^2 + t + c_4$$
$$x_1(0) = 1 = c_4$$
$$x_1(2) = 0 = -\tfrac{4}{3}c_1 + 2c_2 + 2 + 1 \qquad ②$$

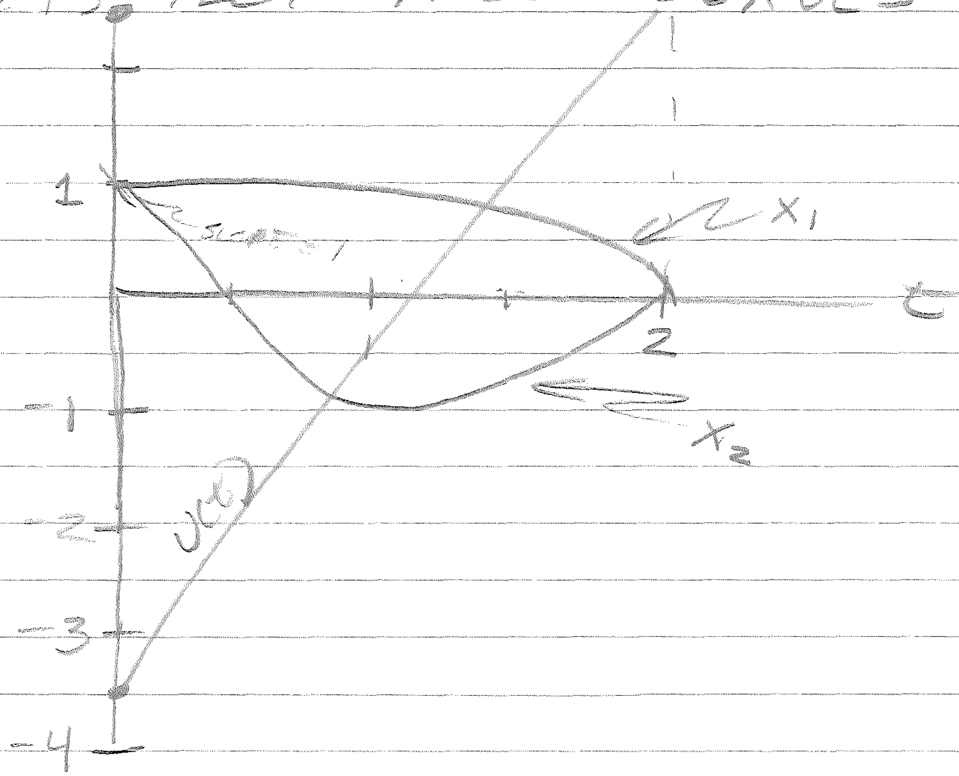
① & ② ARE 2 EQ/ 2 UNKNOWNS.
GIVES

$$c_1 = -3 \quad , \quad c_2 = -7/2$$

$$x_1(t) = \tfrac{1}{2}t^3 - \tfrac{7}{4}t^2 + t + 1$$
$$x_2 = \tfrac{3}{2}t^2 - \tfrac{7}{2}t + 1$$
$$u = \dot{x}_2 = 3t - 7/2$$

} OPTIMAL SOLUTION

LETS PLOT THESE CURVES



2/13/76 [illegible]

$J = \int_{t_0}^{t_f} [illegible] \, dt$; $\psi(x_f, t_f) = 0$

$\phi_a = \phi + \lambda' \psi$

$\left(\frac{\delta \phi_a}{\delta x} - \frac{d}{dt} \frac{\delta \phi_a}{\delta \dot{x}}\right) = 0$

$\delta x' \frac{\delta \phi_a}{\delta \dot{x}} \Big|_{t_0}^{t_f}$

Ex.

$J = \frac{1}{2} \int_{t_0}^{t_f} \|u(t)\|^2_{R(t)} + \|x(t) - r(t)\|^2_{Q(t)} \, dt$

SUBJECT TO $\dot{x} = Ax + Bu$, $x(t_0) = x_0$

R [illegible] POS. [illegible] DEF. [illegible]

Q(t) [illegible] POS. [illegible] SEMI [illegible]

$\phi_a = \frac{1}{2} \left(\|u\|^2_R + \frac{1}{2} \|x - r\|^2_Q + \lambda^T (Ax + Bu - \dot{x}) \right)$

① $\frac{\delta \phi_a}{\delta x} - \frac{d}{dt} \left(\frac{\delta \phi_a}{\delta \dot{x}} \right) = 0$

$\left[\frac{1}{2} Q(x - r) + \frac{1}{2} Q^T (x - r) + A^T \lambda \right] - \frac{d}{dt} \left(\frac{\delta \phi_a}{\delta \dot{x}} \right) = 0$

$\left[Q(x - r) + A^T \lambda \right] + \frac{d\lambda}{dt} = 0$

② $\frac{\delta \phi_a}{\delta u} - \frac{d}{dt} \left(\frac{\delta \phi_a}{\delta \dot{u}} \right) = 0$

$Ru + B^T \lambda - 0 = 0 \Rightarrow u = -R^{-1} B^T \lambda$

$\frac{d\lambda}{dt} = -A^T \lambda - Q(x - r)$

$\dot{x} = Ax + Bu$

[illegible]

[illegible] COND. ON E

$\delta x' \; \delta u' \left. \begin{bmatrix} \frac{\delta \phi_a}{\delta \dot{x}} \\ \frac{\delta \phi_a}{\delta \dot{u}} \end{bmatrix} \right|_{t_0}^{t_f} = 0$; $\delta x^T(t_0) = $

@ $t = t_f$, $\delta x'(t_f) \neq 0$

THUS

$\left. \frac{\delta \phi_a}{\delta \dot{x}} \right|_{t_f} = -\lambda(t) \Big|_{t_f} = -\lambda(t_f) = 0$

THUS, [illegible] DIFF. EQ. ARE

$\dot{\lambda} = -A^T \lambda - Q(x - r)$; $\lambda(t_f) = 0$

$\dot{x} = Ax - BR^{-1}B^T \lambda$; $x(t_0) = x_0$

INEQUALITY CONSTRAINT

EXTREMIZE J SUBJECT TO

$f(u,\dot{u},t)=0$

AND $\Gamma_{MIN} \le \Gamma(u,\dot{u},t) \le \Gamma_{MAX}$

LET $(\Gamma_{MAX_i} - \Gamma_i)(\Gamma_i - \Gamma_{MIN_i}) = \gamma_i^2$

FOR $\gamma_i^2 \ge 0$, THE INEQUALITY IS SATISFIED. USE LAGRANGE MULTIPLIER

ISOPERIMETRIC CONSTRAINTS

EXTREMIZE $J = \int_{t_0}^{t_f} \phi(u,\dot{u},t)\,dt$

SUBJECT TO

$\int_{t_0}^{t_f} e(u,\dot{u},t)\,dt = c$ [illegible]

$z(t) \triangleq \int_{t_0}^{t} e(u,\dot{u},t)\,dt \Rightarrow z(t_0)=0$

$z(t_f)=c$

$\Rightarrow \dot{z}(t) = e \; ; \; z(t_0)=0$

$z(t_f)=c$

THEN $\phi_0 = \phi + \lambda^T(\dot{z}-e)$

$\frac{\partial\phi_0}{\partial u} - \frac{d}{dt}\frac{\partial\phi_0}{\partial\dot{u}} = 0 \Rightarrow u \boxed{\begin{matrix} n \\ m \end{matrix}}$

$\Rightarrow n+m$ EQUATIONS

ALSO $\frac{\partial\phi_0}{\partial z} - \frac{d}{dt}\left(\frac{\partial\phi_0}{\partial\dot{z}}\right)=0 \Rightarrow r$ EQUA.

UNKNOWNS $\overset{n}{x}, \overset{m}{u}, \overset{r}{z}, \overset{r}{\lambda} \Rightarrow 2r+m+n$ UNKNOWNS

OTHER r EQUATIONS FROM

$\dot{z}(t) = e, \; z(t_0)=0, \; z(t_f)=c$

$\frac{\partial}{\partial z}\phi_0 = 0 \Rightarrow \frac{d}{dt}\left(\frac{\partial\phi_0}{\partial\dot{z}}\right)=0$

AND $\frac{d}{dt}\lambda = 0 \Rightarrow \lambda =$ CONST.

EXAMPLE

EXTREMIZE

$J = \int_{t_0}^{t_f} \frac{1}{2}(u_1^2 + u_2^2 + 2\dot{u}_1 \dot{u}_2)\,dt$

SUBJECT TO $\int_{t_0}^{t_f} u_2^2\,dt = C$

LET $z = \int_{t_0}^{t} u_2^2\,dt$

$\Rightarrow \dot{z} = u_2^2$; $z(t_0) = 0$; $z(t_f) = C$

$F_a = \frac{1}{2}(u_1^2 + u_2^2 + 2\dot{u}_1 \dot{u}_2) + \lambda(\dot{z} - u_2^2)$

EULER-LAGRANGE

$\frac{\partial F_a}{\partial u_1} - \frac{d}{dt}\left(\frac{\partial F_a}{\partial \dot{u}_1}\right) = 0$ ①

$u_1 - \frac{d}{dt}\dot{u}_2 = 0$

$\frac{\partial F_a}{\partial u_2} - \frac{d}{dt}\frac{\partial F_a}{\partial \dot{u}_2} = 0$

$u_2 - 2u_2\lambda) - \frac{d}{dt}\dot{u}_1 = 0$ ②

λ = CONST ③

10-14-76 (FRI)

CONTINUOUS OPTIMAL CONTROL

BOLZA PROBLEM

$\dot{x} = f(x, u, t)$; n VECTOR

$M(t_0)\,x(t_0) = m_0$; r VECTOR

$N(t_f)\,X(t_f) = n_{t_f}$; q VECTOR

MINIMIZE

$$J = \Theta(x,t)\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \phi(x,u,t)\,dt$$

USE LAGRANGE MULTIPLIER

$$J_a = \Theta\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f}\Big[\underbrace{\phi + \lambda^T[f(x,u,t)}_{\mathcal{H} = \phi + \lambda^T f} - \dot{x}]\Big]dt$$

$$J_a = \Theta\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f}[\mathcal{H} - \lambda^T \dot{x}]\,dt$$

INTEGRATE BY PARTS

$$J_a = \Theta\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f}\mathcal{H}\,dt - \Big(\lambda^T X\Big|_{t_0}^{t_f}$$

$$- \int_{t_0}^{t_f}\dot{\lambda}^T x\,dt\Big)$$

$$= (\Theta - \lambda^T x)\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f}(\mathcal{H} + \acute{\lambda}^T x)\,dt$$

$$\delta J_a = (\delta x)^T\left(\frac{\delta\Theta}{\delta x}\ \lambda\right)\Big|_{t_0}^{t_f}$$

$$+ \int_{t_0}^{t_f}(\delta X)^T\frac{\delta H}{\delta X} + (\delta x)^T\acute{\lambda}$$

$$+ (\delta u)^T\frac{\delta\mathcal{H}}{\delta u}\Big]\,dt$$

$$\delta J_a = (\delta x)^T \left[\frac{\delta \theta}{\delta x} - \lambda \right]_{t_0}^{t_f}$$

$$+ \int_{t_0}^{t_f} (\delta x)^T \left(\frac{\delta H}{\delta x} + \dot{\lambda} \right) dt$$

$$+ \int_{t_0}^{t_f} (\delta u)^T \frac{\delta H}{\delta u} dt = 0$$

OUR NECESSARY CONDITIONS ARE:

$$(\delta x)^T \left(\frac{\delta \theta}{\delta x} - \lambda \right) = 0$$

$$\delta H / \delta x + \dot{\lambda} = 0$$

$$\delta H / \delta u = 0$$

$$\dot{x} = f(x, u, t) \longleftarrow \text{STATE EQ}$$

$$\dot{\lambda} = - \delta H / \delta x \longleftarrow \text{COSTATE}$$

x = STATE

$$\delta H / \delta u = 0$$

$$H = \phi + \lambda^T f \longleftarrow$$

$$0 = m(t_0) \implies m(t_0)\, x(t_0) = 0$$

$$0 = n(t_0) \implies N(t_f)\, x(t_f) = 0$$

$$J_a = \theta \Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \Big[\phi + \lambda^T (f - \dot{x}) dt$$

$$+ \nu^T n(x, t) \Big] dt$$

$$- \xi^T m(x, t) \Big]_{t_0}$$

PLOWGAN AWAY... GIVE...

$$\delta J_a = (\delta x)^T \left(\frac{\delta \theta}{\delta x} - \lambda \right) \Big|_{t_0}^{t_f}$$

$$+ \int_{t_0}^{t_f} \Big[(\delta x)^T \frac{\delta H}{\delta x} + \delta \dot{x}^T \dot{\lambda}$$

$$+ \delta u^T \, \delta H / \delta \lambda \Big] dt$$

$$- (\delta x)^T \frac{\delta m^T}{\delta x} \xi \Big|_{t_0} + \delta x^T \left(\frac{\delta n}{\delta x} \right)^T \nu \Big|_{t_f}$$

$$\delta J_a = (\delta x)^T \left[\frac{\delta \Theta}{\delta x} + \left(\frac{\delta \eta}{\delta x}\right)^T \nu \right]$$

$$+ \delta x^T \left(\frac{\delta \Theta}{\delta x} - \lambda - \left(\frac{\delta m}{\delta x} \right)^T \xi \right) \Big|_{t_0}$$

$$+ \int_{t_0}^{t_f} (-)$$

OUR NECESSARY CONDITIONS ARE THEN

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

$$\dot{\lambda} = -\delta \mathcal{H} / \delta x$$

$$\delta H / \delta u = 0$$

$$\delta \Theta / \delta x = \lambda - \left(\frac{\delta \eta}{\delta x} \right)^T \nu \Big|_{t_f}$$

$$\delta \Theta / \delta x = \lambda + \left(\frac{\delta m}{\delta x} \right)^T \xi \Big|_{t = t_0}$$

NOW

$$\mathcal{H} = \phi + \lambda^T f = \phi(x, u, t) + \lambda^T f(x, u, t)$$

$$\frac{\delta \mathcal{H}}{\delta t} = \frac{\delta \phi}{\delta t} + \left(\frac{\delta x}{\delta \phi} \right)^T \frac{d\phi}{dx}$$

$$+ \left(\frac{\delta u}{dt} \right)^T \left(\frac{d\phi}{du} \right]$$

$$+ \frac{\delta f^T}{\delta t} + \left(\frac{dx}{dt} \right)^T \frac{d}{dt} f^T + \left(\frac{du}{dt} \right)^T \frac{\delta f^T}{\delta u} \Big] u$$

$$+ \frac{\delta \lambda^T}{\delta t} f$$

SIMPLIFY:

$$\frac{dH}{dt} = \frac{\delta \phi}{\delta x} + \dot{x}^T \left[\frac{\delta \phi}{\delta x} - \frac{\delta f^T}{\delta x} \lambda \right]$$

$$+ \dot{u}^T \left[\frac{\delta \phi}{\delta t} + \frac{\delta f^T}{\delta u} \lambda \right)$$

$$+ \frac{\delta f^T}{\delta t} \lambda + \frac{\delta \lambda'}{\delta t} \lambda$$

LOOK @ SECOND & THIRD:

$$\lambda' = - \frac{\delta H}{\delta t} = - \frac{\delta}{\delta x} (\phi + \lambda^T f)$$

$$= - \delta\phi / \delta x - df^T / dx$$

$$\frac{\delta H}{\delta u} = \frac{\delta}{\delta u} = \frac{\delta}{\delta \phi} (\phi + \lambda^T f) = \frac{d\phi}{du} + \frac{\delta f^T}{\delta u} \lambda \quad = 0$$

$$\frac{dH}{dt} = \frac{\delta \phi}{\delta t} - \dot{x}^T \dot{\lambda} + \dot{u} \cancel{\frac{\delta H}{\delta u}} \rightarrow 0 + \frac{\delta f^T}{\delta t} \lambda + \frac{\delta \lambda'}{\delta t} \lambda$$

$$+ \dot{\lambda}^T f \quad \Rightarrow \text{CANCEL}$$

$$= \frac{\delta \phi}{\delta t} + \frac{\delta f^T}{\delta t} \lambda$$

= 0 IF ϕ & f ARE NOT EXPLICIT FUNCTIONS OF t

$\dot{x} = f(x, u) \Rightarrow$ AUTOTOMOUS OR TIME VARIANT

10-18-76 (MON)

t_0 & t_f FIXED

$$\begin{cases} \dot{x} = f(x, u, t) \\ J = \Theta(x, t)\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \phi(x, u, t)\, dt \\ m\left[x(t_0), t_0\right] = 0 \\ n\left[x(t_f), t_f\right] = 0 \end{cases}$$

$$\delta x'\left(\frac{\delta \Theta}{\delta x} - \lambda\right)\Big|_{t_0}^{t_f} = 0$$

$$\begin{cases} \dot{\lambda} = -\delta H / \delta x \Leftarrow \text{COSTATE EQ.} \quad H = \phi + \lambda^T f \\ \dot{x} = f(x, u, t) = \delta H / \delta \lambda \end{cases}$$

TRANS VERSALITY COND

$$\begin{cases} \frac{\delta H}{\delta u} = 0 \\ \lambda(t_f) = \frac{\delta \Theta}{\delta x} + \left(\frac{\delta n}{\delta x}\right)^T \nu \\ \lambda(t_0) = \delta \Theta / \delta x + \left(\delta m / \delta x\right)^T \xi \end{cases}$$

Ex

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 \quad x_1(0) = 0 \quad f(x, u, t) \\ \dot{x}_3 = u \end{cases}$$

WE WISH TO DERIVE THE SYSTEM $\Rightarrow$ WE REACH THE TERMINAL MANIFOLD

$$x_1^2(1) + x_2^2(1) = 0$$

$(t_f = 1 \; ; \; n(x(t_f), t_f) = 0$

$$\Rightarrow J = \frac{1}{2}\int_0^1 u^2 dt \quad \text{IS MINIMIZED}$$

FIRST, FIND $\mathcal{H}$

$$H = \phi + \lambda^T f = \frac{1}{2}U^2 + \begin{bmatrix} \lambda_1 & \lambda_2 & \lambda_3 \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \\ U \end{bmatrix}$$

$$= \frac{1}{2}U^2 + \lambda_1 x_2 + \lambda_2 x_3 + \lambda_3 U$$

NOW

$$\frac{\delta H}{\delta U} = 0 = U + \lambda_3 \Rightarrow \lambda_3 = -U$$

LOOK @ COSTATE EQUATION

$$\dot{\lambda} = -\frac{\delta H}{\delta x} \Rightarrow \left\{ \begin{bmatrix} \dot{\lambda}_1 \\ \dot{\lambda}_2 \\ \dot{\lambda}_3 \end{bmatrix} = -\frac{\delta H}{\delta x} = \begin{bmatrix} -\delta H/\delta x_1 \\ -\delta H/\delta x_2 \\ -\delta H/\delta x_3 \end{bmatrix} \right.$$

$$= \begin{bmatrix} 0 \\ -\lambda_1 \\ -\lambda_2 \end{bmatrix}$$

OR

$$\left. \begin{array}{l} \dot{\lambda}_1 = 0 \\ \dot{\lambda}_2 = -\lambda_1 \\ \dot{\lambda}_3 = -\lambda_2 \end{array} \right\} \text{COSTATE EQUATIONS}$$

FROM BH

$$\left. \begin{array}{l} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 \\ \dot{x}_3 = U \end{array} \right\} \text{STATE EQ.}$$

$$x_i(0) = 0 \quad i = 1, 2, 3$$

NOW, SINCE $\theta = 0$:

$$\lambda(t_f) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 2x_1(t_f) \\ 2x_2(t_f) \\ 0 \end{bmatrix} \checkmark$$

WE ARE GIVEN

$$\Rightarrow \lambda_1(1) = 2x_1(1) \qquad x_1(0) = 0$$

$$\lambda_2(1) = 2x_2(1) \qquad x_2(0) = 0$$

$$\lambda_3(1) = 0 \qquad x_3(0) = 0$$

ALSO $x_1^2(1) + x_2^2(1) = 1$

SPLIT B.C. ROUGH TO SOLVE.
WE COULD IN PRINCIPLE
HOWEVER, DO IT.

CONSIDER THE PROBLEM:

t_f FREE

$$\dot{x} = f(x,u,t) \quad ; \quad x(t_0) = x_0$$

$$J = \Theta(x,t)\Big|_{t_f} + \int_{t_0}^{t_f} \phi(x,u,t)\,dt$$

$$J_a = \Theta(x,t)\Big|_{t_f} + \nu^T N[x(t_f, t_f)]$$
$$+ \int_{t_0}^{t} [\phi + \lambda^T(f - \dot{x})]\,dt$$

LET $\mathcal{H} = \phi + \lambda^T f$

$$J_a = \Theta\Big|_{t_f} + \nu^T n\Big|_{t_f} + \int_{t_0}^{t_f} \mathcal{H}\,dt$$
$$- \underbrace{\left(\lambda^T x\Big|_{t_0}^{t_f} - \int_{t_0}^{t_f} \dot{\lambda}^T x\,dt\right)}_{\text{FROM PARTS INTEGRATION}}$$

$$J_a = \Theta\Big|_{t_f} + \nu^T n\Big|_{t_f}$$
$$+ \int_{t_0}^{t_f} [\mathcal{H} + \dot{\lambda}^T x)\,dt$$
$$- \lambda^T x\Big|_{t_0}^{t_f}$$

LET

$$\bigodot = \Theta + \nu^T n$$

$$J_a = \Theta\Big|_{t_f} - \lambda^T x\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f}\left[\mathcal{H} + \dot{\lambda}^T(x)\right]dt$$

$$\delta J_a = \left[(\delta x^T)\frac{\delta\Theta}{\delta x}\Big|_{t_f} + (\delta t_f)\frac{\delta\Theta}{\delta t_f}\right.$$

$$\left. + \delta t_f\left(\frac{d\Theta}{dx}\right)^T\frac{dx}{dt}\right]$$

$$- \left[\delta x^T\frac{\delta(\lambda^T x)}{\delta x} + \delta t\left(\frac{\delta(\lambda^T x)}{\delta\lambda}\right)^T\frac{\delta\lambda}{\delta t}\right.$$

$$\left. + \delta t\left(\frac{\delta\lambda^T x}{\delta x}\right)^T\frac{dx}{dt}\right]\Big|_{t_0}^{t_f}$$

$$+ \int_{t_0}^{t_f}\left[(\delta x)^T\left\{\frac{\delta H + \dot{\lambda}^T x}{\delta x}\right\}\right.$$

$$\left. + (\delta u)^T\frac{\delta H}{\delta u}\right]dt$$

$$+ \left(\mathcal{H} + \dot{\lambda}^T x\right)\Big|_{t_f}\delta t_f$$

IDENTICAL FROM Θ ON NEXT PG

$$= \delta t_f\left[\frac{\delta\Theta}{\delta t_f} + \left(\frac{\delta\Theta}{\delta x}\right)^T\dot{x} - \lambda^T\cancel{\dot{X}} = x\dot{\lambda}^T\right.$$

$$\left. + \left(H + \dot{\lambda}^T x\right)\right]\Big|_{t_f}$$

$$+ \delta X^T(t_f)\left[\frac{\delta\Theta}{\delta x} - \lambda\right]\Big|_{t_f}$$

$$+ \int_{t_0}^{t_f}\left[\delta X^T\left(\frac{\delta H}{\delta x} + \dot{\lambda}\right) + \delta U^T\left(\frac{\delta H}{\delta u}\right)\right]dt$$

$$= 0$$

CONDITIONS FROM THIS, ARE

① $\frac{\delta \mathcal{H}}{\delta u} = 0$

COSTATE

② $\dot{\lambda} = -\delta \mathcal{H} / \delta x \; ; \; n(t_f, x(t_f)) = 0$

③ $\lambda = \frac{\delta \Theta}{\delta x} \Big|_{t = t_f}$

④ $\frac{\delta \Theta}{\delta t_f} + \mathcal{H} = 0 \quad @ \; t = t_f$

$$\Rightarrow \frac{\delta \Theta}{\delta t_f} + \frac{\delta n^T}{\delta t} \nu + H = 0 \; @ \; t = t_f$$

SUBJECT TO

⑤ $\dot{x} = f(x, u, t) \quad , x(t_0) = x_0$

$= \frac{\delta \mathcal{H}}{\delta \lambda}$

10-20-76 (WED)

THE LINEAR REGULATOR PROBLEM

$$\dot{x} = Ax + BU$$ FINITE CONTROL ENERGY

$$J = \frac{1}{2} x^T(t_f) S x(t_f) + \frac{1}{2}\int_{t}^{t_f}(x^T Q x + U^T R U)\,dt$$

$(S \,\&\, Q \geq 0$, $R > 0$, & ALL SYMMETRIC

- BEFORE, WE LOOKED AT

$$J = \Theta(x,t)\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \phi(x,u,t)\,dt$$

HERE, WE KNOW $x(t_0)$ → FIXED

$$\Rightarrow J = \Theta[x'(t_f), t_f] + \cancel{\Theta[x(t_0), t_0]} + \int \phi\,dt$$

NOW

$$\delta J = \delta x'\overset{③}{\left(\frac{\delta\Theta}{\delta x} - \lambda\right)}\Big|_{t_0}^{t_f}$$

$$+ \int_{t_0}^{t_f}\left[\delta x^T \overset{②}{\left(\frac{\delta H}{\delta x} + \dot{\lambda}\right)} + \delta U^T \overset{①}{\left(\frac{\delta H}{\delta U}\right)}\right]dt$$

- FOR OUR PROBLEM

$$H = x^T Q x + U^T R U + \lambda^T(Ax + BU)$$

A NEC. CONDITION IS

① $\delta H / \delta U = 0$

$$= RU + B^T\lambda \Rightarrow U = -R^{-1}B^T\lambda$$

② SECOND

$$\dot{\lambda} = -\frac{\delta H}{\delta x} \Rightarrow \dot{\lambda} = -(Qx + A^T\lambda)$$

③ AGAIN, WE'RE GIVEN t_0, THUS

$$\delta x'\left(\frac{\delta\Theta}{\delta x} - \lambda\right)\Big|_{t_f} = 0$$

$$\Rightarrow \frac{\delta\Theta}{\delta x} - \lambda\Big|_{t_f} = 0$$

NOW, HERE $\Theta = \frac{1}{2} x^T S x$

$$\Rightarrow @\ t = t_f$$

$$\lambda = \delta\Theta / \delta x = Sx \Rightarrow \lambda(t_f) = Sx(t_f)$$

OUR EQUATIONS TO DATE ARE

$$\dot{x} = Ax - BR^{-1}B^T\lambda \quad ; \quad x(t_0) = x_0$$

$$\dot{\lambda} = -(Qx + A^T\lambda) \quad ; \quad \lambda(t_f) = Sx(t_f)$$

WE, IN PRINCIPLE CAN SOLVE. NOTE THAT THIS SCHEME IS OPEN LOOP.

DOES THERE EXIST A CORRESPONDING FEEDBACK?

ASSUME:

$\lambda(t) = P(t)x(t)$,

$\Rightarrow u(t) = -R^{-1}B^TP(t)x(t) = K(t)x(t)$

[K WILL BE THE "KALMAN" GAIN]

NOW $\lambda(t_f) = P(t_f)x(t_f) = Sx(t_f)$

$\Longrightarrow S = P(t_f)$

$\dot{\lambda}(t) = P(t)\dot{x} + \dot{P}(t)\cancel{x} = -Qx - A^T\lambda$

$= -Qx - A^TP(t)x(t)$

EQUATING

$P\dot{x} + \dot{P}x = -Qx - A^TPx$

$[\dot{P} + Q + A^TP]x + P\dot{x} = 0$

$[\dot{P} + Q + A^TP]x + P[Ax - BR^{-1}B^TRx] = 0$

$[\dot{P} + Q + A^TP + PA - PBR^{-1}B^TP)x = 0$

THUS

$\dot{P} + Q + A^TP + PA - PBR^{-1}B^TP = 0$

P IS OUR UNKNOWN:

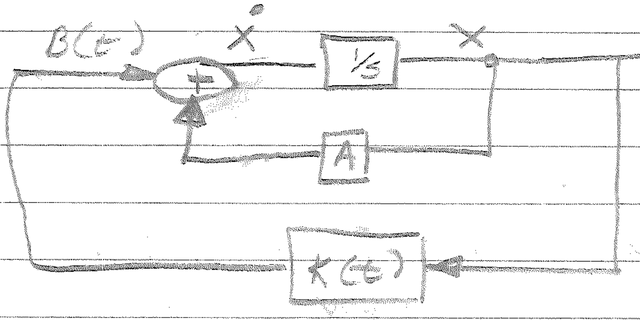
$P = -PA - A^TP + PBR^{-1}B^TP - Q = 0$

RICATTI EQUATION (NONLINEAR)

$P(t_f) = S$

NOTE: WE KNOW BOUNDRY COND. AT $t = t_f$ THUS WE GOTTA "INTEGRATE" THIS THING BACKWARDS)

IN ORDER TO SOLVE REGULATOR PROBLEM WE GOTTA SOLVE RICCATTI EQUATION.



WE GOTTA SOLVE RICATTI SO WE CAN DETERMINE KALMAN GAIN, K(t).

ON RICATTI EQUATION

IF S IS VERY LARGE:

$$P(t)P^{-1}(t)=0$$

$$\dot{P}P^{-1}+P\dot{P}^{-1}=0 \Rightarrow \dot{P}P^{-1}=-P\dot{P}^{-1}$$

$$\dot{P}^{-1}=-P^{-1}\dot{P}P^{-1}$$

THUS, WE MAY WRITE

$$-\dot{P}^{-1}=P^{-1}\dot{P}P^{-1}=P^{-1}PAP^{-1}-P^{-1}A^{T}PP^{-1}+P^{-1}PBR^{-1}B^{T}PP^{-1}$$
$$-P^{-1}QP^{-1}$$

$$\left.=-AP^{-1}-P^{-1}A^{T}+R^{-1}B^{T}-P^{-1}QP^{-1}\right\}\text{GOTTA SOLVE}$$

$$\Rightarrow P^{-1}(t_f)=S^{-1}\Leftarrow$$

PROBLEM

$\dot{x} = -\frac{1}{2} x + u(t)$

MINIMIZE

$J = \frac{1}{2} S x^2(t_f) + \frac{1}{2}\int_{t_0}^{t_f} \left[2x^2 + u^2 \right] dt$

FOR THIS PROBLEM, WE HAVE

A RICCATI EQATION, IN GENERAL

$\dot{P} = -PA - A^TP - PBR^{-1}B^TP - Q$

$= \frac{1}{2}P + \frac{1}{2}P + P(1)(1)^{-1}(1)P - 2$

$= P + P^2 - 2$

NOW, $P(t_f) = S$

SOLUTION (FROM SAGE'S BOOK, IS

$P(t) = -0.5 + 1.5 \tanh\left[-1.5t + \xi_1 \right]$

$P(t) = -0.5 + 1.5 \coth\left[-1.5t + \xi_2 \right]$

$\ni$ ξ_1 & ξ_2 ARE DETERMINED FROM S

CONSIDER

FIRST ORDER RICATTI EQ. SOLUTION (METHODS OF SOLVING)

$* \quad \frac{dy}{dx} + a_2(x)y^2 + a_1(x)y + a_0(x) = 0$

LET $y_1(x)$ BE A PARTICULAR SOLUTION

MAKE A VARIABLE $y = y_1 + \frac{1}{z}$

TO REDUCE (*) TO A FIRST ORDER LINEAR DIFF. EQ. IN z.

$\dot{p} = p^2 + p - 2 \Rightarrow$ PARTICULAR SOLUTION @ $p = 1$

LET, THEN

$p = 1 + 1/z$

$\dot{p} = -1/z^2$

PLUG IN ABOVE

$-\dot{z}/z^2 = \frac{1}{z} + \left(1 + \frac{2}{z} + \frac{1}{z^2}\right) - 1$

$\Rightarrow -\dot{z} = z + \cancel{z^2} + 2z + 1 - \cancel{z^2}$

$\therefore \dot{z} + 3z + 1$

GENERAL SOLUTION IS

$z = ke^{-3t} - \frac{1}{3}$

$p = 1 + \frac{1}{z} = 1 + \frac{1}{ke^{-3t} - 1/3}$

EVALUATE K BY $p(t_f) = 5$

11-3-76 (WED)

THE MAXIMUM PRINCIPLE WITH STATE & CONTROL VARIABLE INEQUALITY CONTRAINTS

MIN: $J = \Theta[x(t_f), t_f] + \int_{t_0}^{t_f} \phi(x, u, t)\,dt$

FOR $\dot{x} = f(x, u, t)$

$\ni \quad x(t_0) = x_0$

$N[x(t_f), t_f] = 0$

$g(u, t) \geq 0 \quad \Longrightarrow (\dot{z})^2 = g(u, t)$

$h(x, t) \geq 0 \quad \Longleftarrow$ OUR NEW EQUATION

LET $\dot{x}_{n+1} = f_{n+1}$

$$\triangleq h_1^2(x, t)\,\mathcal{H}(h_1) + h_2^2(x, t)\,\mathcal{H}(h_2) + \dots + h_s(x, t)\,\mathcal{H}[h_s]$$

$\ni \mathcal{H}(h_s(x, t))$ = MODIFIED HEAVISIDE STEP FUNCTION:

$$\mathcal{H}[h_s(x, t)] = \begin{cases} 0 & ; h_s(x, t) \geq 0 \\ K_s > 0 & ; h_s < 0 \end{cases}$$

IF $h(x, t) \geq 0$, THEN

$$h_1^2(x, t)\,\mathcal{H}(h_1) + \dots + h_s(x, t)\,\mathcal{H}(h_s) = 0$$

ASSUME $x_{n+1}(t_0) = 0$

& $\dot{x}_{n+1}(t) \geq 0 \quad \forall t$

$\dot{x}_{n+1} = 0$ ONLY WHEN

$h(x, t) \geq 0$

10-25-76 (MON)

TRACKING PROBLEM:

$$\dot{x} = Ax + BU \quad \Leftarrow \text{STATE}$$

$$J = \tfrac{1}{2}\|x(t_f) - r(t_f)\|_S$$

$$+ \tfrac{1}{2}\int_{t_0}^{t_f}\left[\|x(t) - r(t)\|_Q^2 + \|u(t)\|_R^2\right]dt$$

(IF $r=0$, WE HAVE REGULATER PROB.)

$$\mathcal{H} = \tfrac{1}{2}\left[\|x-r\|_Q^2 + \|U\|_R^2\right] + \lambda^T(Ax+BU)$$

$(x-r)^TQ(x-r)$ $$\dot{\lambda} = -\frac{\delta H}{\delta x} = -\left[Q(x-r) + A^T\lambda\right] \Leftarrow \text{COSTATE}$$

$$\delta H/\delta U = 0 \Rightarrow RU + B^T\lambda = 0 \Rightarrow U = R^{-1}B^T\lambda$$

THUS

$$\dot{x} = Ax - BR^{-1}B^T\lambda$$

$$\dot{\lambda} = -Q(x-r) - A^T\lambda$$

IN MATRIX FORM:

$$\begin{bmatrix}\dot{x}\\ \dot{\lambda}\end{bmatrix} = \begin{bmatrix}A & -BR^{-1}B\\ -Q & -A^T\end{bmatrix}\begin{bmatrix}x\\ \lambda\end{bmatrix} + \begin{bmatrix}0\\ Qr\end{bmatrix}$$

LET'S SOLVE:

$$\begin{bmatrix}x(t)\\ \lambda(t)\end{bmatrix} = \Phi(t,t_0)\begin{bmatrix}x(t_0)\\ \lambda(t_0)\end{bmatrix}$$

$$+ \int_{t_0}^{t}\Phi(t,\tau)\begin{bmatrix}0\\ Q(\tau)r(\tau)\end{bmatrix}d\tau$$

APPLY BOUNDRY CONDITIONS:

$$\begin{bmatrix} x(t_f) \\ \lambda(t_f) \end{bmatrix} = \Phi(t_f, t_0) \begin{bmatrix} x(t_0) \\ \lambda(t_0) \end{bmatrix} + \int_{t_0}^{t} \Phi(t_f, \tau) \begin{bmatrix} 0 \\ Q(\tau) R(\tau) \end{bmatrix} d\tau$$

$$= \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} x(t_0) \\ \lambda(t_0) \end{bmatrix} + \int_{t_0}^{t} \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} 0 \\ Q(\tau) r(\tau) \end{bmatrix} d\tau$$

$$= \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} x(t_0) \\ \lambda(t_0) \end{bmatrix} + \begin{bmatrix} f_1(t_f) \\ f_2(t_f) \end{bmatrix}$$

NOW WE KNOW

$$\lambda(t_f) = \delta \Theta(x(t_f), t_f) / \delta x(t_f)$$

$$\Theta = \tfrac{1}{2} \| x(t_f) - r(t_f) \|_S^2$$

$$\Rightarrow \lambda(t_f) = S [x(t_f) - r(t_f)]$$

ALSO

$$\lambda(t_f) = \phi_{21} x(t_0) + \phi_{22} \lambda(t_0) + f_2(t_f)$$

SUBSTITUTING:

$$S x(t_f) - S r(t_f) = \phi_{21} x(t_0) + \phi_{22} \lambda(t_0) + f_2(t_f)$$

BUT

$$x(t_f) = \phi_{11} x(t_0) + \phi_{12} \lambda(t_0) + f_1(t_0)$$

$$\Rightarrow S[\phi_{11} x(t_0) + \phi_{12} \lambda(t_0) + f_1] - S r(t_f) = \phi_{21} x(t_0) + \phi_{22} \lambda(t_0) + f_2$$

NOW, IN GENERAL

$$\begin{bmatrix} x(t_f) \\ \lambda(t_f) \end{bmatrix} = \Phi(t_f, t) \begin{bmatrix} x(t) \\ \lambda(t) \end{bmatrix} + \int_{t}^{t_f} \phi(t_f, \tau) \begin{bmatrix} 0 \\ Q(\tau) r(\tau) \end{bmatrix}$$

$$= \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} x(t) \\ \lambda(t) \end{bmatrix} + \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$$

SO, GOING THRU THE SAME GARBAGE GIVES:

$$s[\phi_{11}x(t)+\phi_{12}\lambda(t)+f_1(t)]-s\,r(t_f)$$

$$=\phi_{21}x(t)+\phi_{22}\lambda(t)+f_2(t)$$

OR

$$[\phi_{22}-s\phi_{12}]\lambda(t)$$

$$=(s\phi_{11}-\phi_{21})x(t)+s(f_1-r)-f_2$$

$$\Rightarrow \lambda(t)=[\phi_{22}-s\phi_{12}]^{-1}(s\phi_{11}-\phi_{21})x(t)$$

$$+[\phi_{22}-s\phi_{12})(sf_1-sr-f_2)]$$

$$\overset{\Delta}{=} P(t)x(t)+\xi$$

ONCE WE GET λ OUR OPTIMAL CONTROL IS

$$u=-R^{-1}B^T u=R^{-1}B^T(Px+\xi)$$

$$=Kx+n \longleftarrow$$ THIS NOT IN TRACKING PROBLEM

NOW

$$\dot{\lambda}(t)=\dot{P}x+P\dot{x}+\dot{\xi}=\dot{P}x+P(Ax-BR^{-1}B^T(Px+\xi))+\dot{\xi}$$

$$=-Q(x-r)-A^T\lambda$$

$$\Rightarrow(\dot{P}+Q+PA+A'P-PBR^{-1}B^TP]x$$

$$+[\dot{\xi}-A^T\xi-PB^{-1}B]$$

IN GENERAL, MOST MUST = 0:

$\dot{P} + Q + PA + A^T P - PBR^{-1}P = 0 \Leftarrow$ RICATTI

(SAME AS REG. PROBLEM)

$\dot{\xi} - A^T\xi - PBR^{-1}B^T\xi - Qr$

BOUNDRY CONDITIONS:

$P(t_f) = S$

$\xi(t_f) = -Sr(t)$

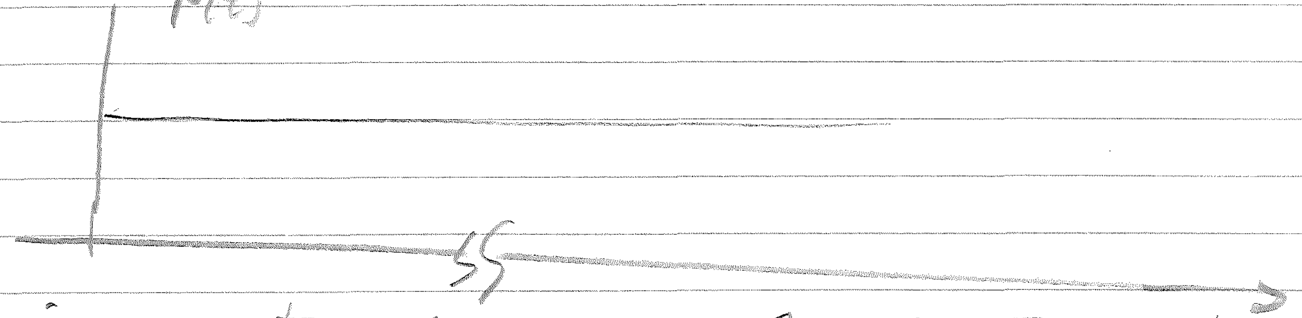
PROBLEM IS ESSENTIALL SOLVED, REMAINS TO PLUG IN.

LET $t_f = \infty$

DON'T WORRY TO MUCH PROBLEMS, DROP FINAL PENALTY TERM:

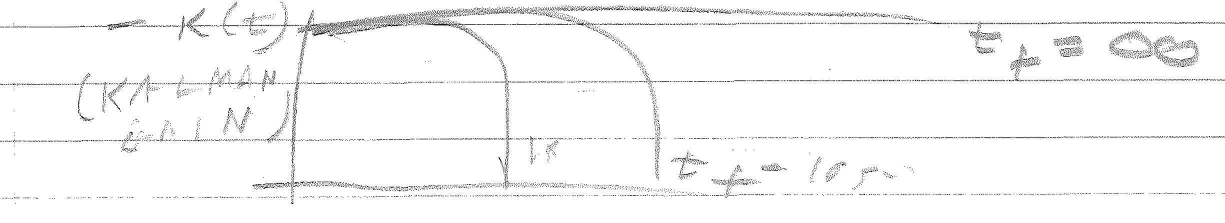
$$J = \frac{1}{2}\int_{t_0}^{\infty}\left[(\|X(t_f) - r(t)\|_Q - \|u(t)\|_R^2\right]$$

SO, LET'S INTEGRATE BACKWARDS



$\dot{P} = 0$ & RICATTI BECOMES LINEAR & WE HAVE OUR STEADY ~~STATE~~ EQUATION

$\cancel{\dot{P}} + Q + PA + A^T P - PBR^{-1}B^T P = 0$



THE BOLZA PROBLEM
(CONTROL & STATE INEQUALITY
CONTRAINTS)

$$V = \theta[x(t_f), t_f] + \int_{t_0}^{t_f} \phi(x, u, t)\, dt$$

SUBECT TO

i $\dot{x} = f(x, t)$

ii. $N = \{x(t_f), t) = 0)$ END POINT

iii $M(x(t_0), t_0) = 0$ INITIAL COND

$[x(t_0) =$

iv – $g(x, u, t) \geq 0$ CONTROL INEQ

v – $h(x(t), t) \geq 0$ STATE INEQ. CON.

10/27/76 (WED)

$$J = \Theta[x(t_f), t_f] + \int_{t_0}^{t_f} \phi(x, u, t)\,dt$$

MINIMIZE SUBJECT TO:

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

$$x(t_0) = x_0$$

$$N[x(t_f), t_f] = 0$$

$$g[x, u, t] \geq 0$$

$$\longrightarrow \text{LET } (\dot{z}_i)^2 = g_i \geq 0 \quad ; \quad z_i^2 = \begin{bmatrix} z_1^2 \\ \vdots \\ z_p^2 \end{bmatrix}$$

USE LAGRANGE MULTIPLIELS

$$J = \Theta + \xi^T x(t_0) + \nu^T N[x(t_f), t_f]$$

$$+ \int_{t_0}^{t_f} [\underbrace{\phi + \lambda^T(f}_{H} - \dot{x}) - \Gamma^T(g - \dot{z}^2)]\,dt$$

$$= \Theta + \xi^T x(t_0) + \nu^T N[x(t_f) t_f)]$$

$$+ \int_{t_0}^{t_f} [H - \lambda^T \dot{x} - \Gamma^T(g - \dot{z}^2)]\,dt$$

$$\Rightarrow H = \phi + \lambda^T f = H[x, u, t]$$

$$\text{LET} \quad u = \dot{w} \quad \text{w.l.o.g.}\; w(t_0) = 0$$

FINDING VARIATION

$$\text{LET } \Phi = H - \lambda^T \dot{x} - \Gamma^T(g - \dot{z}^2) = \Phi(x, \dot{x}, w, \dot{w}$$

$$\Theta = \Theta + \nu^T N$$

$$J_a = \Theta(x(t_f), t_f) + \xi^T x(t_0)$$

$$+ \int_{t_0}^{t_f} \Phi\,dt$$

$$\delta J_a = \delta x' \left.\frac{\delta \Theta}{\delta x}\right|_{t_f} + \delta t_f \left.\left(\frac{\delta \Theta}{\delta t_f} + \frac{\delta \Theta}{\delta x}\frac{\delta x}{\delta t}\right)\right|_{t_f}$$

$$\int_{t_0}^{t_f} \Big[\delta x^T \frac{\delta \Phi}{\delta x} + \delta w^T \frac{\delta \Phi}{\delta w} + \delta z^T \frac{\delta \Phi}{\delta z}$$

$$+ \delta \dot{x}^T \frac{\delta \Phi}{\delta \dot{x}} + \delta \dot{w}^T \frac{\delta \Phi}{\delta \dot{w}} + \delta \dot{z}^T \frac{\delta \Phi}{\delta \dot{z}} \Big] dt$$

$$+ \Phi \Big|_{t_f} \delta t_f = 0$$

WE HAVE ASSUMED $\phi = \phi(x, \dot{x}, w, \dot{w}, z, \dot{z}, t)$

NOW, AS WE'VE SHOWN BEFORE

$$\int_{t_0}^{t_f} \delta \dot{X}^T \frac{\delta \Phi}{\delta \dot{X}} dt + \delta X^T \frac{\delta \Phi}{\delta t}$$

$$= \delta X^T \frac{\delta \Phi}{\delta \dot{X}} \Bigg|_{t_0}^{t_f} - \int_{t_0}^{t_f} \delta X^T \left(\frac{\delta \Phi}{\delta X} - \frac{d}{dt} \frac{\delta \Phi}{\delta \dot{X}} \right) dt$$

NOW (SETTING $\delta\phi / \delta w = \delta \phi / \delta z = 0$)
AND INTEGRATING TERS AS WE
DID ABOVE, WE GET

$$\delta J_a = \delta x^T \frac{\delta \Theta}{\delta x}\Bigg|_{t_f} + \delta t_f \left(\frac{\delta \Theta}{\delta t_f} + \left(\frac{\delta \Theta}{\delta x}\right)^T \dot{x} + \Phi \right)$$

$$+ \delta x^T \frac{\delta \Phi}{\delta \dot{x}}\Bigg|_{t_0}^{t_f} - \int_{t_0}^{t_f} \delta x^T \left(\frac{\delta \Phi}{\delta x} - \frac{d}{dt}\left(\frac{\delta \Phi}{\delta \dot{x}}\right) \right)^{t_f} dt$$

$$+ \delta w^T \frac{\delta \Phi}{\delta \dot{w}}\Bigg|_{t_0}^{t_f} + \int_{t_0}^{t_f} \delta w^T \frac{d}{dt} \left(\frac{\delta \Phi}{\delta \dot{w}} \right) dt$$

$$+ \delta z^T \frac{\delta \phi}{\delta \dot{z}}\Bigg|_{t_0}^{t_f} + \int_{t_0}^{t_f} \delta z^T \frac{d}{dt} \left(\frac{\delta \Phi}{\delta \dot{z}} \right) dt = 0$$

AGAIN

$$\mathcal{H} = \phi + \lambda^T f$$

$$\Theta = \theta + \nu^T N$$

$$\Phi = \phi + \lambda^T (\hat{f} - \dot{x}) - \Gamma (g - (\dot{z}^2))$$

CONSIDER TERM:

$$\delta x^T \frac{\delta \Phi}{\delta \dot{x}}\Big|_{t_0}^{t_f} = \delta x^T(-\lambda)\Big|_{t_0}^{t_f}$$

$$= -\delta x^T(t_f)\lambda(t_f) + \delta x^T\overset{0}{\cancel{(t_0)}} \lambda \lambda(t_0)$$

CONSIDEDER:

$$\delta w^T \frac{\delta \Phi}{\delta \dot{w}}\Big|_{t_0}^{t_f} = \delta w^T\left(\frac{\delta H}{\delta \dot{w}} - \frac{\delta g^T}{\delta w}\Gamma\right)\Big|_{t_0}^{t_f}$$

$$= \delta w^T\left(\frac{\delta H}{\delta \dot{w}} - \frac{\delta g}{\delta \dot{w}}\right)\Gamma\Big|_{t_f} - \overset{0}{\cancel{0}}$$

CONSIDER

$$\delta z^T \delta\Phi/\delta\dot{z}\Big|_{t_0}^{t_f} = \delta z^T 2\dot{z}^T\Gamma\Big|_{t_0}^{t_f}$$

$$= \delta z^T(t_f)\begin{bmatrix} 2\Gamma_1 z_1 \\ 2\Gamma_2 \cdots \\ \vdots \\ 2\Gamma_n e_f \end{bmatrix} = 0$$

BUT SINCE $z(t_0) = 0$

THUS, MAKING THESE SUBSTITUTION INTO VARIATION:

$$\xrightarrow{} ④ \qquad \xrightarrow{} ⑤$$

$$\delta J_a = \delta(x^T)\frac{\delta \Theta}{\delta x}\Big|_{t_f} + \delta t_f\left[\frac{\delta\Theta}{\delta t_f} + \left(\frac{\delta\Theta}{\delta x}\right)^T\dot{x} + \Phi\right]\Big|_{t_f}$$

$$- \delta x^T(t_f)\lambda(t_f) - \int_{t_0}^{t_f}\delta x^T\underbrace{\left[\frac{\delta\Phi}{\delta x} - \frac{d}{dt}\frac{\delta\Phi}{\delta \dot{x}}\right]}_{①}dt$$

$$+ \delta w^T\left(\frac{\delta H}{\delta \dot{w}} - \frac{\delta g^T}{\delta w}\Gamma\right)\Big|_{t_f}$$

$$+ \int_{t_0}^{t_f}\delta w^T \frac{d}{dt}\left(\delta\Phi/\delta\dot{w}\right)dt \longrightarrow ②$$

$$+ \delta z^T 2\dot{z}^T\Gamma\Big|_{t_f} + \int_{t_0}^{t_f}\delta z^T\frac{d}{dt}\left(\frac{\delta\Phi}{\delta z}\right)dt \longrightarrow ③$$

$$= 0$$

① EULER - LAGRANGE: $\frac{\delta \phi}{\delta x} - \frac{d}{dt} \frac{\delta \phi}{\delta \dot{x}} = 0$

② $d/dt \;\; \delta\phi/\delta\dot{w} = 0 \implies \delta\phi/\delta\dot{w} = C$

$$\left.\frac{\delta\phi}{\delta w}\right|_{t_f} = C\Big|_{t_f} = 0 \implies C = 0$$

$$\implies \frac{\delta\phi}{\delta\dot{w}} = 0 \quad \forall \; t'$$

③ $\frac{d}{dt} \;\; \delta\phi/\delta\dot{z} = 0 \implies \delta\phi/\delta\dot{z} = C_2$

$$\implies \delta\phi/\delta\dot{z} = 0$$

④ $\delta \bar{x}^T(t_f) \left[\frac{\delta\Theta}{\delta x} - \lambda \right]\Big|_{t_f} = 0$

⑤ $\left[\frac{\delta\Theta}{\delta t_f} + \frac{\delta\Theta^T}{\delta x} \dot{x} + \underline{\Phi} \right]\Big|_{t_f} \delta t_f = 0$

WE WISH TO SOLVE FOR $\mathcal{H}$ AGAIN

$$\phi = H - \lambda^T \dot{x} - \Gamma^T \left[g - (\dot{z})^2 \right]$$

$$\implies \mathcal{H} = \underline{\Phi} + \lambda^T \dot{x} + \Gamma^T \left[\underset{0 \;\forall\; \text{ADMIS. INPUTS}}{\underbrace{g - \dot{z}^2}} \right]$$

THUS

$$\mathcal{H} = \dot{x}^T \lambda + \underline{\Phi}, \qquad \frac{\delta\phi}{\delta \dot{x}} = -\lambda$$

$$= \underline{\Phi} - \dot{x}\left(\frac{\delta\phi}{\delta\dot{x}}\right)$$

RECALL CORNER CONDITIONS

$$\left.\delta\Phi/\delta\dot{x}\right|_{t=t_{0^-}} = \left.\frac{\delta\Phi}{\delta\dot{x}}\right|_{t=t_1}$$

$$\left.\Phi - \dot{x}^T \frac{\delta\Phi}{\delta\dot{x}}\right|_{t=t_-} = \left.\Phi - \dot{x}^T \frac{\delta\Phi}{\delta\dot{x}}\right|_{t=t_+}$$

WE CAN GENERALIZE THIS TO THE WEIERSTRAUSS E FUNCTION.

$$E > 0$$

$$E = \Phi[x, \dot{\underline{X}}, t] - \Phi(x, \dot{x}, t) - (\dot{\underline{X}} - \dot{x})^T \, \delta\Phi/\delta\dot{x}$$

$\ni$ x = OPTIMUM VECTOR

$\underline{X}$ = ADMISSABLE VECTOR

10-29-76 (FRI)

$$\mathcal{H} = \Phi - \dot{x}^T \delta\Phi/\delta\dot{x} = \Phi - \underline{\dot{x}^T \dot{w}^T \dot{z}^T} \begin{bmatrix} \delta\Phi/\delta\dot{x} \\ \delta\Phi/\delta\dot{w} \\ d\Phi/d\dot{z} \end{bmatrix}$$

$$\lambda = - \delta\Phi/\delta\dot{x}$$

$$E = \Phi(x, \dot{\underline{X}}, t) - \Phi(x, \dot{x}, t)$$
$$- (\dot{\underline{x}} - \dot{x})^T \delta\Phi/\delta\dot{x} \geq 0$$

x = OPTIMAL VECTOR

$\underline{X}$ = ADMISSABLE VECTOR

NOW

$$\Phi(x, w, z, \dot{\underline{X}}, \dot{\underline{W}}, \dot{z}, t) - \Phi(x, w, z, \dot{x}, \dot{w}, \dot{z}, t]$$

$$- (\dot{\underline{X}} - \dot{x}]^T \delta\Phi/\delta\dot{x} - (\underline{W} - \dot{w})^T \cancel{\delta\Phi/\delta\dot{w}}^{0}$$

$$- (\dot{\underline{z}} - z)^T \cancel{\delta\Phi/\delta\dot{z}}^{0} \geq 0$$

THEN

$$\mathcal{H}[x, \dot{\underline{W}}, \lambda, t] - \dot{\underline{X}}^T \lambda - [\mathcal{H}(x, \dot{w}, \lambda, t)$$

$$- \underline{\dot{x}^T \lambda}) + \underline{X}^T \lambda - \dot{x}^T \lambda \geq 0$$

THUS

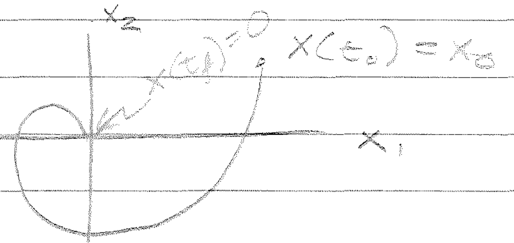
$$\mathcal{H}(x, \dot{\underline{W}}, \lambda, t) \geq \mathcal{H}(x, \underset{\uparrow}{\dot{w}}, \lambda, t)$$

OPTIMAL CONTROL

THIS IS A NECESSARY CONDITION
FOR OPTIMAL CONTROL.
(P'S MINIMUM PRINCIPLE)

$\Rightarrow$ PONTRYAGIN'S MAX. PRINCIPLE

EXAMPLE



$J = t_f;\ x(t_0) = x_0$

$x(t_f) = 0$

ALSO LET $\dot{x}(t_f) = 0$

WE WILL SHOW THAT $\mathcal{H} = -1$

RECALL

$$\frac{\delta\Theta}{\delta t_f} + \left(\frac{\delta N^T}{\delta t_f}\right)\nu + \mathcal{H} = 0$$

$$\frac{\delta\Theta}{\delta t_f} = 1 \qquad (\text{NOTE } \phi = 0)$$

$$N(t_f, x(t_f)\} = x(t_f) = 0$$

$$\Rightarrow \left(\frac{\delta N}{\delta t_f}\right)^T \nu = \left(\frac{\delta x(t_f)}{\delta t_f}\right)\nu = 0$$

$$\therefore \mathcal{H} = -1$$

EXAMPLE

MINIMIZE $J = t_f$

SUBJECT TO $\dot{x} = Ax + BU$

WITH $\|u(t)\| \leq 1$

FIND $\hat{u}$ = OPTIMAL CONTROL.

USE P'S MAX PRINCIPLE:

$$\mathcal{H}(x, u, \lambda, t) \geq \mathcal{H}(x, \hat{u}, \lambda, t)$$

$$\mathcal{H} = \phi + \lambda^T f = 0 + \lambda^T(Ax + BU)$$

THEN

$$\lambda^T(\cancel{Ax} + BU) \geq \lambda^T(\cancel{Ax} + B\hat{U})$$

$$\Rightarrow \lambda^T BU \geq \lambda^T B\hat{U}$$

LOOK @ OPTIMAL CONTROL $\Rightarrow$

WE REQUIRED THAT

$$\|U(t)\| \leq 1 \Longrightarrow |U_i(t)| \leq 1$$

$$U = \begin{bmatrix} U_1 \\ \vdots \\ U_m \end{bmatrix}$$

$$U^T \underbrace{B^T \lambda}_{q} \geq \hat{U}^T \underbrace{B^T \lambda}_{q}$$

THUS

$$U^T q \geq \hat{U}^T q$$

$$\underline{\left[U_1 \cdots U_p \right.} \left. \begin{matrix} q_1 \\ \vdots \\ q_p \end{matrix} \right] = \sum_{n=1}^{r} U_n q_n$$

$$\Rightarrow \sum_{i=1}^{r} U_i q_i \geq \sum_{i=1}^{r} \hat{U}_i q_i$$

THUS, WE WANNA MINIMIZE.

FIND $\min\limits_{U_i} \sum_{i=1}^{r} U_i q_i$

$$= \sum_{i=1}^{r} \min_{U_i} U_i q_i$$

$$= \sum_{i=1}^{r} \min_{|U_i| \leq 1} U_i q_i$$

MINIMUM OCCURS WHEN $U_i = -1$

$$\Rightarrow \min_{U_i \leq 1} U_i q_i = -|q_i|$$

OPTIMAL CONTROL IS

$\hat{U}_i = -1$ FOR $q_i > 0$

$\hat{U}_i = 1$ " $q_i < 0$

$\hat{U}_i$ UNDEFINED FOR $q_i = 0$

OR:

$$\hat{u}_i = -\operatorname{sgn} q_i$$

$$= -\operatorname{sgn} B^T \lambda$$

$$\begin{bmatrix} \hat{U}_1 \\ \hat{U}_2 \\ \vdots \\ \hat{U}_r \end{bmatrix} = -\operatorname{sgn} \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_r \end{bmatrix} \Rightarrow \hat{U} = -\operatorname{sgn} q$$

$$= -\operatorname{sgn} B^T \lambda$$

EXAMPLE (?)

MINIMIZE $\begin{cases} J = \theta + \int_{t_0}^{t_f} \phi \, dt \\ \dot{x} = Ax + BU \end{cases}$; $x(t_0) = x_0$

LET $J = \int_{t_0}^{t_f} \| u \|^2 dt$

$\dot{x} = -x + f$

FIND $f(x) = 0$ LET $\frac{df}{dt} = -f$

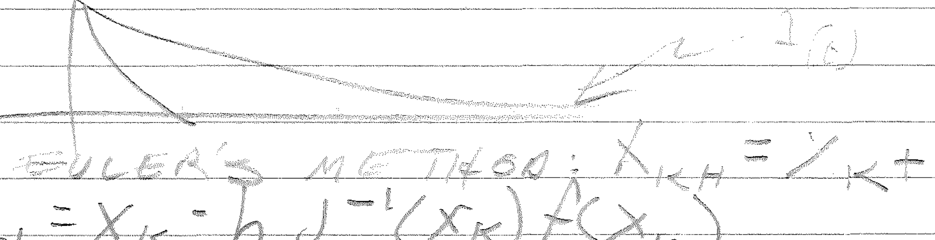
$f(t) = f(0)\, e^{-t}$

$\rightarrow 0$ AS $t \rightarrow \infty$

$$\frac{\delta f[x(t)]}{dt} = -f$$

$$\Rightarrow \frac{\delta f}{\delta x} \cdot \frac{dx}{dt} = -f$$

$$\therefore \frac{dx}{dt} = \left(\frac{df}{dx}\right)^{-1} f = -J^{-1}(f), x(t_0) = x_0$$



USE EULER'S METHOD: $X_{K+1} = X_K + h(-J^{-1} f)$

$$X_{K+1} = X_K - h J^{-1}(X_K) f(X_K)$$

(FOR $h = 1$, WE GOT NEWTON'S METHOD.

BACK TO PROBLEM

$$\frac{df}{dt} = -f \underbrace{+u}$$ ← ADD A CONTROL

$$f[x(0)] = f_0$$ ← BC →

$@\, t = t_f$, WE WANT $f[x(t_f)] = 0$

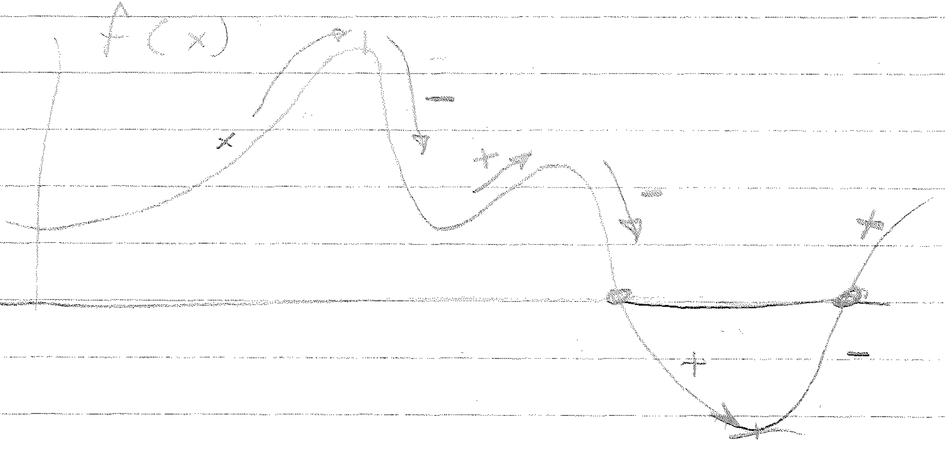
11/1/76 (MON)

Optimally Controlled Iterative Schemes

$f(x) = 0$

BRANIN'S METHOD: $\dot{x} = \pm J^{-1}(x) f(x)$

$$\dot{f} = \pm f \Rightarrow f[x(t)] = f[x(0)]\, e^{\pm t}$$

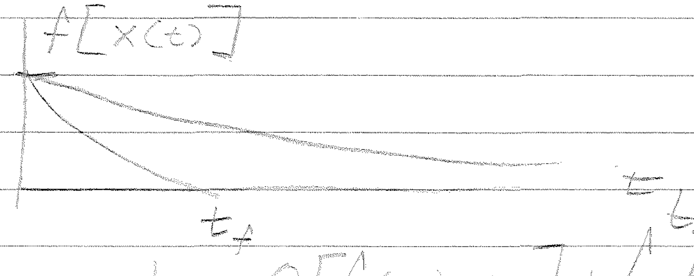


CONSIDER

$$f(x) = 0 \qquad \dot{f} = \pm f + u$$

$$f(x(t)) = f(x(0))\, e^{\pm t} = f_0\, e^{\pm t}$$

ALSO, $f(x(t_f)) = 0$



LET $\quad J = \theta[f(t_f), t_f] + \int_{t_0}^{t_f} \phi(f, \dot{f}, u, t)\, dt$

SUBJECT TO

$$\dot{f} = -f + u$$

B.C. $\Rightarrow \quad f(t_0) = f_0$

AND $\quad f(t_f) = 0$

LET'S PICK A MORE SPECIFIC PERFORMANCE INDEX

LET $\quad J = \frac{1}{2} \int_{t_0}^{t_f} \|u\|^2 dt$

COSTATE EQUATION IS

$$\dot{\lambda} = -\delta H / \delta f = \lambda \qquad ①$$

$$\dot{f} = \delta H / \delta \lambda = -f + u \qquad ②$$

$$\left(H = \tfrac{1}{2}\|u\|^2 + \lambda^T(-f + u) \right)$$

ALSO

$$\delta H / \delta u = 0 \Rightarrow u + \lambda = 0 \qquad ③$$

$$\dot{\lambda} = \lambda \Rightarrow \lambda = \lambda_0 e^t$$

$$u = -\lambda \Longrightarrow u = -\lambda_0 e^t$$

$$\dot{f} = -f + u = -f - \lambda_0 e^t$$

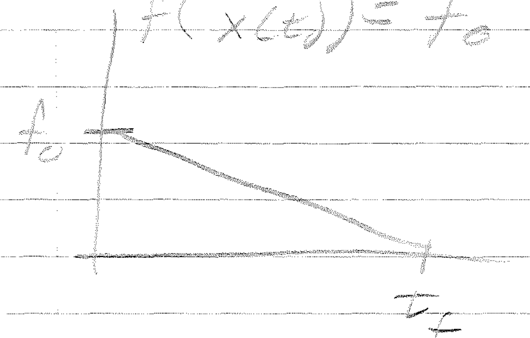
HOMO $\longrightarrow f(t) = k e^{-t}$ + PARTICULAR

SOLUTION IS

$$\hat{u} = \left(\frac{2 e^{-t_f} e^t}{e^{-t_f} - e^{t_f}} \right) f_0 = - \frac{e^{-(t_f - t)} f_0}{\sinh t_f}$$

THEN, IT TURNS OUT THE OPTIMAL TRAJECTORY IS

$$\hat{f}(x(t)) = f_0 \sinh(t_f - t)/\sinh t_f$$



NOTE, THOUGH, WE CAN'T HERE SOLVE FOR $x(t)$.

NOW,

$$\dot{f} = -f + \hat{u}$$

$$\delta f/\delta x \cdot \delta x/\delta t = -f + \hat{u}$$

$$\dot{x} = (\delta f/\delta x)^{-1}(-f + \hat{u})$$

$$= (\delta f/\delta x)^{-1}\left[-f - \frac{e^{-(t_f - t)}}{\sinh t_f} f_0\right]$$

WE MUST INTEGRATE THIS BABY. (TO DO CORRECTLY WE SHOULD USE JACOBIAN SIGN CHANGE)

INEGRATE

$$\dot{x} = (\delta f/\delta x)^{-1}\left(-f - \frac{e^{-(t_f - t)}}{\sinh t_f} f_0\right)$$

DIVY $[t_0, t_f]$ INTO N INTERVALS. X_N WILL BE OUR SOLUTION.

USING EULER:

$$X_{k+1} = X_k - h\left(f + \frac{e^{-(t_f - t)}}{\sinh t_f} f_0\right)$$

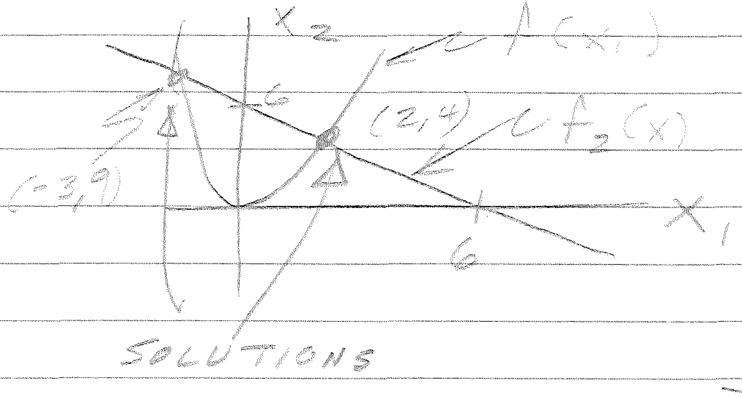
$$= X_k - h \coth(t_f - t_k)\, J^{-1}(x_k) f(x_k)$$

$$K = 0, 1, 2 \ldots N-1$$

EXAMPLE:

$f_1(x) = x_1^2 - x_2 = 0$

$f_2(x) = x_1 + x_2 - 6 = 0$



ASSUME $t_f = 1$ SEC $\left.\begin{matrix} \\ \end{matrix}\right\} \Rightarrow N = 25$

$h = 0.04$

INITIAL GUESSES:

$x_0 (3,6) \Longrightarrow x_{25} (1.99, 3.979)$

NOW

$x_{k+1} = x_k - h \coth(t_f - t_k) J^{-1}(x_k) f(x_k)$

COMPARE TO EULERS

$x_{k+1} = x_k - J^{-1}(x_k) f(x_k)$

SO WE HAVE A TIME VARYING STEP SIZE

REFERENCES:

K.S CHAO, R.J.P de FIGUERIDO "OPTIMALLY CONTROLLED ITERATIVE SCHEMES FOR OBTAINING SOLUTIONS OF A NONLINEAR EQ" INT J. OF CONTROL, 1973, VOL 18, pp. 377-384

$$\dot{x}_{n+1} = f_{n+1}$$

$$x_{n+1}(t) = \int_{t_0}^{t} \dot{x}_{n+1}(t)\,dt + \cancelto{0}{x_{n+1}(t_0)}$$

APPLY LAGRANGE:

$$J_a = \Theta + \xi^T x(t_0) + \nu^T N[x(t_f), t_f] + \int_{t_0}^{t_f} \Big[\phi + \lambda^T (f - \dot{x}) - \Gamma^T (g - \dot{z}^2) + \lambda_{n+1}^T (f_{n+1} - \dot{x}_{n+1}) \Big] dt$$

$$\delta J_a = 0$$

SO, WE GOT

$$H = \phi + \lambda^T f$$

$$\Phi = H - \lambda^T \dot{x} - \Gamma^T (g - z^2)$$

$$\tilde{\Phi} = \Phi + \lambda_{n+1} (f_{n+1}^{(x,t)} - \dot{x}_{n+1})$$

GIVES

$$\frac{\delta \tilde{\Phi}}{\delta x} - \frac{d}{dt} \frac{\delta \tilde{\Phi}}{\delta \dot{x}} = 0$$

(1) $(*)$ $\dfrac{\delta \Phi}{\delta x} + \lambda_{n+1} \dfrac{\delta f_{n+1}}{\delta x} - \dfrac{d}{dt} \dfrac{\delta \Phi}{\delta \dot{x}} = 0$

(2) $(*)$ $\dot{\lambda}_{n+1}(t) = 0$

$$\cancelto{0}{\frac{\delta \tilde{\Phi}}{\delta x_{n+1}}} - \frac{d}{dt} \frac{\delta \tilde{\Phi}}{\delta \dot{x}_{n+1}} = 0$$

$$= -\frac{d}{dt}(\lambda_{n+1}) \Longrightarrow \dot{\lambda}_{n+1} = 0$$

FROM PREVIOUS PROBLEM

(3) $\frac{\delta \Phi}{\delta \dot{w}} - \frac{\delta \bar{\Phi}}{\delta u} = 0$

(4) $\delta \bar{\Phi} / \delta \dot{z} = 0$

(5) $\lambda = \delta \Phi / \delta x + \left(\frac{\delta N}{\delta x} \right)^T \nu \quad ; \quad t = t_f$

(6) $\frac{\delta \Phi}{\delta t} + \left(\frac{\delta N}{\delta t} \right)^T \nu$

$- \dot{x}^T \frac{\delta \bar{\Phi}}{\delta x} + \hat{\Phi} = 0$

(7)

$$\delta z^T \begin{bmatrix} 2 \dot{z}_1 \Gamma_1 \\ \vdots \\ 2 z_r \Gamma_r \end{bmatrix} = 0$$

(8) $\frac{\delta H}{\delta \dot{w}} - \left(\frac{\delta \Phi}{\delta \dot{w}} \right)^T \Gamma = 0$

BANG-BANG CONTROL

$$\dot{x} = f(x,u,t) = f(x,t) + G(x,t)\,u(t)$$
$$J = \Theta[x(t_f), t_f] + \int_{t_0}^{t_f} [\phi(x,t) + h^T(x,t)\,u]\,dt$$

HAMILTONIAN:

$$\mathcal{H} = \phi + h^T u + \lambda^T (f + Gu)$$

H WILL BE MINIMUM @ OPTIMAL CONTROL:

$$H[x,\hat{u},t] \geq H[x,u,t]$$

$$\phi(x,t) + h^T(x,t)\,\hat{u} + \lambda^T f(x,t) + \lambda^T G\hat{u}$$
$$\leq \phi + h^T u + \lambda^T (f + Gu)$$

OR, CANCELLING

$$(h^T + \lambda^T G)\,\hat{u} \leq (h^T + \lambda^T G)\,u$$

WE MUST BOUND CONTROL:

$$a_i \leq u_i \leq b_i$$

NOW, IF $(h^T + \lambda^T G)_i > 0$

$$\Rightarrow \hat{u}_i = a_i$$

BUT, IF $(h^T + \lambda^T G)_i < 0$

$$\Rightarrow \hat{u}_i = b_i$$

THUS, "BANG BANG"

(FOR $(h^T + \lambda^T G)_i = 0$, $\hat{u}_i$ = UNDEFINED.

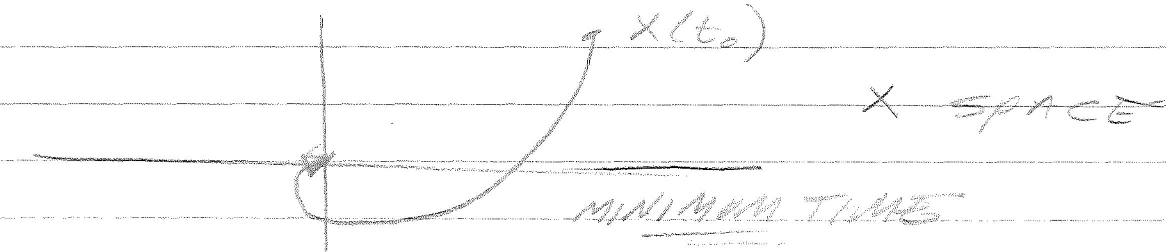
THIS IS CALLED A SINGULAR PROBLEM

IN ORDER TO SOLVE BANG-BANG, WE GOTTA SOLVE

$$\dot{x} = f + Gu = \delta H / \delta \lambda$$
$$\dot{\lambda} = - \delta \mathcal{H} / \delta x$$

MINIMUM TIME PROBLEM

$\dot{x} = Ax + bu$; U IS A SCALER



$$J = \int_{t_0=0}^{t_f} dt = t_f$$

$\exists \; -1 \le u < 1$

$$\mathcal{H} = 1 + \lambda^T(Ax + bu)$$

USING MAX PRINCIPLE

$$1 + \lambda^T(Ax) + \lambda^T b\hat{u} \le 1 + \lambda^T bu$$

$$\lambda^T b\hat{u} \le \lambda^T bu$$

$$\hat{u} = -\mathrm{sgn}(\lambda^T b)$$

TEST WILL COVER UP TO 3, INCLUDING MAX PRINC, CLOSED BOOK

11/8/76 (MON)

MINIMUM TIME PROBLEM

$\dot{x} = Ax + bu$

$J = t_f = \int_0^{t_f} dt \Rightarrow \phi = 1$

$-1 \leq u \leq 1$

SOLUTION:

$$H(x, \hat{u}, t) \leq H(x, u, t)$$

$$H \triangleq \phi + \lambda^T(Ax + bu) = 1 + \lambda^T(Ax + bu)$$

$$\cancel{1} + \lambda^T(\cancel{Ax} + b\hat{u}) \leq \cancel{1} + \lambda^T(\cancel{Ax} + bu)$$

$$\lambda^T b\hat{u} \leq \lambda^T bu \Rightarrow \hat{u} = -\mathrm{sgn}(\lambda^T b)$$

SUBSTITUTE INTO OUR ORIGINAL:

$$\left\{ \begin{array}{ll} \dot{x} = Ax + bu & \\ \quad = Ax - b\,\mathrm{sgn}(\lambda^T b) & ; x(0) = x_0 \\ \Rightarrow \dot{\lambda} = -\delta H / \delta x = -A^T \lambda & \lambda(t_f) = \lambda_{t_f} \end{array} \right.$$

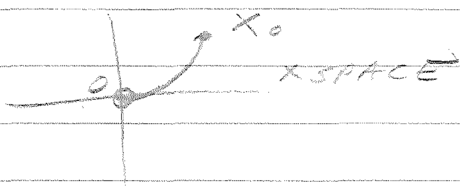
NOW $\frac{\delta H}{\delta t} = 0 \Rightarrow H = $ CONST ON OPTIMAL

$$H + \cancel{\frac{\delta \Theta}{\delta t_f}}^{0} + \frac{\delta N^T}{\delta t_f} \nu = 0$$

$$x(t_f) = 0$$

$$\dot{x}(t_f) = 0$$

$$\Rightarrow H = 0$$



$$\Rightarrow \lambda(t) = e^{-A^T(t - t_f)} \lambda(t_f)$$

LET $\tau = t_f - t$, $t \in [t_0, t_f]$

$$x(t) = x(-\tau + t_f) \triangleq \xi(\tau)$$

$$d\tau = -dt$$

$\dot{x} = Ax + bu$ BECOMES $\Longrightarrow$

$$-\frac{\delta\xi(\tau)}{\delta\tau} = A\xi - b\,\mathrm{sgn}\,\lambda^T b$$

BUT

$$\Rightarrow \lambda(t) = e^{-A^T(t-t_f)}\lambda(t_f)$$

$$= e^{A^T\tau}\lambda(t_f)$$

$$\Rightarrow -\frac{\delta\xi(\tau)}{\delta\tau} = A\xi - b\,\mathrm{sgn}\left[e^{A^T\tau}\lambda(t_f)\right]^T b$$

$$\therefore \frac{\delta\xi(\tau)}{\delta\tau} = -A\xi + \underbrace{b\,\mathrm{sgn}\left(\lambda^T(t_f)\,e^{At}b\right)}_{\text{FORCING}}$$

HOMO SOLUTION:

$$\xi(t) = \xi(0)e^{-At} + \text{PARTIC. SOLN}$$

$$= \overset{=x(t_f)=0}{\cancel{\xi(0)}}e^{-At} + \int_0^\tau e^{-A(\tau-s)} \times b\,\mathrm{sgn}\left[\lambda^T(t_f)e^{As}b\right]ds$$

$$x(t_f) = \xi(0)$$

$$\therefore \xi(t) = \int_0^\tau e^{-A(\tau-s)}b\,\mathrm{sgn}\left[\lambda^T(t_f)e^{As}b\right]ds$$

EXAMPLE

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) \qquad ; x(0) = x_0$$

$$\Phi = e^{At} = I + At + \frac{A^2 t^2}{2!} + \cdots \to 0$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} t + \cancel{\frac{1}{2!} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} t^2}$$

$$= \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$$

NOW

$$\xi(\tau) = \int_0^\tau \underbrace{\begin{bmatrix} 1 & -\tau+s \\ 0 & 1 \end{bmatrix}}_{e^{-A(\tau-s)}} \underset{\substack{\uparrow \\ b}}{\begin{bmatrix} 0 \\ 1 \end{bmatrix}} \operatorname{sgn} \underset{\substack{\uparrow \\ \lambda(t_f)}}{\underline{\lambda_{t_f} \lambda_{2t_f}}}$$

$$\times \underset{\substack{\uparrow \\ e^{As}}}{\begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}} \underset{\substack{\uparrow \\ b}}{\begin{bmatrix} 0 \\ 1 \end{bmatrix}} ds$$

$$= \int_0^\tau \begin{bmatrix} s-\tau \\ 1 \end{bmatrix} \operatorname{sgn} \left[s\lambda_1(t_f) + \lambda_2(t_f) \right] ds$$

WHEN $s\lambda_1(t_f) + \lambda_2(t_f)$, WE HAVE SWITCHING. CALL THIS SWITCHIN TIME NAMELY, $\tau = \tau_s$:

$$\tau_s \lambda_1(t_f) + \lambda_2(t_f) = 6$$

$$\Rightarrow \tau_s = \frac{-\lambda_2(t_f)}{\lambda_1(t_f)}$$

@ $\tau = \tau_s$:

$$\xi(\tau_s)_{new} = \int_0^{\tau_s} \begin{bmatrix} s-\tau_s \\ 1 \end{bmatrix} \text{sgn}\, \xi \lambda_1 \text{ — } o.T.$$

$$u = -\text{sgn}\, \lambda^T b$$

SWITCHING WHEN

$$\lambda^T b = 0$$

$$\underline{\lambda_1 \ \lambda_2} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \Rightarrow \lambda_1 = 0$$

$$\lambda^T b = \lambda(t_f)\, e^{A(t - t_f)} b = \tau_s \lambda_1(t_f) + \lambda_2(t_f)$$

$$\Rightarrow \xi(\tau_s) = \int_0^{\tau_s} \begin{bmatrix} s-\tau_s \\ 1 \end{bmatrix} \text{sgn}\left[s\lambda_1 - \tau_s \lambda_1\right]$$

$$= \int_0^{\tau_s} \begin{bmatrix} s-\tau_s \\ 1 \end{bmatrix} \text{sgn}\, \lambda_1 \underbrace{(s-\tau_s)}_{-q} ds$$

LET $s - \tau_s = -q$

$dq = -ds$

$s = 0, \ q = \tau_s$

$s = \tau_s, \ q = 0$

$$\xi(\tau_s) = \int_{\tau_s}^{0} \begin{bmatrix} -q \\ 1 \end{bmatrix} \text{sgn}\left[\lambda_1(t_f) q\right] (-dq)$$

$$= \int_0^{\tau_s} \begin{bmatrix} -q \\ 1 \end{bmatrix} \text{sgn}\left(\lambda_1(t_f) q\right) dq$$

$$q = \tau_s - s > 0$$

$$\Rightarrow \text{sgn}\, \lambda_1(t_f) q = \text{sgn}\, \lambda_1(t_f)$$

$$\xi(t_f) = -\int_0^{\tau_s} \begin{bmatrix} -q \\ 1 \end{bmatrix} \operatorname{sgn} \lambda(t_f)\, dx$$

$$= -\operatorname{sgn} \lambda_1(t_f) \left[\int_0^{\tau_s} \begin{bmatrix} -q \\ 1 \end{bmatrix} dq \right]$$

$$= -\operatorname{sgn} \lambda_1(t_f) \begin{bmatrix} -\frac{1}{2} q^2 \\ q \end{bmatrix} \Bigg|_0^{\tau_s}$$

$$= \left[-\operatorname{sgn} \lambda_1(t_f) \right] \begin{bmatrix} -\frac{1}{2} \tau_s^2 \\ \tau_s \end{bmatrix}$$

(i) $\lambda_1(t_f) > 0$

$$\begin{bmatrix} x_1(t_s) \\ x_2(t_s) \end{bmatrix} = -\begin{bmatrix} -\frac{1}{2} \tau_s^2 \\ \tau_s \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \tau_s^2 \\ -\tau_s \end{bmatrix} \geq 0$$

(ii) $\lambda_1(t_f) < 0$

$$\begin{bmatrix} x_1(t_f) \\ x_2(t_f) \end{bmatrix} = + \begin{bmatrix} -\frac{1}{2} \tau_s^2 \\ \frac{1}{2} \tau_s \end{bmatrix} \geq 0$$

CONSIDER (i):

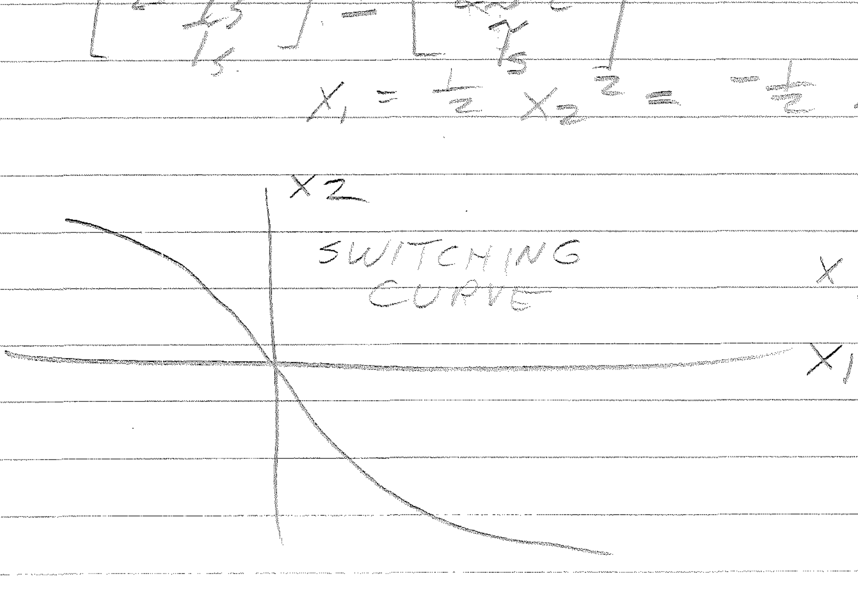
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} (\tau_s)^2 \\ -\tau_s \end{bmatrix} = \begin{bmatrix} \frac{1}{2} x_2^2 (\tau_s) \\ -\tau_s \end{bmatrix}$$

$$= -\frac{1}{2} x_2 |x_2|$$

FOR (ii)

$$\begin{bmatrix} \frac{-1}{2} \tau_s^2 \\ \tau_s \end{bmatrix} = \begin{bmatrix} \frac{1}{2} x_2^2 \\ \tau_s \end{bmatrix}$$

$$x_1 = \frac{1}{2} x_2^2 = -\frac{1}{2} x_2 |x_2| < 0$$



$$x_1 + \frac{1}{2} x_2 |x_2| = 0$$

11/10/76 (WED)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t)$$

$x(t_0) = x_0 \qquad x(t_f) = 0$

$J = \int_{t_0}^{t_f} dt$

$-1 \le u \le 1$

USE PONT'S MAX PRINCIPLE:

$H^*[x, \hat{u}, t] \le H(x, u, t)$

GIVES

$$\left.\begin{aligned} \hat{u} &= - \operatorname{sgn} \lambda^T b \\ \hat{u} &= - \operatorname{sgn} \left(\begin{bmatrix} \lambda_1 & \lambda_2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) \end{aligned}\right\}$$

$$= - \operatorname{sgn} \lambda_2 = \begin{cases} 1 \, ; & \lambda_2 < 0 \\ -1 \, ; & \lambda_2 > 0 \end{cases}$$

$$\left(\text{NOW} \; \begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = u(t) \end{cases} \Rightarrow \dot{x}_2 = \pm 1 \Rightarrow x_2(t) = \pm t + C \right)$$

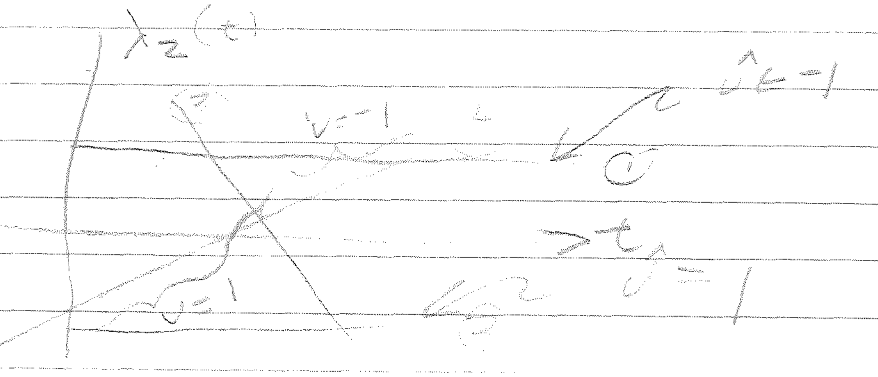
LOOK @ COSTATE:

$\dot{\lambda} = - \delta H / \delta x = - A^T \lambda$

$$\begin{bmatrix} \dot{\lambda}_1 \\ \dot{\lambda}_2 \end{bmatrix} = - \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$$

$\dot{\lambda}_1 = 0 \Rightarrow \lambda_1 = C_1$

$\dot{\lambda}_2 = - \lambda_1 \Rightarrow \lambda_2 = - C_1 t + C_2$



$x_2(t) = \pm t + C_3$

$\dot{x}_1 = x_2 \Longrightarrow x_1 = \pm \frac{1}{2} t^2 + C_3 t + C_4$

NOW $x_2^2 = t^2 \pm 2C_3 t + C_3^2$

$\frac{1}{2} x_2^2 = \frac{1}{2} \pm C_3 t \; \frac{1}{2} C_3^2$

FOR $\hat{u} = 1$

$x_1(t) = \frac{1}{2} x_2^2 - \frac{1}{2} C_3^2 + C_4$

$= \frac{1}{2} x_2^2 + C_5$

FOR $\hat{u} = -1$

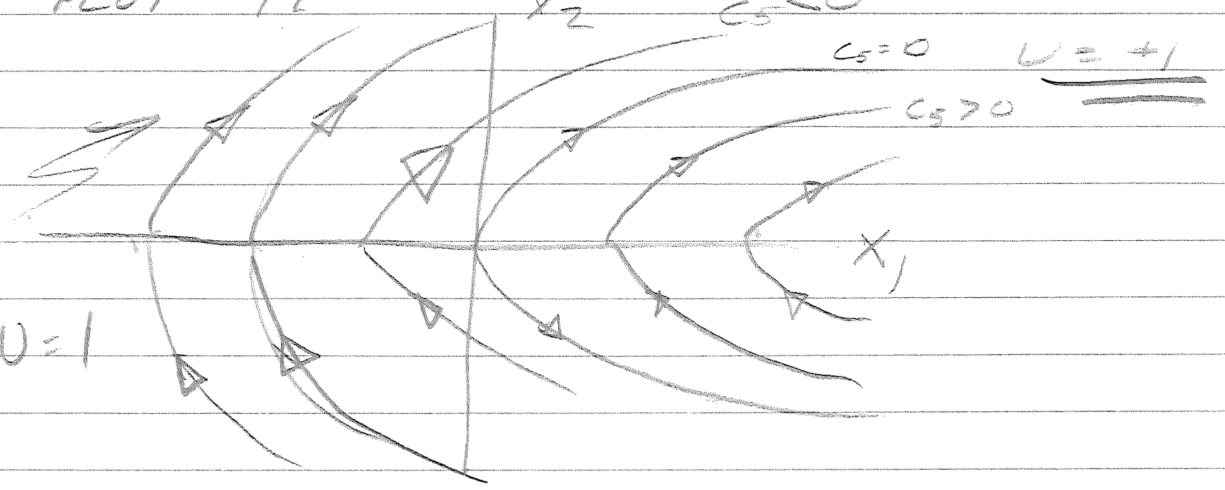
$x_1(t) = -\left(\frac{1}{2} t^2 - C_3 t\right) + C_4$

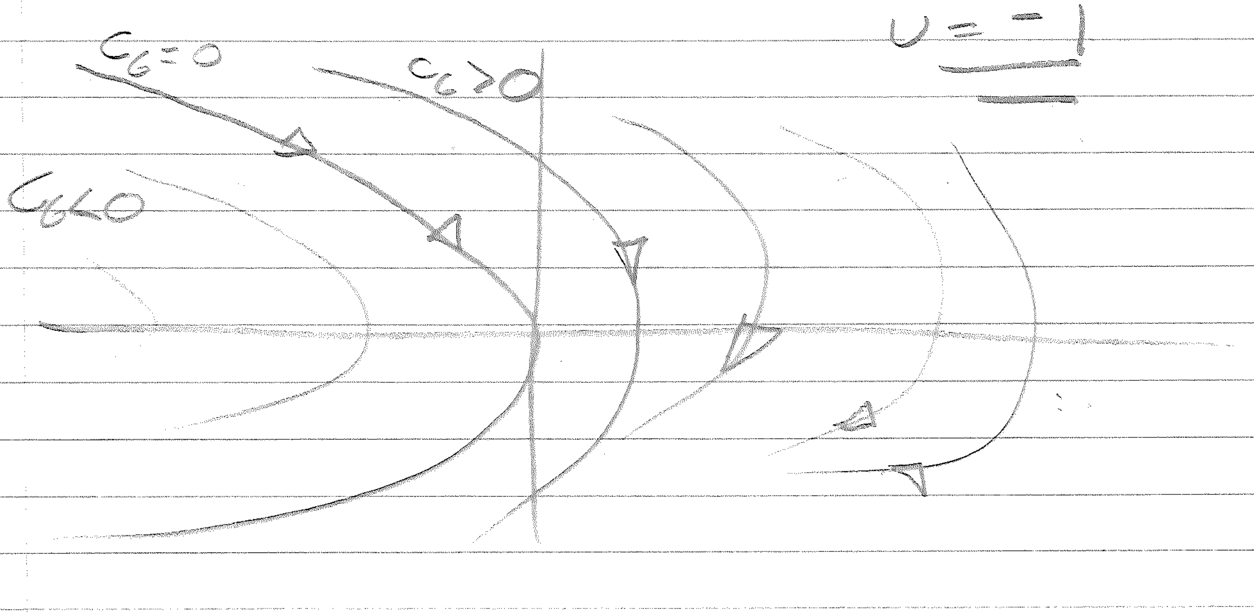
$= -\left(\frac{1}{2} x_2^2 - \frac{1}{2} C_3^2\right) + C_4$

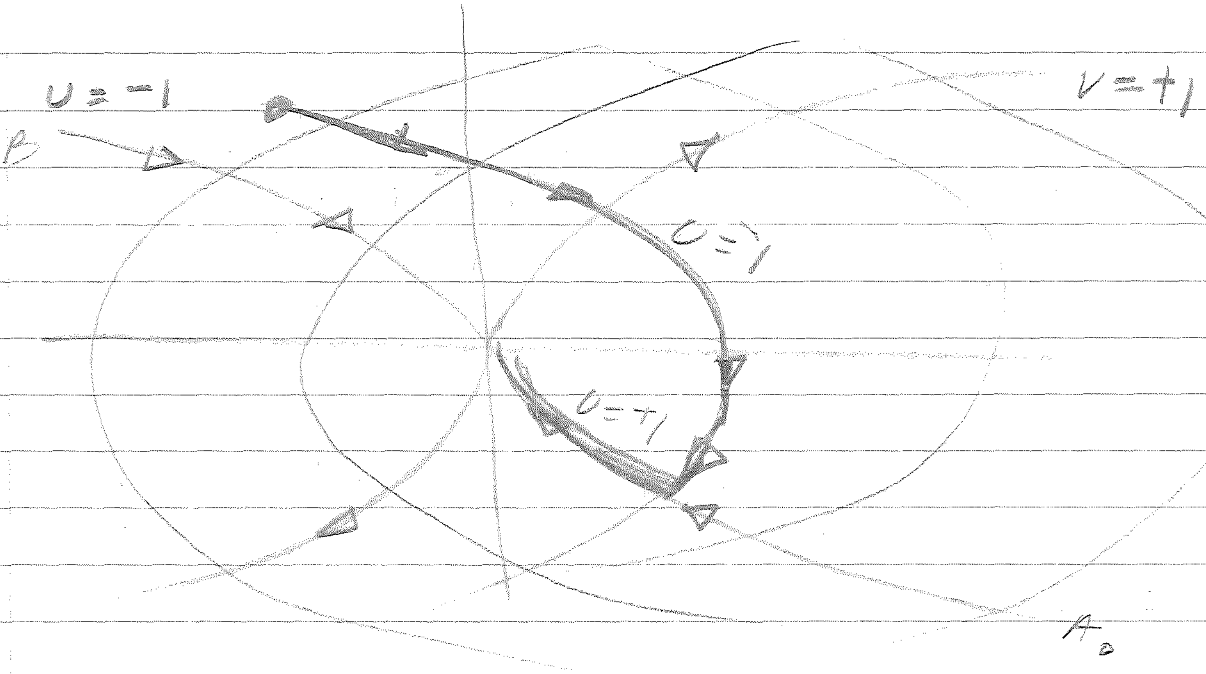
$= -\frac{1}{2} x_2^2 + C_6$

PLOT IT

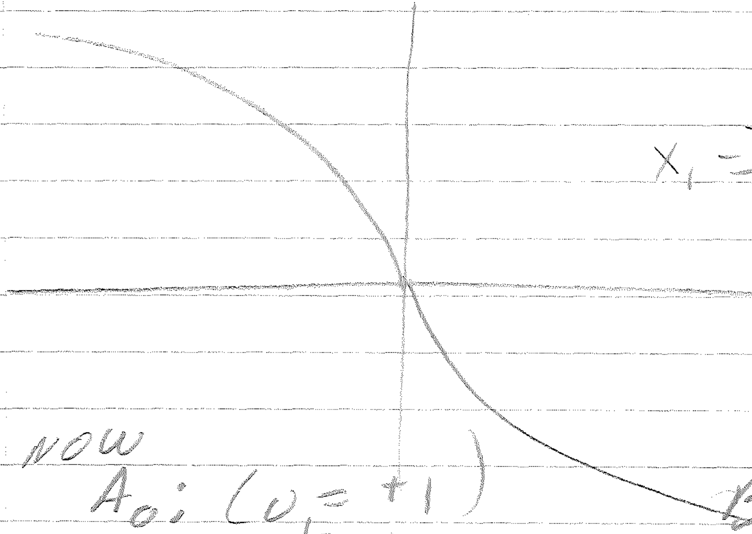
FROM $\dot{X}_2 = U = 1$







IF WE'RE ON A OR B WE CAN GO DIRECTLY TO ZERO, GET SWITCHING CURVE



$X_1 = -\frac{1}{2} X(X_1)$

NOW

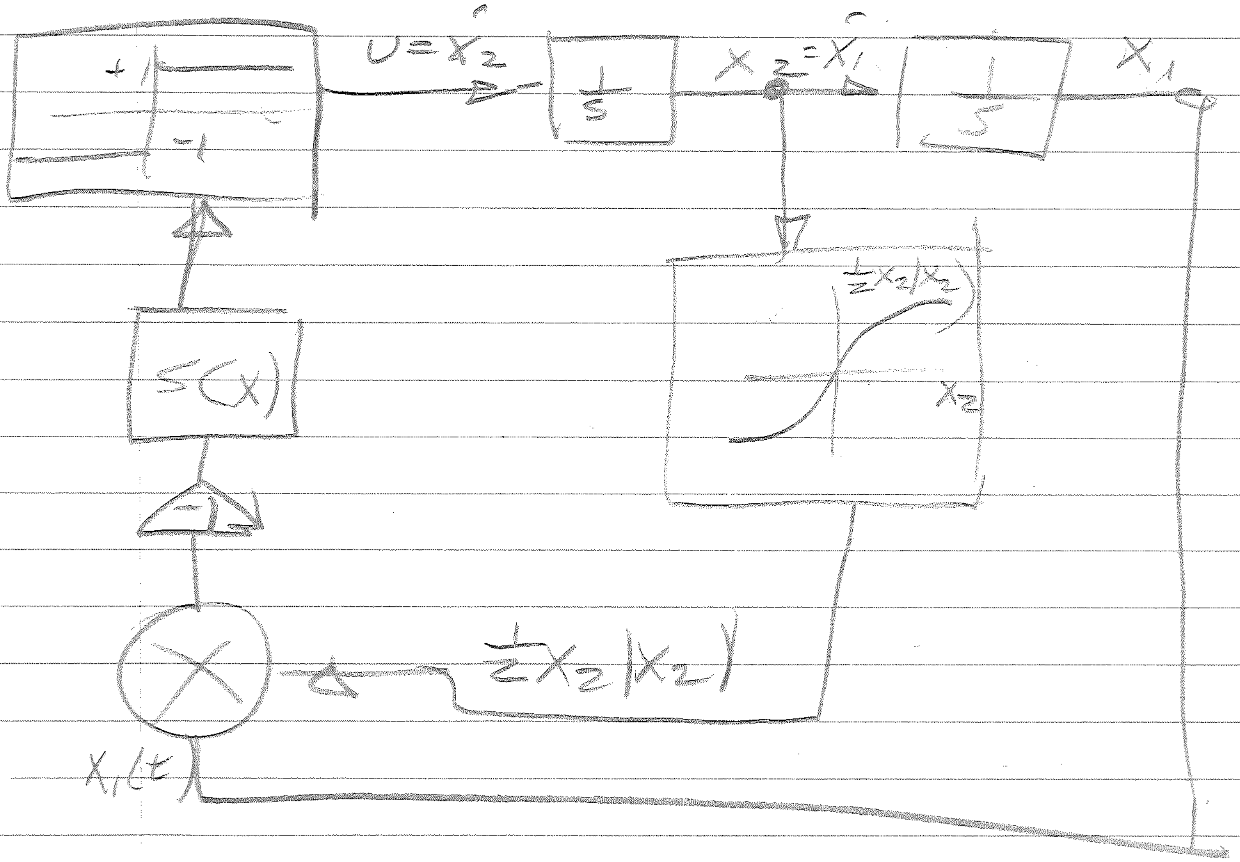
A_0: $(U = +1)$ $X = \frac{1}{2} X_2^2$

B_0: $(U = -1)$ $X(t) = -\frac{1}{2} X_2^2$

COMBINES TO GIVE

$X - \frac{1}{2} |X_2| X_2$

LOOK @ IMPLEMENTATION ; $U = -\mathrm{sgn}(S)$



MINIMUM FUEL PROBLEM

THE [illegible] RATE OF FUEL CONSUMPTION

$$J = \int_{t_0}^{t_f} c|u|dt$$

$$= \int_{t_0}^{t_f} \sum_{i=1}^{m} c_i|u_i|dt$$

= MASS OF FUEL [illegible]

$$\frac{dM}{dt} = \sum_{i=1}^{m} c_i|u_i|$$

$$\Rightarrow M(t) = M(t_0) - \int_{t_0}^{t} \sum_{i=1}^{m} c_i|u_i|dt$$

$$M(t_0) - M(t) = \sum_{i=1}^{m} \int_{t_0}^{t} c_i|u_i|dt$$

OUR PROBLEM WILL BE

MINIMIZE

$$J = \int_{t_0}^{t_f} \left(k + \sum_{i=1}^{m} c_i|u_i| + \varphi\right)dt$$

SUBJECT TO

$$\dot{x} = f(x, t) + G(x, t)u$$

$$x(t_0) = x_0, \quad x(t_f) = 0$$

$$-1 \le u_i \le 1$$

NOW

$$H = k + \sum_{i=1}^{m} c_i|u_i| + \varphi$$

$$+ \lambda^T(f + GU)$$

APPLY THE MAXIMUM PRINCIPLE $\Longrightarrow$

$\mathcal{H}(x, \hat{u}, t) \le \mathcal{H}(x, u, t)$

$\Rightarrow \cancel{K} + \sum c_i |\hat{u}_i| + \cancel{\phi(x,t)} + \lambda^T [f(x,t) + G(x,t)\hat{u}]$

$\le \cancel{K} + \sum c_i |u_i| + \cancel{\phi(x,t)} + \lambda^T [f(x,t) + G \ldots]$

OR

$\sum c_i |\hat{u}_i| + \lambda^T G \hat{u} \le \sum c_i |u_i| + \lambda^T G u$

IS SATISFIED IF ONE ... TO MINIMISE

$\lambda^T G \hat{u} = \underline{[\lambda_1 \lambda_2 \cdots \lambda_n]} \begin{bmatrix} g_{11} & \cdots & g_{1m} \\ \\ g_{n1} & \cdots & g_{nm} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}$

$\longrightarrow \quad \underline{g}_1 \qquad \underline{g}_m$

$\lambda^T G \hat{u} = [\lambda^T \underline{g}_1 \; \cdots \; \lambda^T \underline{g}_m] \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix}$

$= \lambda^T \sum_{i=1}^{m} \underline{g}_i u_i$

$\sum_{i=1}^{m} c_i |\hat{u}_i| + \lambda^T \sum_{i=1}^{m} \underline{g}_i \hat{u}_i$

$\le \sum_{i=1}^{m} c_i |u_i| + \lambda^T \sum_{i=1}^{m} \underline{g}_i u_i$

USING MAX PRINCIPLE
(TO MINIMISE →

THUS

$$c_i|\hat{u}_i| + \lambda^T \underline{g}_i \hat{u}_i \geq c_i'|u_i| + \lambda^T \underline{g}_i u_i$$

EXAMINE

$$c_i|u_i| + \lambda^T \underline{g}_i u_i$$

(i) FOR $u_i > 0$

$$c_i u_i + \lambda^T \underline{g}_i u_i = c_i + \lambda^T \underline{g}_i) u_i$$

ASSUME $c_i = 1$

$$u_i + \lambda^T \underline{g}_i u_i = (1 + \lambda^T \underline{g}_i) u_i$$

(ii) $u_i \leq 0$

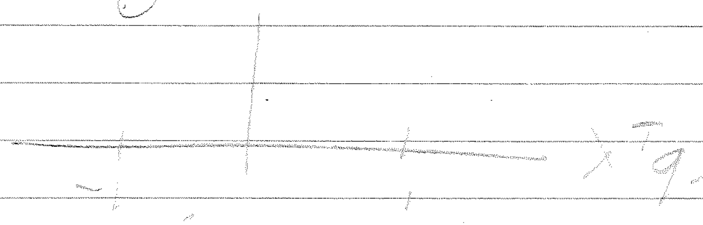
$$-\cancel{u_i} + \lambda^T \underline{g}_i u_i = (-1 + \lambda^T \underline{g}_i) u_i$$

THUS FOR $c_i = 1$:

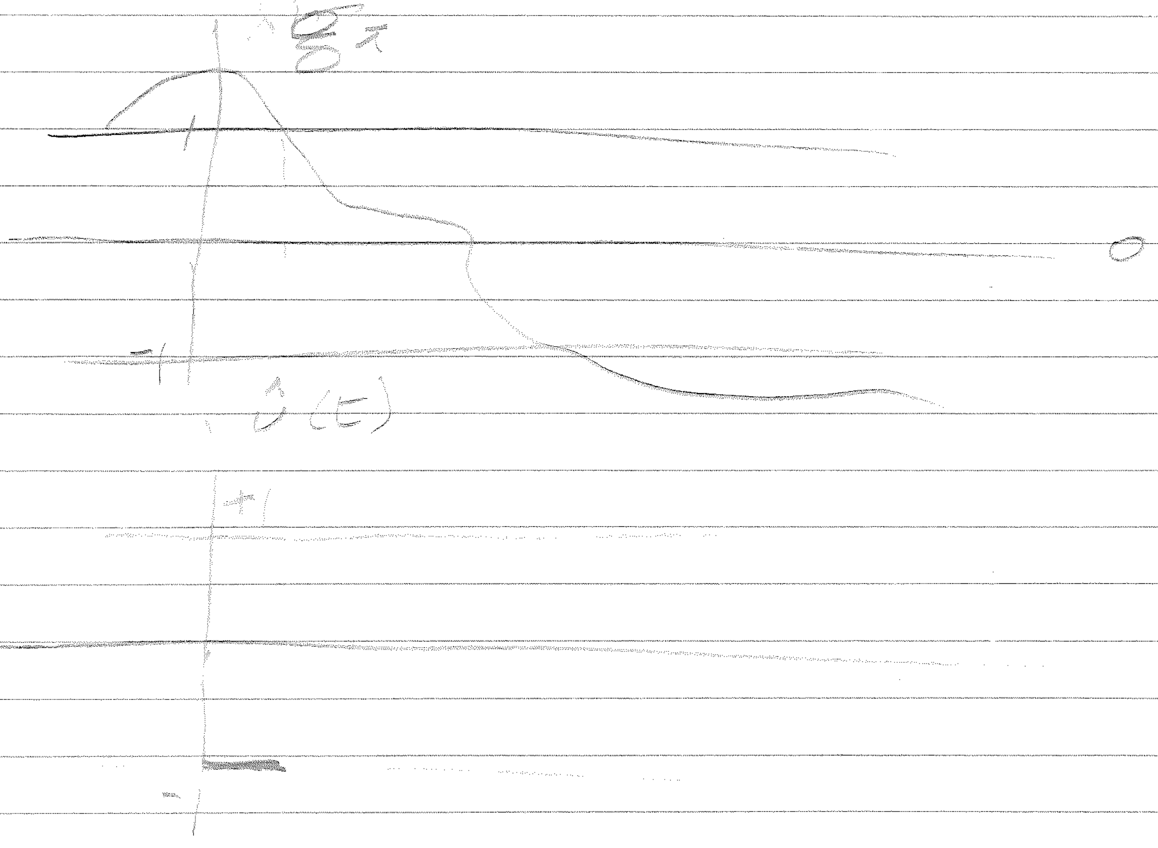
$$u_i + \lambda^T \underline{g}_i u_i = \begin{cases} (1 + \lambda^T \underline{g}_i) u_i \; ; & u_i > 0 \\ (-1 + \lambda^T \underline{g}_i) u_i \; ; & u_i \leq 0 \end{cases}$$

WANT TO MINIMIZE [illegible], LET'S CONSIDER THE FOLLOWING CASES —

(i) $\lambda^T g_i > 1$



$$* = \begin{cases} 1 + (>1)(\geq 0) \Longrightarrow \hat{v}_i = 0 \\ -1 + (>0)(\leq 0) \leq 0 \Longrightarrow \hat{v}_i = -1 \end{cases}$$



(ii) $\lambda^T g = 1$

$$* = \begin{cases} (1+1)(\geq 0) = \geq 0 \Rightarrow v_i = 0 \\ (-1+1)(\leq 0) = \leq 0 \Longrightarrow v_i = \text{ANYTHING} \end{cases}$$

NONPOSITIVE

[illegible]

≤ 1

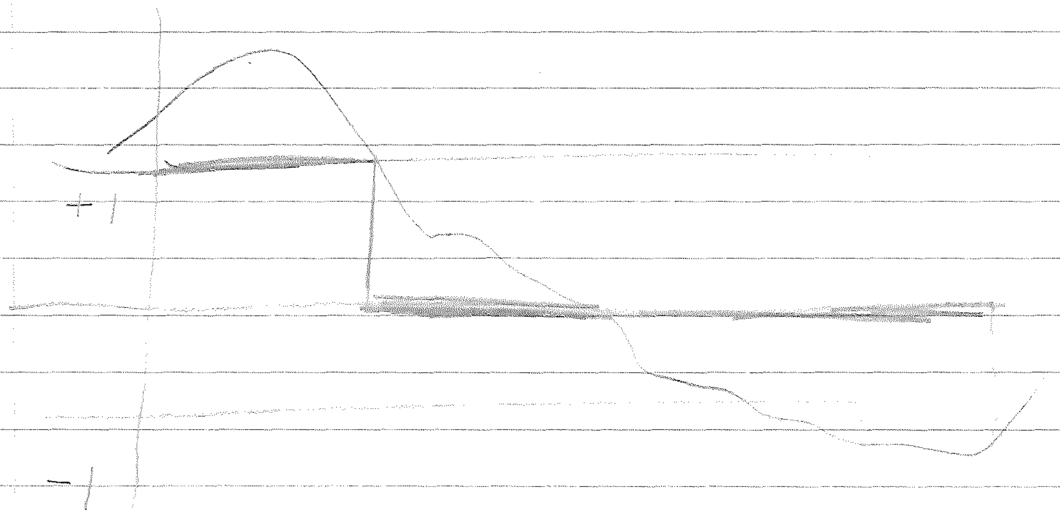
148

NG

(iii)

$0 < \lambda^T g_i < 1$

$$* = \begin{cases} (1 + (<1))(\geq 0) \geq 0 \Rightarrow \hat{U}_i = 0 \\ (-1) + (<1))(\leq 0) \geq 0 \Rightarrow \hat{U}_i = 0 \end{cases}$$



(iv)

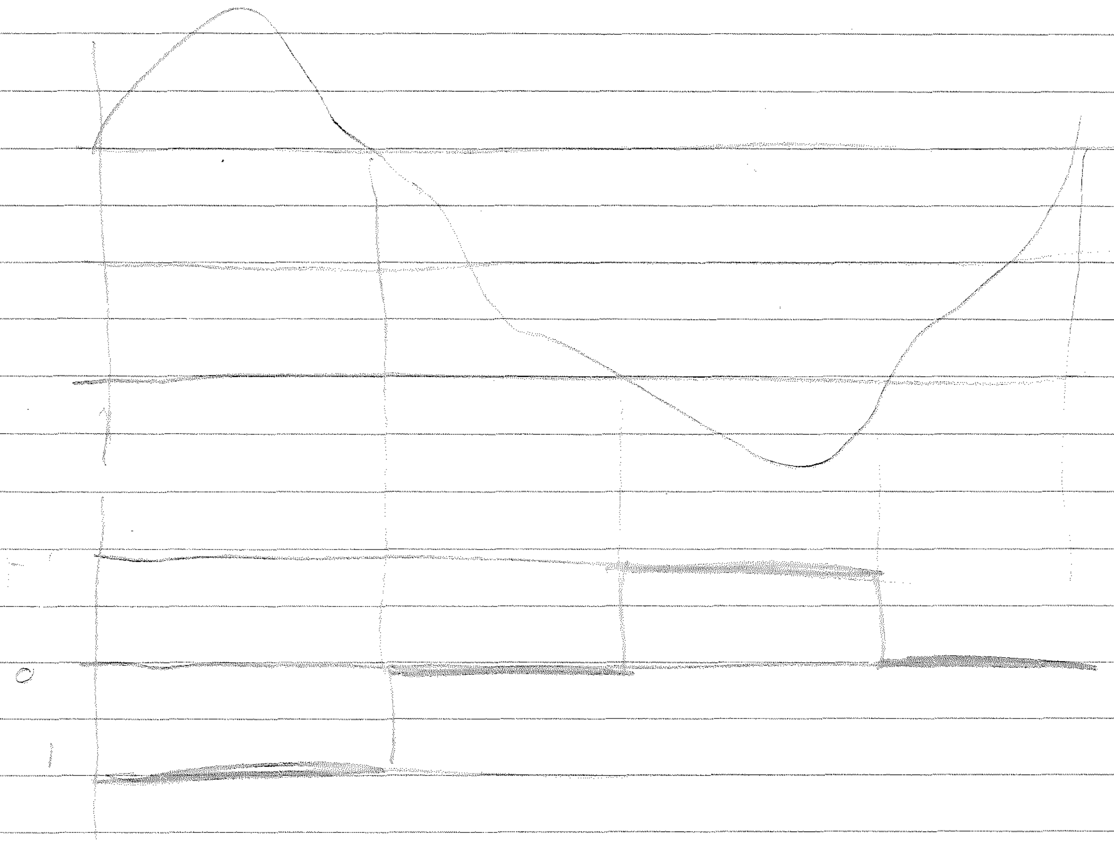
$-1 < \lambda^T g_i < 0$

$$* = \begin{cases} (1 + (>-1))(\geq 0) \geq 0 \Rightarrow \hat{U}_i = 0 \\ (-1 + (>-1)) \leq 0 \geq 0 \Rightarrow \hat{U}_i = 0 \end{cases}$$

(v) $\lambda^T g_i < -1 \Rightarrow U_i = 1$

(vi) $\lambda^T g = -1 \Rightarrow \hat{U} = $ ANY CONS VAL

MONNECK THE CONTROL



⇒ BANG OFF BANG CONTROL

11-15-76

EXAMPLE

$$J = \int_{t_0}^{t_f} \left[\frac{1}{2} x^T Q x + |u(t)| \right] dt$$

$$\dot{x} = Ax + bu \qquad x(t_0) = x_0 \qquad x(t_f) = 0$$

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad b = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad Q = \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix}$$

$$\mathcal{H} = \frac{1}{2} x^T Q x + |u(t)| + \lambda^T (Ax + bu)$$

$$= \frac{1}{2} q_1 x_1^2 + \frac{1}{2} q_2 x_2^2 + |u| + \begin{bmatrix} \lambda_1 & \lambda_2 \end{bmatrix} \left[\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \right]$$

$$= \frac{1}{2} q_1^2 x_1^2 + \frac{1}{2} q_2 x_2^2 + |u| + \lambda_1 x_2 + \lambda_2 u$$

LAST TIME WE SHOWED THAT:

$$\hat{u} = \begin{cases} 1 & \lambda^T g_i < -1 \\ 0 & -1 < \lambda^T g_i < 1 \\ -1 & \lambda^T g_i > 1 \end{cases}$$

FOR $\dot{x} = f(x,t) + G(x,t)u$

(g_i COLUMN OF G MATRIX)

FOR THIS PROBLEM, $G = b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$$\lambda^T b = \begin{bmatrix} \lambda_1 & \lambda_2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \lambda_2$$

$$\Rightarrow \quad u = \begin{cases} 1 & \lambda_2 < -1 \\ 0 & |\lambda_2| < 1 \\ 1 & \lambda_2 > 1 \end{cases}$$

WE HAVE THIS ① & STATE EQ.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ u \end{bmatrix}$$

COSTATE EQUATION. ②

$$\dot{\lambda} = -\frac{\delta H}{\delta x} = -(Qx + A^T\lambda)$$

OR

$$\begin{bmatrix} \dot{\lambda}_1 \\ \dot{\lambda}_2 \end{bmatrix} = -\left[\begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}\right]$$

OR

$$\begin{cases} \dot{\lambda}_1 = -q_1 x_1 & ; \lambda_1(t_f) = 0 \\ \dot{\lambda}_2 = -q_2 x_2 + \lambda_1 & ; x_2(t_f) = 0 \end{cases} \quad ③$$

SOLVE USING ① ② & ③

(i) FOR $\hat{u} = 1$

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = 1 \implies x_2 = t + c_1$$

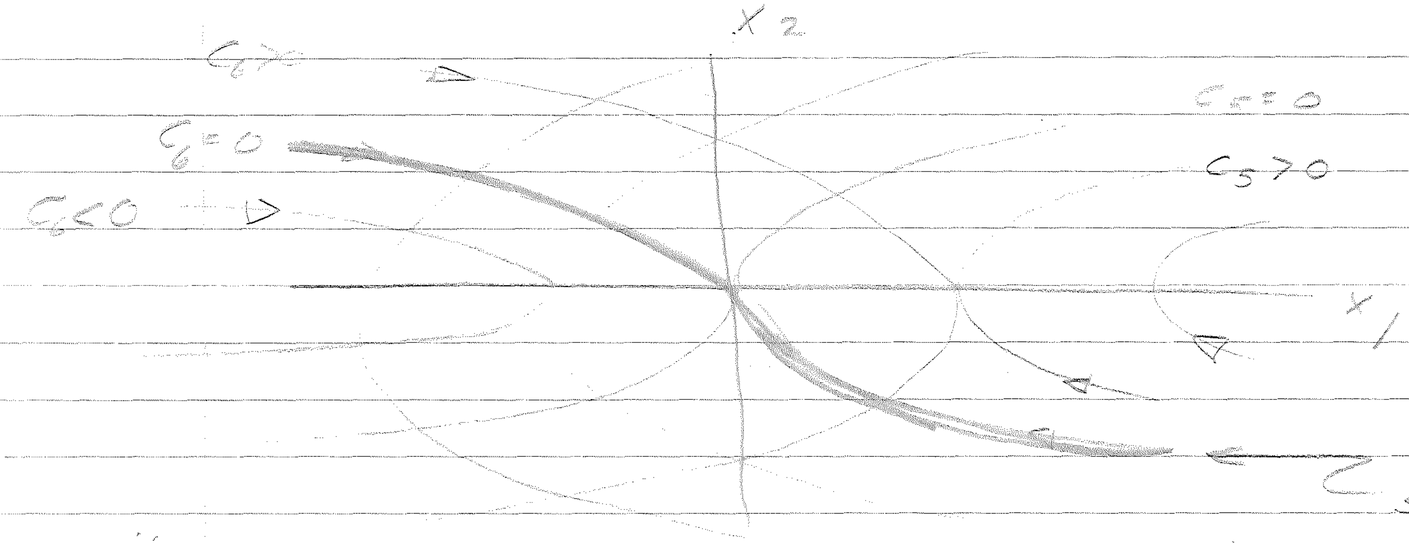
$$\therefore \dot{x}_1 = t + c$$

$$\implies x_1 = \frac{1}{2}t^2 + c_1 t + c_2$$

THIS IS SAME PROBLEM
WE WORKED BEFORE.

(ii) FOR $u = -1$ $\implies x_1 = \frac{1}{2}x_2^2 + c_5$

$$\begin{cases} x_1 = -\frac{1}{2}x_2^2 + c_5 \ ; u = -1 \\ x_1 = \frac{1}{2}x_2^2 + c_5 \ ; u = +1 \end{cases}$$



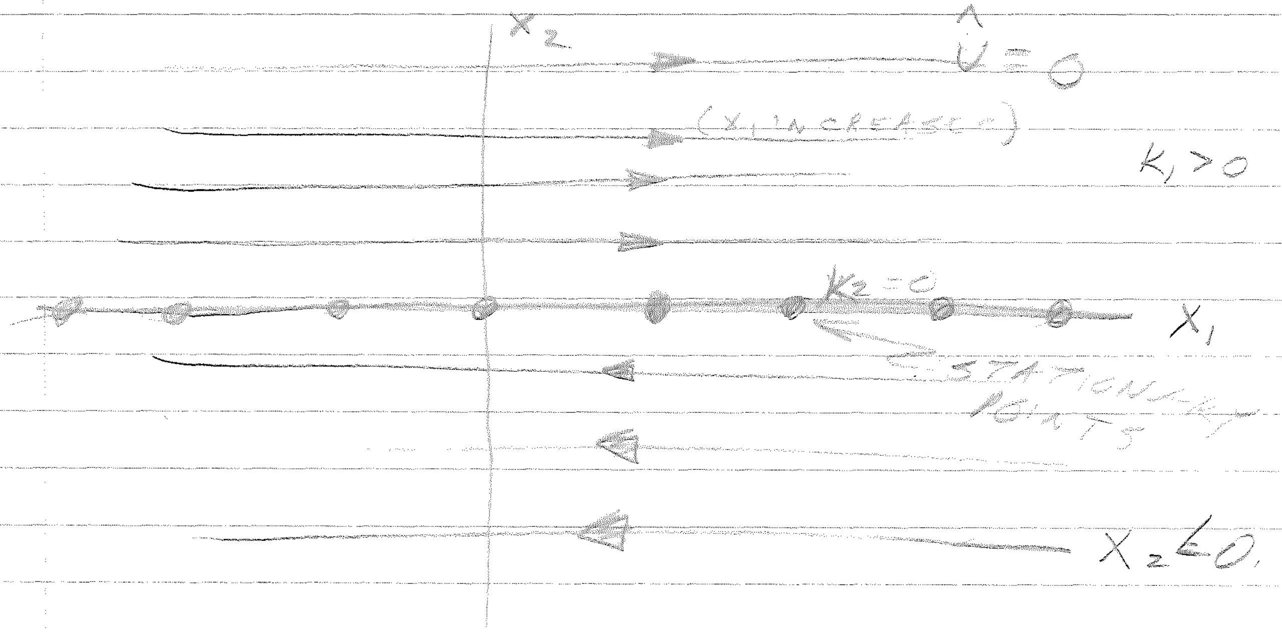
SWITCHING LINE
$x_1 + \frac{1}{2} x_2 |x_2| = 0$

iii WHAT ABOUT $u = 0$:

$\dot{x}_1 = x_2$

$\dot{x}_2 = 0 \quad \Rightarrow x_2 = K_1$

$\Rightarrow x_1 = K_1 t + K_2$



IF $K_1 = 0$, $x_1 = K_2$

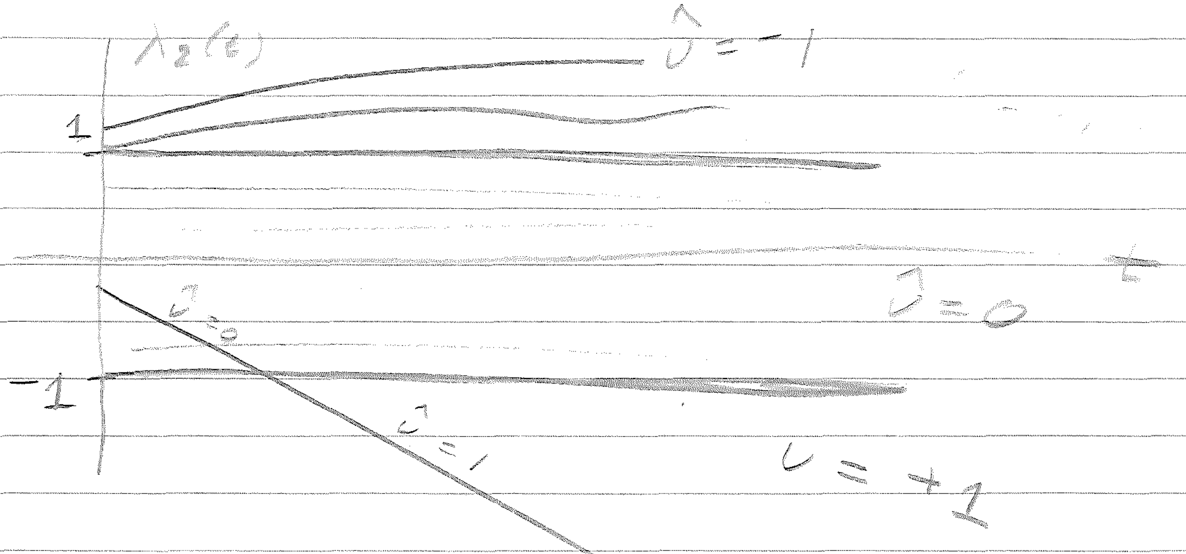
SOLVE FOR λ's:

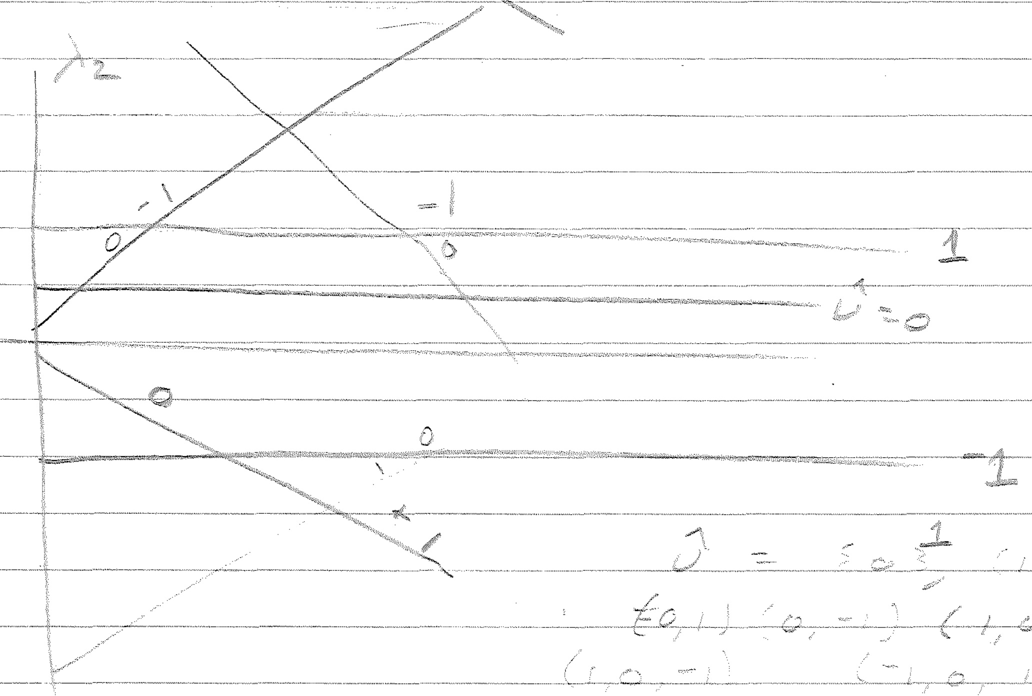
CONSIDER THE CASE FOR $Q = 0$

$\Rightarrow \lambda_1' = 0 \Rightarrow \lambda_1 = a_1$

$\lambda_2' = \lambda \Rightarrow \lambda_2 = a_1 t + a_2$

λ_2 IS A STRAIGHT LINE, CAN CHANGE SIGN ONLY ONCE. LET'S LOOK @ VARIOUS POSSIBILITIES





$\hat{u} = \{0\}^{\underline{1}}$ (1) (-1)

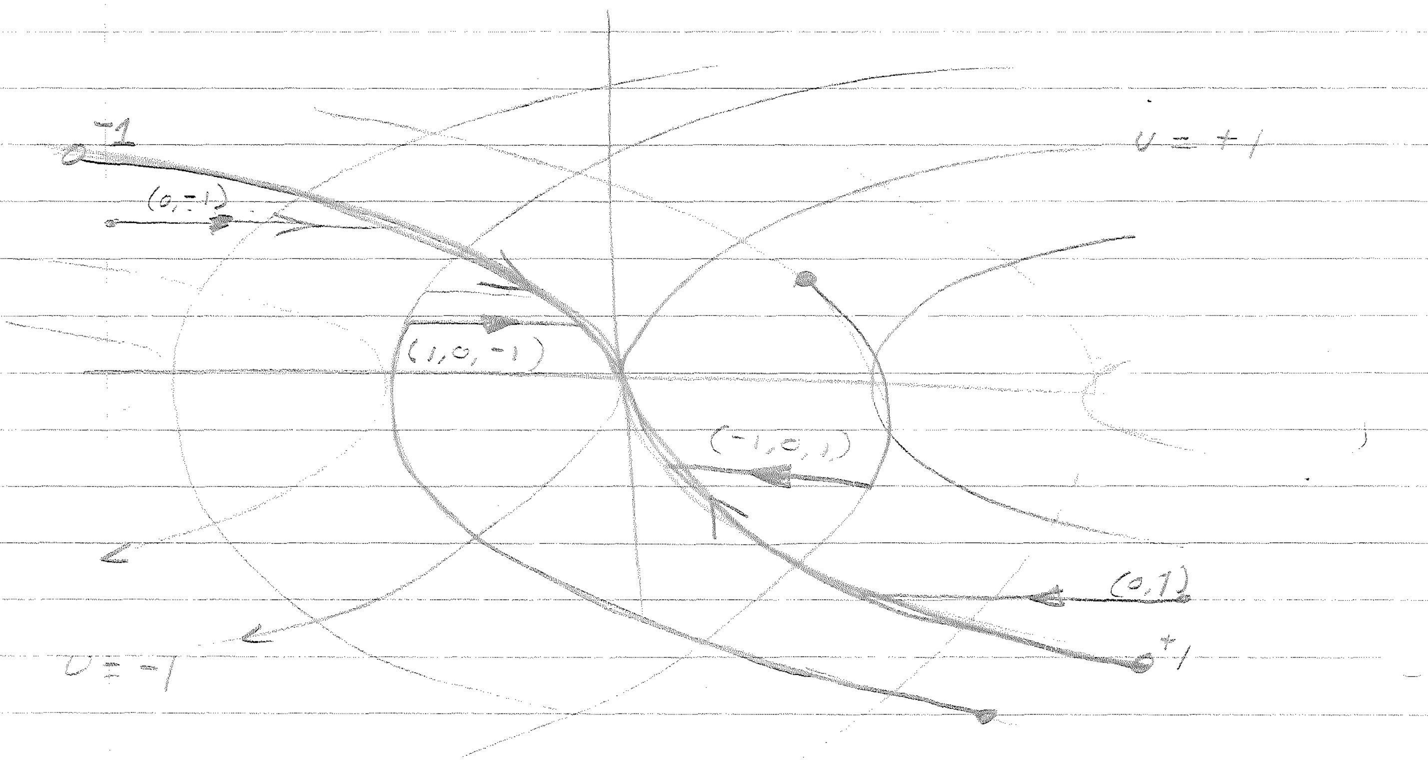
$\{0, 1\}$ $(0, -1)$ $(1, 0)^*$ $(-1, 0)^*$

$(1, 0, -1)$ $(-1, 0, 1)$

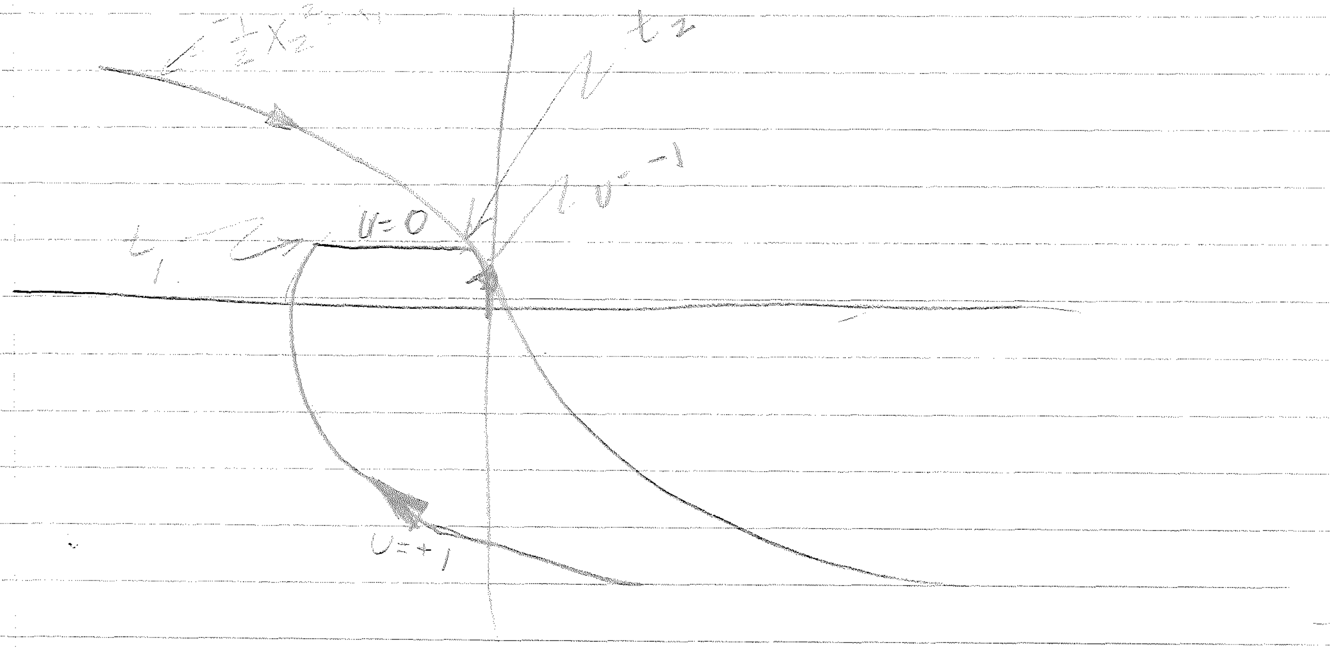
1 — WON'T WORK CAUSE OF STATIONARY PTS.

* CAN'T DO THESE SINCE THEY ARE TRPLT

OUR 6 CONTROL COMBS ARE THUS
{1} {-1}, {0,1} {1,0} {1,0,-1} (-1,0,1}



LOOK @ 1,0,-1 CASE



@ $t = t_2$, WE HAVE INTERSECTION

$X(t_2) = \frac{1}{2} X_2^2(t_2)$

AND $X_2(t_2) = x_2(t_1) = k_1$

@ $t = t_1$

$X_1(t_1) = k_1 t_1 + k_2 = x_2(t_1) t_1 + k_2$

SOLVING FOR k_2:

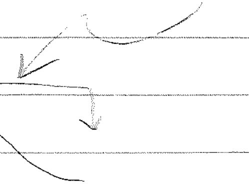
$k_2 = x_1(t_1) - x_2(t_1) t_1$

THEN

$X_1(t) = k_1 t + k_2$

$= X_2(t_1) t + x_1(t_1) - x_2(t_1) t_1$

$= \underbrace{x_1(t_1) + X_2(t_1)\,(t - t_1)}_{}$, $t_1 < t < t_2$



@ $t = t_2$, WE GET

$x_1(t_2) = x_1(t_1) + x_2(t_1)(t_2 - t_1)$

TURNS OUT $t_2 - t_1 = \infty$

11-17-76

FROM LAST TIME

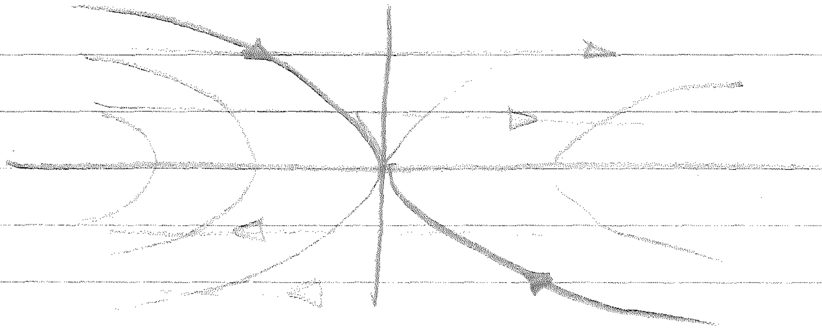
$$J = \int_{t_0}^{t_f} |u(t)|\, dt$$

$$\dot{x} = Ax + bu$$

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

WE FOUND

$$u = \begin{cases} 1 & \lambda_2 < 1 \\ 0 & -1 < \lambda_2 < 1 \\ -1 & 1 < \lambda_2 \end{cases}$$

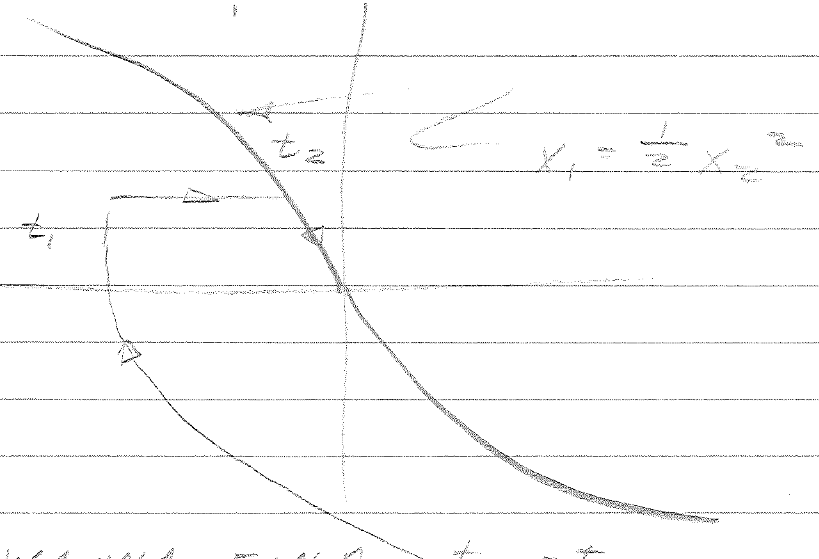


FOR $\hat{u} = 0$

$$\dot{x} = x_2, \quad \dot{x}_2 = u = 0 \Rightarrow x_2 = k_1$$

$$x_1 = k_1 t + k_2$$

CONSIDER THE $u = (-1, 0, 1)$



WE WANNA FIND $t_2 - t_1$

AT $t = t_2$, $x_1(t_2) = \frac{1}{2} x_2^2(t_2)$

ETC.

LAST LECTURE, WE SHOWED ~~THAT~~

$$x(t) = x_2(t_1)t + \overbrace{(x_1(t_1) - x_2(t_1)t_1)}^{= k_2}$$
$$= x_1(t_1) + x_2(t_1)(t - t_1)$$

CONSIDER $x_1(t_2) =$

$$x_1(t_2) = x_1(t_1) + x_2(t_1)(t_2 - t_1)$$

PAUSE TO SOLVE FOR λ:

$$\dot{\lambda} = -\frac{\delta H}{\delta x} = -A^T \lambda$$
$$= |u| + \lambda^T (Ax + Bu)$$

$$\begin{bmatrix} \dot{\lambda}_1 \\ \dot{\lambda}_2 \end{bmatrix} = -\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{matrix} \lambda_1 \\ \lambda_2 \end{matrix}$$

$\dot{\lambda}_1 = 0 \Rightarrow \lambda_1 = a$
$\dot{\lambda}_2 = -\lambda_1 = -at + b$

@ $t = t_1$ u_1 SWITCHES FROM 1 TO ZERO @ $t = t_2$ IT SWITCHES FROM ZERO TO MINUS 1.

SO

@ $t = t_1$, $u = 1 \Rightarrow u = 0$, THUS $\lambda_2 = -1$
@ $t = t_2$, $u = 0 \Rightarrow u = -1$, THUS $\lambda_2 = 1$

THEN

$$\begin{cases} \lambda_2(t_1) = -at_1 + b = -1 \\ \lambda_2(t_2) = -at_1 + b = 1 \end{cases}$$

ALGEBRA

$$(-at_1 + b) - (-at_2 + b) = -1 - (+1)$$
$$a(t_2 - t_1) = -2$$
$$\Rightarrow t_2 = -2/a$$

H IS EXPLICITLY TIME IND. THUS IT IS MINIMUM OPTIMUM TRAJECTORY. H MUST BE ZERO ON OPTIMAL TRAJECTORY. $x(t_f) = 0$

$$\mathcal{H} = [\varphi(t)] [+ \lambda^T (Ax + bu)] = \lambda_1 x_2$$

$$\mathcal{H} = 0 \text{ ON } AB$$
$$\hat{u} = 0,\ \mathcal{H} = \lambda_1 x_2 = 0 = a x_2$$
$$\Rightarrow a = 0$$

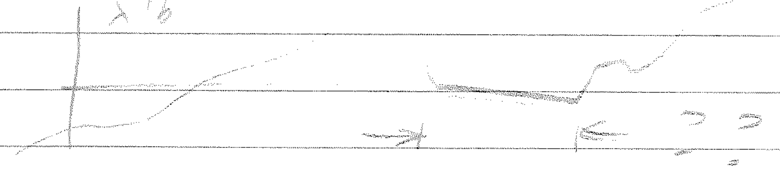
$$\therefore t_2 - t_1 = \frac{-2}{0} = \infty$$

[illegible] GET RID OF THIS WITH

$$J = \int_{t_0}^{t_f} (|u| + k)\, dt$$

THE SINGULAR PROBLEM

BANG BANG



$\dot{x} = Ax + bu$

$$\hat{u} = -\text{sgn}\,\lambda^T b = \begin{cases} 1 & \lambda^T b < 0 \\ -1 & \lambda^T b > 0 \\ & \text{INDETERMINED} \\ & \text{IF } \lambda^T b = 0 \end{cases}$$

$-1 < u < +1$

EXAMPLE

$$J = \frac{1}{2}\int_0^2 x^2\,dt \qquad t_f = 2$$

FIND $\hat{u}$ TO BRING $x(t_0)$ TO ZERO

$-1 \leq u \leq 1 \qquad \dot{x} = u$

$$\mathcal{H} = \frac{1}{2}x^2 + \lambda^T u$$

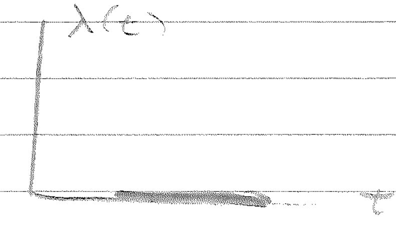
$$\hat{u} = -\text{sgn}\,\lambda = \begin{cases} 1 & ; \lambda < 0 \\ -1 & ; \lambda > 0 \\ ? & ; \lambda = 0 \end{cases}$$

WHEN $\lambda = 0$, $\mathcal{H} = \frac{1}{2}x^2 \qquad \frac{\delta \mathcal{H}}{\delta u} = \lambda = 0$

$\lambda = 0$

$\dot{\lambda} = 0$

$\ddot{\lambda} = 0$



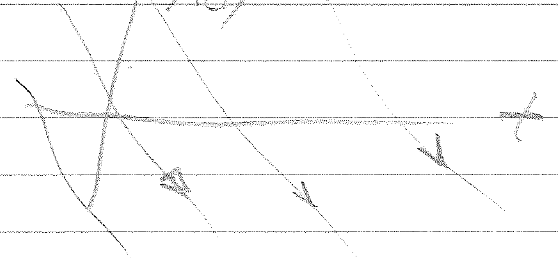
NOW

$\dot{x} = \frac{\delta \mathcal{H}}{\delta \lambda} = u \quad (- -$

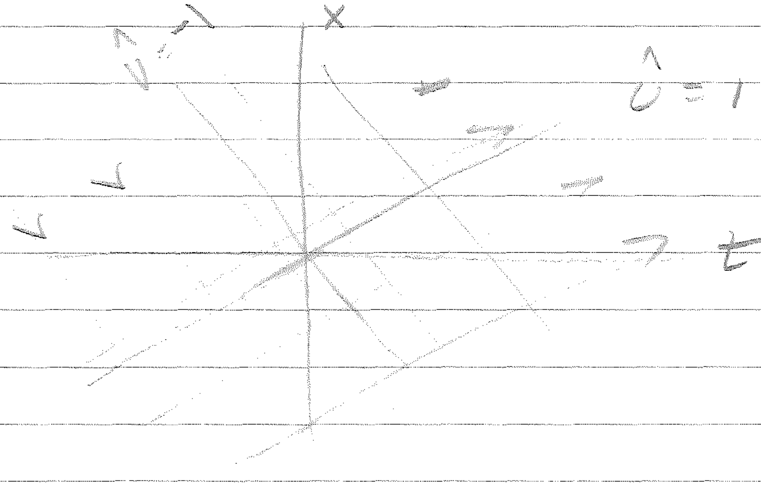
$\dot{\lambda} = -\delta\mathcal{H}/\delta x = -x$

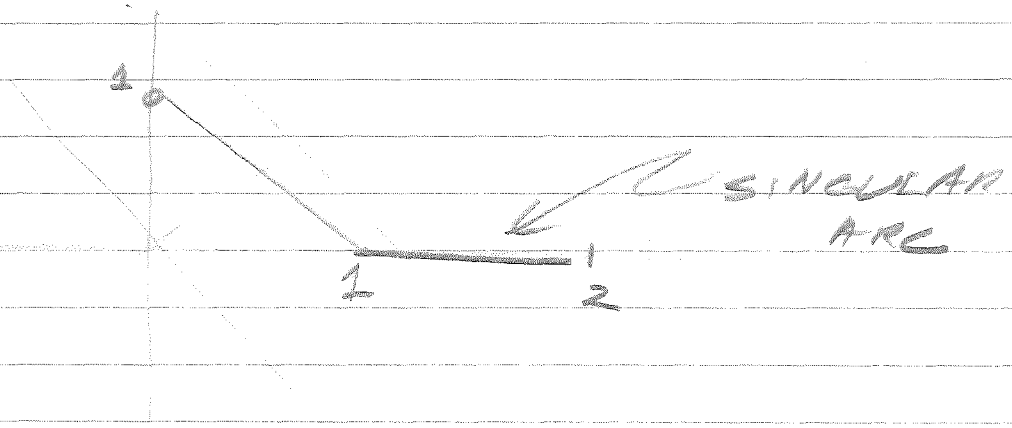
ASSUME $\hat{u} = -1 \quad (\lambda > 0)$

$\Rightarrow \dot{x} = u = -t + c$



FOR $\hat{u} = 1$ $\quad x = +t + C_2$





11-19-76

SINGULAR PROBLEM

$J = \frac{1}{2}\int_0^2 x^2 dt \qquad t_f = 2$

$\dot{x} = u(t) \qquad x(0) = 1 \qquad , x(2) = 0$

$|u(t)| \leq 1$

$\mathcal{H} = \frac{1}{2}x^2 + \lambda^T u$

FROM PONT'S MAX PRINCIPLE:

$$\hat{u} = -\text{sgn}\,\lambda = \begin{cases} -1 & ; \lambda > 0 \\ 1 & \lambda < 0 \\ = ? & \lambda = 0 \end{cases}$$

$\frac{\delta \mathcal{H}}{\delta u} = \lambda = 0$ FOR SINGULARITY

$\mathcal{H}$ NOT EXPLICIT FUNCTION OF $\mathcal{H}$ HERE.

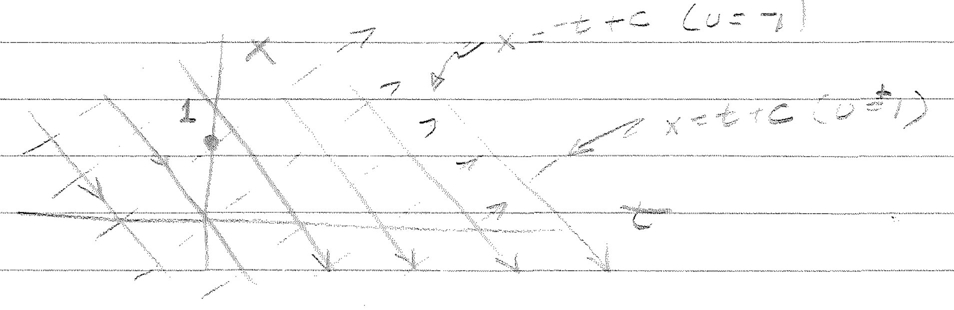
LET'S SOLVE FOR SINGULARITY

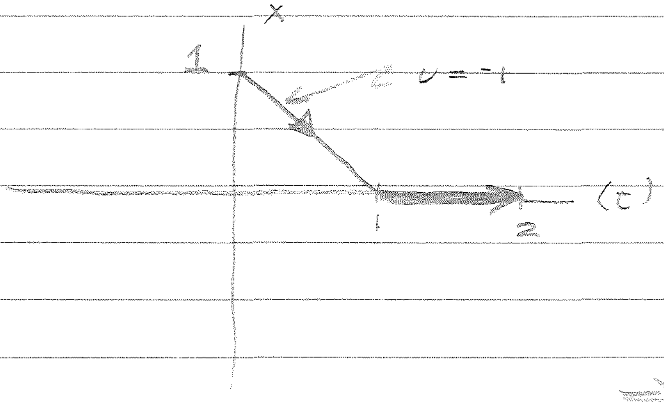
$\dot{x} = u \qquad x(0) = 1 \Leftarrow$ STATE

$\dot{\lambda} = -\frac{\delta \mathcal{H}}{\delta x} = \bar{x} \qquad x(2) = 0 \Leftarrow$ COSTATE

$\hat{u} = \pm 1$ FOR BANG-BANG OPTIMAL

$\Rightarrow \dot{x} = \pm 1 \Rightarrow x \overset{\pm}{=} t + c$





$u = -1$

$x = -t + c$

$x(0) = 1 = c \Rightarrow c = 1$

$x = -t + c$

$\Rightarrow \dot{\lambda} = -x = t - 1$

$\lambda(t) = \frac{1}{2}t^2 - t + c'$

$\lambda(t) = \frac{1}{2}t^2 - t + c'$

$c' = \lambda(0)$

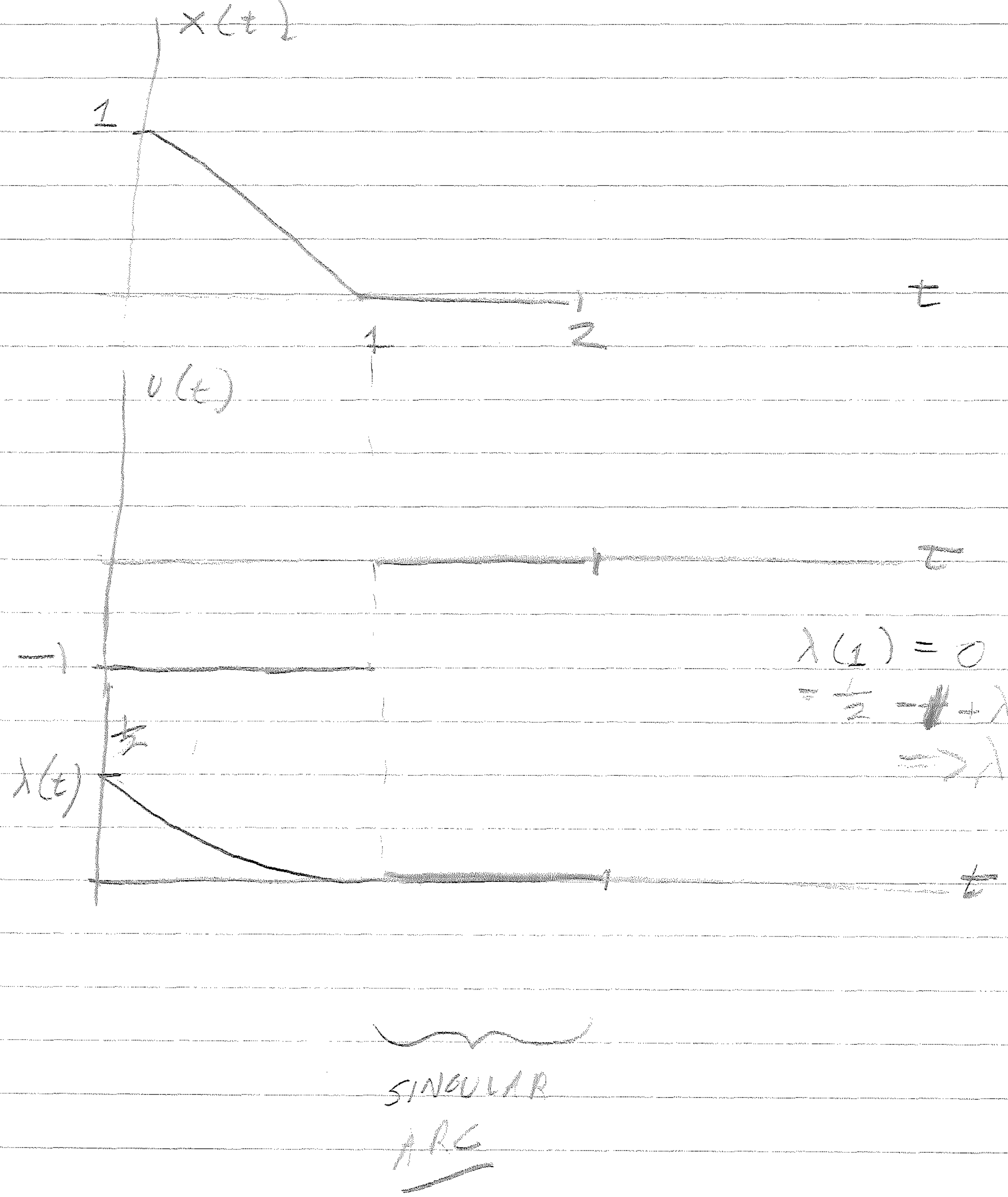
$\Longrightarrow \lambda(t) = \frac{1}{2}t^2 - t + \lambda(0)$

DURING SINGULARITY

$\lambda(t) = 0 \Longrightarrow \dot{\lambda}(t) = 0 \Longrightarrow \ddot{\lambda}(t) = 0$

$\Longrightarrow \dot{\lambda} = -x = 0$

$\ddot{\lambda} = 0 = -x = u$



$\lambda(1) = 0$

$= \frac{1}{2} - 1 + \lambda(0)$

$\Longrightarrow \lambda(0) = \frac{1}{2}$

$\overline{\Pi} =$ [illegible] B

$\mathcal{H}(x,\hat{u},t) \leq \mathcal{H}(x,$ [illegible]

SINGULAR PROBLEM

$$\begin{cases} J = \Theta[x(t_f), t_f] + \int_{t_0}^{t_f} (\phi + h' u)\, dt \\ \dot{x} = f(x,t) + G(x,t)\, u(t); \quad \mathcal{H} = \phi + h^T u + \lambda^T (f + Gu) \end{cases}$$

SPECIAL CASE IS THE LINEAR SYSTEM:

$\dot{x} = Ax + Bu$

$J = \frac{1}{2} x^T(t_f) S x(t_f)$

$\quad + \int_{t_0}^{t_f} x^T Q x \, dt$

THEN

$\mathcal{H} = x^T Q x^T + \lambda^T (Ax + Bu)$

THUS FOR OPTIMAL CONTRL

$\lambda^T B \hat{u} \leq \lambda^T B u$

$\Rightarrow \hat{u} = -\operatorname{sgn}(\lambda^T B)$

$\frac{\delta \mathcal{H}}{\delta u} = + B^T \lambda = 0$ FOR SINGU.

$\hat{u} = ?$

EXAMPLE CONSIDER

$J = \frac{1}{2} \int_0^{t_f} x_1^2 (dt)$ $\quad t_f$ IS FIXED ?

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} u(t)$$

$$= \begin{cases} x_1 + u & x_1(0) = x_{10}, \; x_1(t_f) = 0 \\ -u & x_2(0) = x_{20}, \; x_2(t_f) = 0 \end{cases}$$

HERE $\mathcal{H} = \frac{1}{2} x_1^2 + \underline{\lambda_1 \; \lambda_2} \begin{bmatrix} x_2 + u \\ -u \end{bmatrix}$

$= \frac{1}{2} x_1^2 + \lambda_1 (x + u) - \lambda_2 u$

$= \frac{1}{2} x_1^2 + \lambda_1 x_2 + (\lambda_1 - \lambda_2) u$

$\frac{\delta \mathcal{H}}{\delta u} = 0 = \lambda_1 - \lambda_2$

① $\begin{cases} \dot{\hat{x}}_1 = x_2 + u & x_1(0) = x_{10} \\ \dot{\hat{x}}_2 = -u & x_2(0) = x_{20} \end{cases}$

$$\begin{bmatrix} \dot{\lambda}_1 \\ \dot{\lambda}_2 \end{bmatrix} = \frac{-\delta \mathcal{H}}{\delta x_1} = - \begin{bmatrix} x_1 \\ \lambda_1 \end{bmatrix} \Rightarrow \begin{matrix} \dot{\lambda}_1 = -x_1 \\ \dot{\lambda}_2 = -\lambda_1 \end{matrix} \quad ②$$

$x_1(t_f) = 0$

$x_2(t_f) \equiv 0$

SINGULAR SOLUTION WHEN

$\lambda_1 = \lambda_2$ ③

$\Longrightarrow \dot{\lambda}_1 = \dot{\lambda}_2$ ④

$\ddot{\lambda}_1 = \ddot{\lambda}_2$ ⑤

FROM ④ & ③

$-\dot{x}_1 = \lambda_1$

FROM ⑤

$-\ddot{x}_1 = -\dot{\lambda}_1$

USING STATE EQUATIONS:

$\dot{x}_1 = x_2 + u = -\dot{x}_1$

THEN OPTIMAL CONTROL (CLOSED LOOP) ON THE SINGULAR ARC IS

$\hat{u} = -x_1 - x_2$

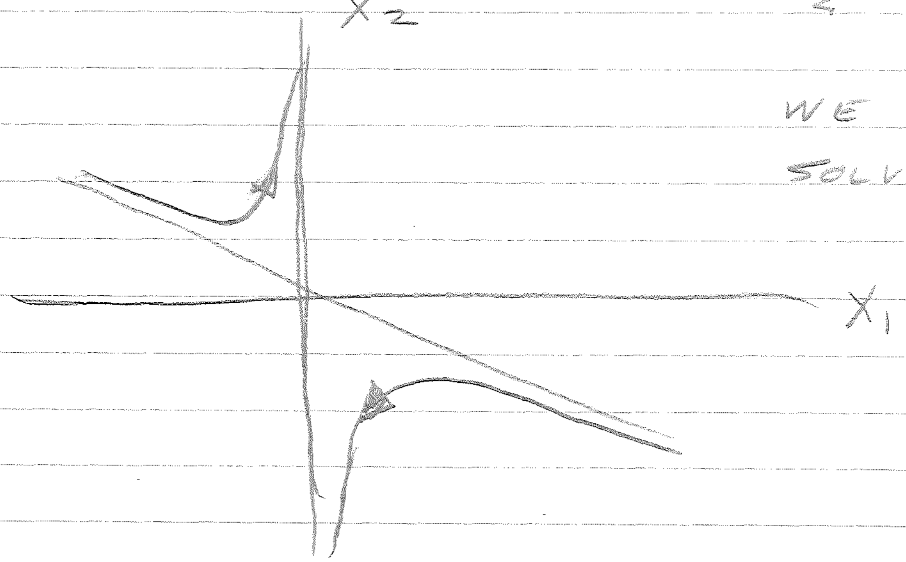
FINON ARC ON SINGULAR ARC

$\mathcal{H} = C = \text{CONSTANT}$

$\Rightarrow \frac{1}{2}x_1^2 + \lambda_1 x_2 = 0$

OR $\frac{1}{2}x_1^2 + x x_2 = 0$

THIS IS AN SINGULAR $\underline{\underline{\underline{ARC}}}$



WE ALSO GOTTA SOLVE $|u| \le K$

11-22-76

EX. $J = \frac{1}{2}\int_0^{t_f} x_1^2 \, dt$ $|U| \leq K$

$\dot{x}_1 = x_2 + U$ $x_1(t_0) = x_{10}$ $x_1(t_f) = 0$

$\dot{x}_2 = -U$ $x_2(t_0) = x_{20}$ $x_2(t_f) = 0$

$\mathcal{H} = \frac{1}{2}x_1^2 + \lambda_1 x_2 + (\lambda_1 - \lambda_2)U$

SINGULAR PROBLEM WHEN

$\delta H/\delta U = \lambda_1 - \lambda_2 = 0$

SINGULAR SOL. EXISTS : IF $\lambda_1 = \lambda_2$

$\lambda_1 = \lambda_2$

$\dot{\lambda} = -\delta H/\delta x \Rightarrow \begin{cases} \dot{\lambda}_1 = -x_1 \\ \dot{\lambda}_2 = -\lambda_1 \end{cases}$ ②

$\dot{\lambda}_1 = \dot{\lambda}_2 \Longrightarrow x_1 = \lambda_1$

$\dot{\lambda}_1 = \dot{\lambda}_2 \Longrightarrow \frac{d}{dt}\dot{\lambda}_1 = \frac{d}{dt}\dot{\lambda}_2$

$\frac{d}{dt}(-x_1) = \frac{d}{dt}(-\lambda_1)$

$+\dot{x}_1 = \dot{\lambda}_1$

$x_2 + U = \dot{\lambda}_1$

$x_2 + U = -x_1 \Rightarrow \underbrace{U = -(x_1 + x_2)}_{\text{FEEDBACK CONTROL}}$

FOR $\lambda_1 = \lambda_2$, $\mathcal{H} = \frac{1}{2}x_1^2 - \lambda_1 x_2$

$= \frac{1}{2}x_1^2 + X_1 X_2$

FOR $\mathcal{H}$ NOT EXPLICITLY A FUNC. OF TIME, $\mathcal{H} = C$

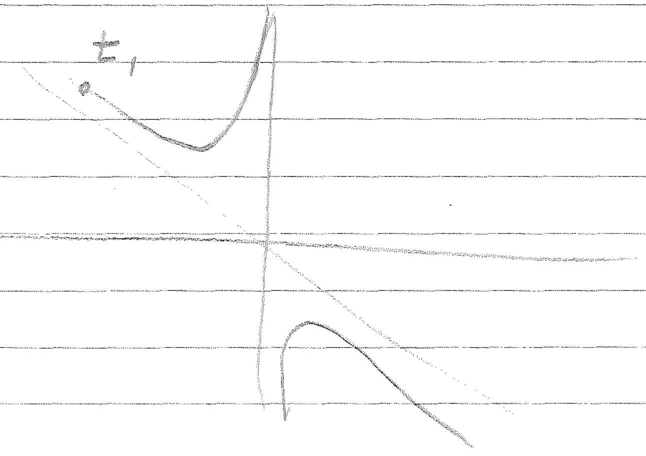
$\Rightarrow \mathcal{H} = \frac{1}{2}x_1^2 - \lambda_1 x_2 = C$ = CONSTANT

ON THE SINGULAR ARC:

$\dot{x}_1 = x_2 + U$

$= x_2 + [-(x_1 + x_2)] = -x_1$

⟹



THUS

$X(t) = X_1(t_1)e^{-(t-t_1)}$

SOLVE FOR X_2

$\dot{X}_2 = -U = X_1 + X_2$

$*$ $= X_1(t_1)e^{-(t-t_1)} + X_2$ ⟵ FORCING FUNCTION

$$X_2(t) = \underbrace{K_1 e^{+(t-t_1)}}_{\text{HOMO SOL}} + \underbrace{X_{2P}}_{\text{PARTICULAR SOL.}}$$

ASSSUME

$X_{2P} = K_2 e^{-(t-t_1)}$

SUBSTITUTING INTO $*$ GIVES

$-K_2 e^{-(t_1-t_1)} = X_1(t_1)e^{-(t-t_1)} + K_2 e^{-(t-t_1)}$

$-2K_2 = X_1(t_1) \Rightarrow K_2 = -\frac{1}{2}X_1(t_1)$

THUS $\Rightarrow$

$X_2(t) = K_1 e^{(t-t_1)} - \frac{1}{2}X_1(t_1)e^{-(t-t_1)}$

WHAT IS K_1?

@ $t = t_1$

$X_2(t_1) = K_1 - \frac{1}{2}X_1(t_1)$

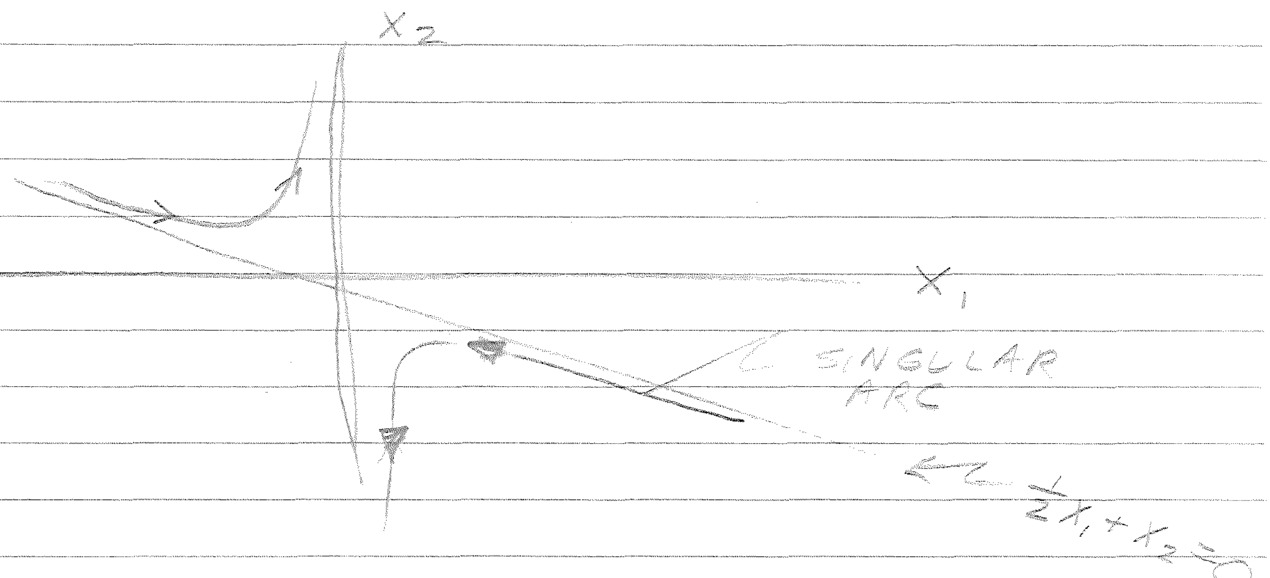
$\Rightarrow K_1 = X_2(t_1) + \frac{1}{2}X_1(t)$

THUS

$$x_2(t) = \left[x_2(t_1) + \frac{1}{2}x_1(t)\right]e^{(t-t_1)}$$

$$- \frac{1}{2}x_1(t)e^{-(t-t_1)}$$

LET'S PLOT THIS BURGER:



CONSIDER CASE WHEN $u = \pm K$

$$\Rightarrow \begin{cases} \dot{x}_1 = x_2 \pm K \\ \dot{x}_2 = \mp K \end{cases}$$

FOR $K = +1 \Rightarrow x_2 = -Kt + x_{20}$

$$\Rightarrow \dot{x}_1 = -Kt + x_{20} + K$$

$$\Rightarrow x_1(t) = -\frac{K}{2}t^2 + (x_{20} + K)t + x_{10}$$

OR

$$\begin{cases} x_1(t) = -\frac{K}{2}t^2 + (x_{20} + K)t + x_{10} \\ x_2(t) = -Kt + x_{20} \end{cases}$$

SOLVE FOR t IN x_2

$$\Rightarrow t = x_{20} - x_2 / K$$

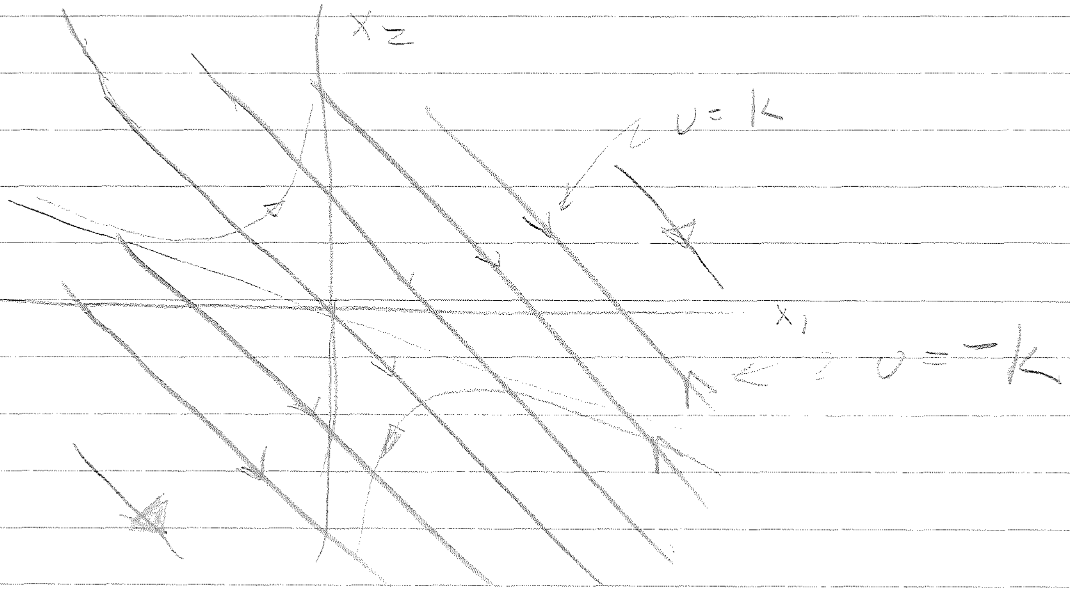
PLUG IN ~~FOR~~ x_1 EQ $\Longrightarrow$

$$x_1 = -\frac{k}{2}\left(\frac{x_{20}-x_2}{k}\right)^2 + \frac{(x_{20}+k)}{k}(x_{20}-x_2)$$

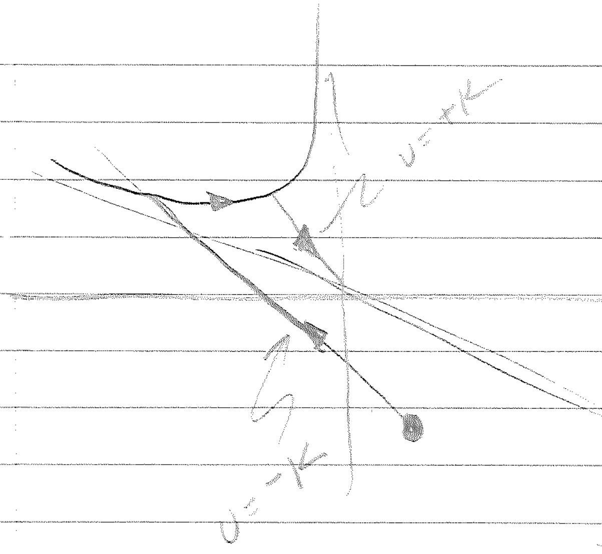
$$+ x_{10}$$

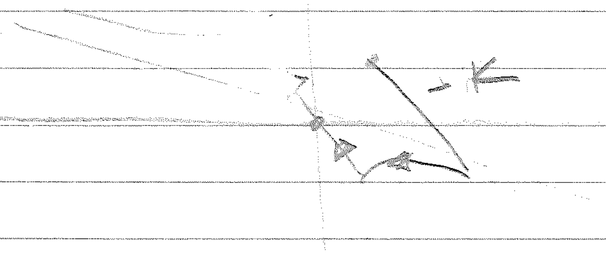
FOR BIG K $(k \Rightarrow \infty)$

$$x_1(t) = (x_{20}-x_2) + x_{10}$$
$$= -x_2 + (x_{10}+x_{20})$$



FOR BIG NEGATIVE K, THIS SAME SOLUTION IS STILL CORRECTED. NOTE DIRECTION DIFFERENCE THOU IN ABOVE GRAPH.



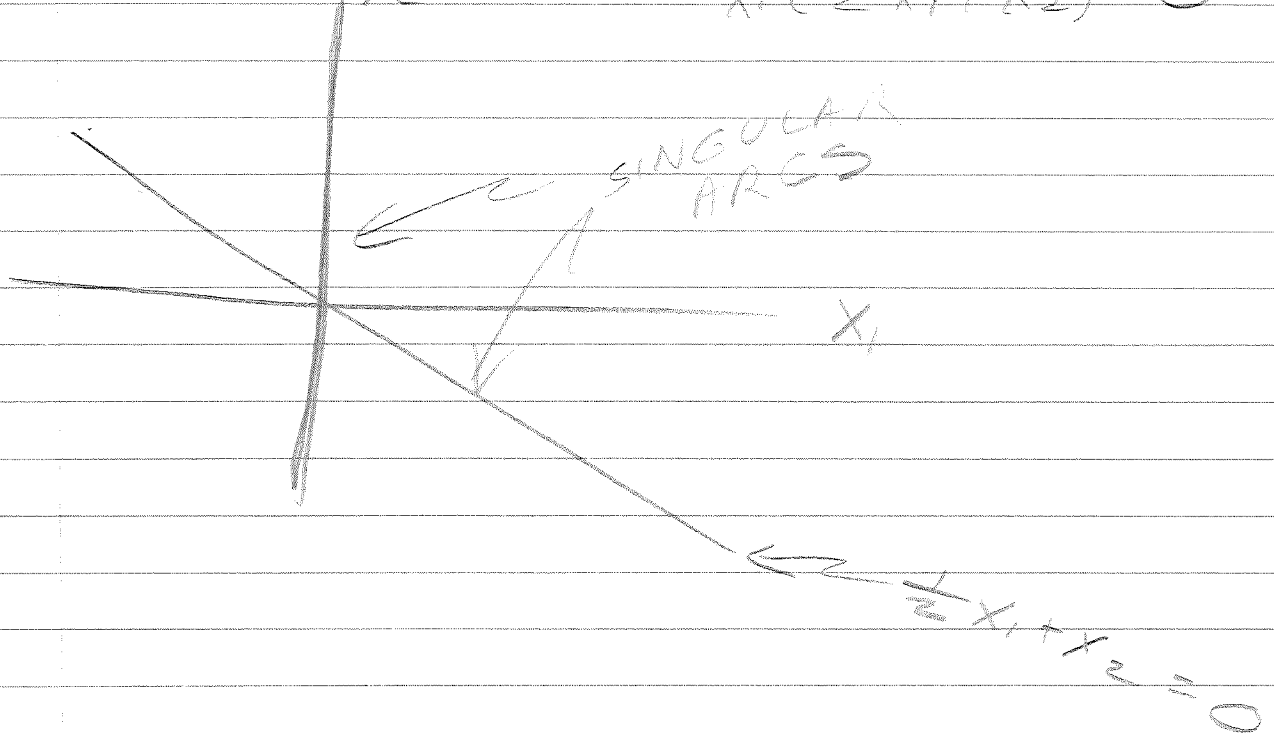


WE HAVE FIXED t_f. LET $t_f \to \infty$
TURNS OUT

$$\mathcal{H} = 0 = \underbrace{\cancel{\frac{\partial\phi}{\partial t_f}}}_{0} + \underbrace{\cancel{\frac{\partial\psi^T}{\partial t_f}}}_{0} \nu + \mathcal{H} = 0$$

$$t_f = \infty: \quad \mathcal{H} = 0 = \tfrac{1}{2}x_1^2 + x_1 x_2$$

$$\Longrightarrow x_1\left(\tfrac{1}{2}x_1 + x_2\right) = 0$$

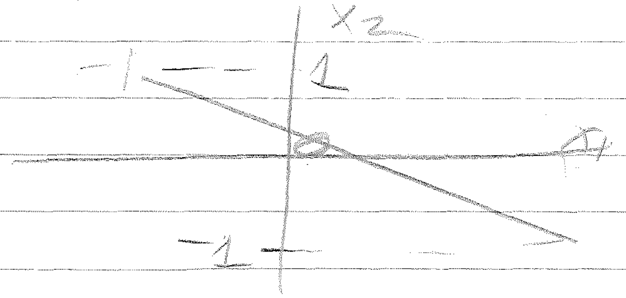


ON ARC
$u = -(x_1 + x_2)$

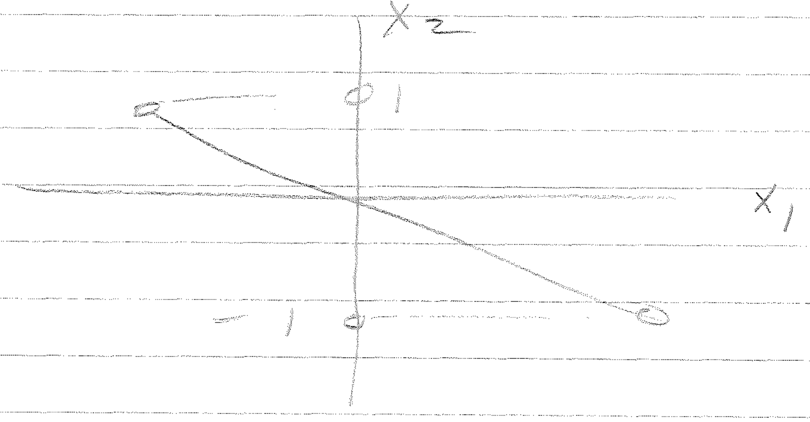
CONSIDER

(i) $x_1 = 0$
$\hat{u} = -x_2$
ASSUME $-1 \le u \le 1$
$\Rightarrow |u| \le 1$
$\Rightarrow |-x_2| \le 1$

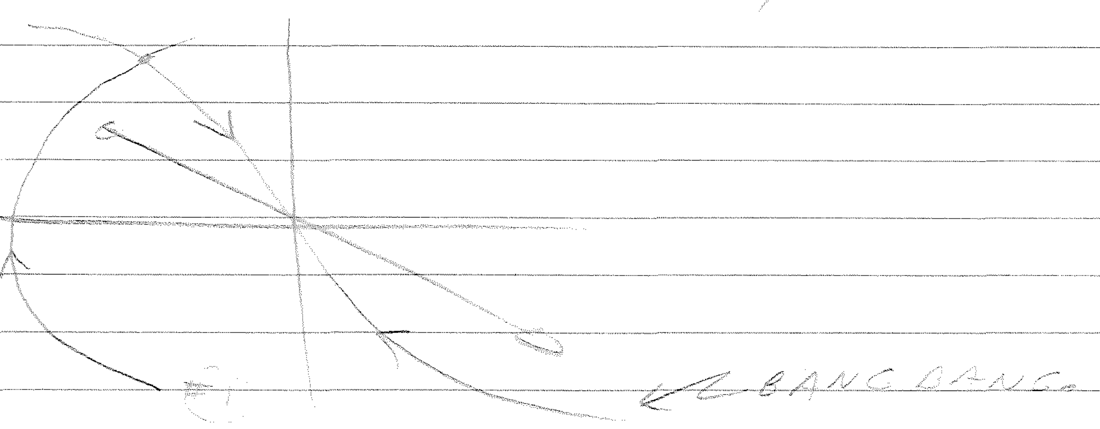
THUS SINGULAR ARC IS BALANCED



(ii) $\frac{1}{2} x_1 + x_2 = 0 \Rightarrow x_1 = -2x_2$
$\hat{u} = -x_1 - x_2 \Rightarrow u = 2x_2 - x_2 = x_2$
$|u| \le 1 \Rightarrow |x_2| \le |x_2| \le 1$



FOR BIG INITIAL CONDITIONS, JUST BANG BANG



(a) $\dot{x}_2 = -u = (x_1 - x_2) = U_2$

@ $x = x_1 = 0$, $\mu = x_2$

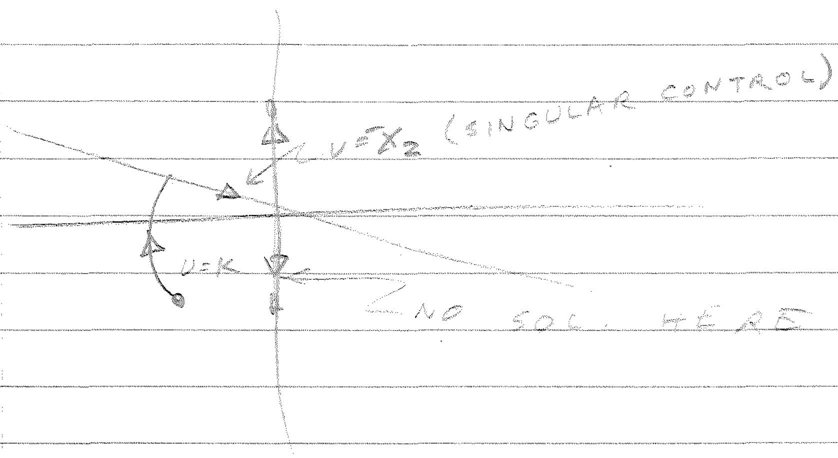
$\dot{x}_2 = x_2 \Rightarrow x_2(t) = x_2(t_1)\, e^{+(t-t_1)}$

(b) ON $x_1 = -2x_2$

$\dot{x} = -u$

$= (x_1 + x_2)$

$= -x_2 \Rightarrow x_2(t) = x(t_1)\, e^{-(t-t_1)}$



11-24-76

[illegible]

[illegible] $= 14$

[illegible] 1

[illegible] EATON

[illegible]

$\dot{x} = f(x, u, t)$ $x(t_0) = x_0$

$\dot{\lambda} = -\partial H/\partial x$ $\lambda(t_f) = \lambda_f$

~~[illegible]~~ $= 0$

[illegible]

[illegible] $X(t)$ [illegible]

[illegible] $X(t)$ [illegible]

[illegible] X_0 [illegible]

[illegible] HAVE [illegible]

$\dot{x}_f = a_{11}x + a_{12}\lambda$

$\dot{\lambda}_f = a_{21}x + a_{22}\lambda_f$

[illegible] λ_p

[illegible]

[illegible]

F[illegible] P[illegible] SO[illegible]

$\dot{x}_p =$ [illegible] $\lambda_p + 0$

$\dot{\lambda}_p = a_{21}x_p + a_{22}\lambda_p + 0$

$x_p(0) = x_0$

$\lambda(0) = 0$

[illegible] SOLUTION [illegible]

$x = Cx_h + x_p$

$y = Cy_h + y_p$

THUS

$x(0) = Cx_h(0) + x_p(0) = x_0$

$y(0) = Cy_h(0) + y_p(0) = 0$.

[illegible]

[illegible] $= Cy_h$ [illegible]

$C = \frac{y_0 + y_p(t)}{y_h(t)}$

EXAMPLE $\dot{x} = -2x - y + 6$ $x(0) = 3$

$\dot{y} = 4x + 3y$ $y(0) = 0$

$\dot{x}_h = -2x_h - y_h$ $x_h(0) = 0$

$\dot{y}_h = 4x_h + 3y_h$ $y_h(0) = 1$

$\ddot{x}_h = -2\dot{x}_h - \dot{y}_h = -2\dot{x}_h - 4x_h - 3y_h$

$= -2\dot{x}_h - 4x_h - 3(-2x_h - \dot{x}_h)$

$\ddot{x}_h = \dot{x}_h + 2x_h$

$\ddot{x}_h - \dot{x}_h - 2x_h = 0$

[illegible]

$x_h = K_1e^{-t} + K_2e^{2t}$

[illegible]

$\dot{x}_h(0) = -2x_h(0)$ [illegible]

$= -2($ [illegible] $) - y_h(0) =$ [illegible]

$2(K_1 + K_2) = -1$ [illegible]

x_h' [illegible]

[illegible]

x_h [illegible] e^{2t}

$\lambda_h = -2x_h - \dot{x}_h$

$= -\frac{?}{3}e^{-t} + \frac{2}{3}e^{2t}$

$- \frac{1}{3}(e^{-t} - \frac{2}{3}e^{2t})$

$= -\frac{1}{2}e^{-t} + \frac{4}{3}e^{2t}$

Now for particular type

$\dot{x} = -2x - \lambda + 6$ $\quad x(0) = 3$

$\dot{\lambda} = -4x + 3\lambda$ $\quad \lambda(0) = 0$

$\dot{x}_p = -2x_p - \lambda_p + 6$ $\quad x_p(0) = 6$

$\dot{\lambda}_p = -4x_p + 3\lambda_p$ $\quad \lambda_p(0) = 0$

$x_p = c_1 e^{-t} + c_2 e^{2t}) + A$

$\lambda_p = c_1 e^{-t}$ [illegible] $e^{2t}) + B$

$0 = -2A - B + 6 \Rightarrow$ [illegible]

$0 = 4A + 3B$ [illegible]

TURNS OUT A-SO THAT

$c_1 = -4$, $c_2 = 2$

[illegible]

$x = c($ [illegible] $) - 4e$ [illegible] $2e$ [illegible]

[illegible] $(4e$ [illegible]

$y = At + B$

[illegible]

[illegible] $) = -193$

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] $y = 45$ [illegible]

11-29-76

QUASILINEARIZATION

$\dot{x} = a_{11}x + a_{12}\lambda + f_1 \qquad x(0) = x_0$

$\dot{\lambda} = a_{21}x + a_{22}\lambda + f_2 \qquad \lambda(t_f) = \lambda_f$

THESE EQUATIONS FROM

$$\left[\begin{array}{l} \dot{x} = f(x, u, t) \; ; \; x(0) = x_0 \\ \dot{\lambda} = -\delta H/\delta x \quad \text{PLUG} \\ \delta H/\delta u = 0 \quad \text{PLUG} \\ \lambda(t_f) = \delta\theta/\delta x + \left(\frac{\delta N^T}{\delta x}\right) \cdot \nu \end{array}\right.$$

NOTE SPLIT B.C.'S

FIND HOMO SOLN:

$$\begin{bmatrix} x_h \\ \lambda_h \end{bmatrix} : \begin{bmatrix} x_h(0) \\ \lambda_h(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

PART. SOLN:

$$\begin{bmatrix} x_p \\ \lambda_p \end{bmatrix} = \begin{bmatrix} x_0 \\ 0 \end{bmatrix}$$

GENERAL SOLUTION IS

$$\begin{bmatrix} x \\ \lambda \end{bmatrix} = C\begin{bmatrix} x_h \\ \lambda_h \end{bmatrix} + \begin{bmatrix} x_p \\ \lambda_p \end{bmatrix}$$

@ $t = 0$

$$\begin{bmatrix} x(0) \\ \lambda(0) \end{bmatrix} = C\begin{bmatrix} x_h(0) \\ \lambda_h(0) \end{bmatrix} + \begin{bmatrix} x_p(0) \\ \lambda_p(0) \end{bmatrix}$$

$$= C\begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} x_0 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} x_0 \\ C \end{bmatrix}$$

$\lambda(t) = C\lambda_h + \lambda_p \Rightarrow \lambda(t_f) = C\lambda_h(t_f) + \lambda_p(t_f)$

$$C = \frac{\lambda(t_f) - \lambda_p(t_f)}{\lambda_h(t_f)}$$

LINEARIZATION

$$\left.\begin{aligned}\dot{x} &= f(x,\lambda,t)\\ \dot{\lambda} &= g[x,\lambda,t\end{aligned}\right\} \qquad \begin{aligned}x(0)&=x_0\\ \lambda(t_f)&=\lambda_f\end{aligned}$$

LET $x^0(t), \lambda^0(t)$ BE A KNOWN SOLUTION (INITIAL GUESS)

$$\dot{x}^{(1)}(t) = f[x^0, \lambda^0, t] + \underbrace{\left.\frac{\delta f}{\delta x}\right|_{\substack{x^0\\ \lambda^0}}}_{\to \text{USE JACOBIAN}} \left[x^{(1)} - x^{(0)}\right] + \left.\frac{\delta f}{\delta \lambda}\right|_{\substack{x^0\\ \lambda^0}} \left[\lambda^{(1)} - \lambda^{(0)}\right] + \text{H.O.T.}$$

$$\dot{\lambda}^{(1)}(t) = g(x^0, \lambda^0, t) + \left.\frac{\delta g}{\delta x}\right|_{\substack{x^0\\ \lambda^0}} \left(x^{(1)} - x^{(0)}\right) + \left.\frac{\delta g}{\delta \lambda}\right|_{\substack{x^0\\ \lambda^0}} \left(\lambda^{(1)} - \lambda^{(0)}\right) + \text{H.O.T.}$$

OR

$$\dot{x}^{(1)}(t) = \overbrace{\left.\frac{\delta f}{\delta x}\right|_{\substack{x^0\\ \lambda^0}}}^{a_1(t)} x^{(1)} + \overbrace{\left.\frac{\delta f}{\delta \lambda}\right|_{\substack{x^0\\ \lambda_0}}}^{a_2} \lambda^{(1)} + \overbrace{\left(f(x^0, \lambda^0, t) - \left.\frac{\delta f}{\delta x}\right|_{\substack{x^0\\ \lambda^0}} x^0 - \left.\frac{\delta f}{\delta \lambda}\right|_{\substack{x^0\\ \lambda_0}} \lambda^0\right)}^{e_1(t)}$$

SAME FOR $\lambda^{(1)}$

USING $x^{(1)}$ & $\lambda^{(1)}$, SOLVE FOR $x^{(2)}$ $\lambda^{(2)}$ ETC. COMPARING NORMS TWIXT ~~NEW~~ TWO SOLN., WE CAN PLACE A STOP ON THE ITTERATION SCHEME

$$\left\| \begin{bmatrix} x^{(k+1)} \\ \lambda^{(k+1)} \end{bmatrix} - \begin{bmatrix} x^{(k)} \\ \lambda^{(k)} \end{bmatrix} \right\| \le \epsilon$$

Ex

$$\dot{x} = x^2 + u \qquad x(0) = 3$$

$$J = \int_0^1 (2x^2 + u^2)\,dt$$

$$\mathcal{H} = 2x^2 + u^2 + \lambda(x^2 + u)$$

$$\dot{\lambda} = -\frac{\delta \mathcal{H}}{\delta x} = -4x - 2\lambda x$$

$$\frac{\delta \mathcal{H}}{\delta u} = 0 = 2u + \lambda = 0 \Longrightarrow u = -\lambda/2$$

THUS

$$\dot{x} = x^2 - \lambda/2 \quad ; \quad x(0) = 3$$

$$\dot{\lambda} = -4x - 2\lambda x \quad ; \quad \lambda(t_f) = 0 \quad \left(\text{SINCE } N = 0\right)$$

NEXT STEP IS LINEARIZATION.

$$\dot{x}^{(1)} \cong \left({x^0}^2 - \frac{\lambda^0}{2}\right) + (2x^0)\left(x^{(1)} - x^{(0)}\right) + \left(-\frac{1}{2}\right)\left(\lambda^{(1)} - \lambda^{(0)}\right)$$

$$\dot{\lambda}^{(1)} \cong \left(-4x^0 - 2\lambda^0 x^0\right) + (-4)\left(x^{(1)} - x^{(0)}\right) + (-2x^0)\left(\lambda^{(1)} - \lambda^0\right)$$

OR

$$\dot{x}^{(1)} = \overbrace{2x^{(0)}}^{a_{11}} x^{(1)} \overbrace{- \tfrac{1}{2}}^{a_{12}} \lambda^{(1)} \overbrace{- (x^0)^2}^{e_1}$$

$$\dot{\lambda}^{(1)} = \underbrace{(-4 - 2\lambda^0)}_{a_{21}} x^{(1)} \underbrace{- 2x^0}_{a_{22}} \lambda^{(1)} + \underbrace{(2\lambda^0 x^0)}_{e_2}$$

LET'S LOOK @ GENERAL VECTOR CASE:

$$\dot{x}^{(i+1)} = f\Big|_{x^i,\lambda^i} + \frac{\delta f}{\delta x}\Big|_{x^i,\lambda^i}\left(x^{i+1} - x^i\right) + \frac{\delta f}{\delta \lambda}\Big|_{x^i,\lambda^i}\left(\lambda^{i+1} - \lambda^i\right)$$

$$\dot{\lambda}^{(i+1)} \cong \left(-\delta H/\delta x\right)\Big|_{x^i,\lambda^i} + \left(-\frac{\delta^2 H}{\delta x^2}\right)\Big|_{x^i,\lambda^i}\left(x^{i+1} - x^i\right)$$

$$+ \left(-\frac{\delta^2 H}{\delta \lambda \delta x}\right)\Big|_{x^i,\lambda^i}\left(\lambda^{i+1} - \lambda^i\right)$$

THUS

$$\dot{Y} = \begin{bmatrix}\dot{x}\\ \dot{\lambda}\end{bmatrix}^{i+1} = \left[\begin{array}{c|c} & \\ \hline & \end{array}\right] \underbrace{\begin{bmatrix}x\\ \lambda\end{bmatrix}^{i+1}}_{Y} + \begin{bmatrix}e_1\\ e_2\end{bmatrix}^{i}$$

QUESTION IS: HOW WE SOLVE?

$$\dot{Y}^{i+1} = A^i(t)\,Y^{i+1} + E^i(t) \leftarrow 2n \text{ EQ'S}$$

$$\left.\begin{cases} Y^{i+1} = Y_h + Y_p \\ Y_h(0) = \begin{bmatrix}0\\ 1\end{bmatrix} \quad Y_p(0) = \begin{bmatrix}x_0\\ 0\end{bmatrix}\end{cases}\right\} \text{SCALAR CASE}$$

FOR VECTOR CASE:

$$Y = \begin{bmatrix}x\\ \lambda\end{bmatrix} = \begin{bmatrix}x_1\\ \vdots\\ x_n\\ \lambda_1\\ \vdots\\ \lambda_n\end{bmatrix} \Rightarrow Y_{h_1} = \begin{bmatrix}0\\ 0\\ 0\\ \vdots\\ 0\\ \hline 1\\ 0\\ 0\\ 0\\ \vdots\\ 0\end{bmatrix} \quad Y_{h_2} = \begin{bmatrix}0\\ 0\\ \vdots\\ 0\\ 0\\ 1\\ 0\\ 0\\ 0\\ \vdots\\ 0\end{bmatrix}$$

$$Y_{h_n} = \begin{bmatrix}0\\ 0\\ \vdots\\ 0\\ 0\\ 0\\ \vdots\\ 1\end{bmatrix}$$

AND HOM. SOLN' IN GENERAL CAN BE WRITTEN

$$c_1 Y_{h_1} + c_2 Y_{h_2} + \dots + c_n Y_{h_n}$$

12/11/76

$$\begin{cases} \dot{x} = f(x, u, t) & x(t_0) = x_0 \\ \dot{\lambda} = -\delta H / \delta x \\ \delta H / \delta u = 0 \end{cases}$$

$$\begin{cases} \dot{x} = f(x, \lambda, t) & x(0) = x \\ \dot{\lambda} = g(x, \lambda, t) & \lambda(t_f) = \lambda_f = \frac{\delta \Theta}{\delta x} \end{cases}$$

THEN

$$\begin{cases} \dot{x}^{i+1} \cong f(x^i, \lambda^i, t) + \frac{\delta f}{\delta x}\Big|_{x^i, \lambda^i} (x^{i+1} - x^i) \\ \qquad + \frac{\delta f}{\delta \lambda}\Big|_{x^i, \lambda^i} (\lambda^{i+1} - \lambda^i) \\ \dot{\lambda}^{i+1} \cong g(x^i, \lambda^i, t) \end{cases}$$

$$+ \frac{\delta g}{\delta x}\Big|_{\lambda^i, x^i} (x^{i+1} - x^i)$$

$$+ \frac{\delta g}{\delta \lambda}\Big|_{\lambda^i, x^i} (\lambda^{i+1} - \lambda^i)$$

THESE ARE IN THE FORM

$$\begin{bmatrix} \dot{x}^{i+1} \\ \dot{\lambda}^{i+1} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x^{i+1} \\ \lambda^{i+1} \end{bmatrix} + \begin{bmatrix} E \end{bmatrix}$$

IN SCALAR CASE

$$\begin{bmatrix} \dot{x} \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$\begin{bmatrix} x_h(0) \\ \lambda_h(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x_p(0) \\ \lambda_p(0) \end{bmatrix} = \begin{bmatrix} x_0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ \lambda \end{bmatrix} = c \begin{bmatrix} x_h \\ \lambda_h \end{bmatrix} + \begin{bmatrix} x_p \\ \lambda_p \end{bmatrix}$$

$$C = \left[\lambda(t_f) - \lambda_p(t_f)\right] / \lambda_h(t_f)$$

CONSIDER THE MORE GENERAL VECTOR CASE AGAIN

$$\begin{bmatrix} x_h \\ \lambda_h \end{bmatrix}^{\cdot} = \begin{bmatrix} \dot{x}_h \\ \dot{\lambda}_h \end{bmatrix} = A \begin{bmatrix} x_h \\ \lambda_h \end{bmatrix}$$

SUCH THAT

$$\begin{bmatrix} x_{h_1}(0) \\ \lambda_{h_1}(0) \end{bmatrix} = \left[\begin{array}{c} x_{1h_1}(0) \\ x_{2h_1}(0) \\ \vdots \\ x_{nh_1}(0) \\ \hline \lambda_{1h}(0) \\ \lambda_{2h}(0) \\ \vdots \\ \lambda_{nh}(0) \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \\ \hline 1 \\ 0 \\ \vdots \\ 0 \end{array}\right]$$

$$\begin{bmatrix} x_{h_2}(0) \\ \lambda_{h_2}(0) \end{bmatrix} = \left[\begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \\ \hline 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{array}\right]$$

$$\vdots$$

$$\Rightarrow \begin{bmatrix} x_{h_n}(0) \\ \lambda_{h_n}(0) \end{bmatrix} = \left[\begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \\ \hline 0 \\ \vdots \\ 0 \\ 1 \end{array}\right]$$

ON TO PARTICULAR SOLUTION:

$$x_p : \begin{bmatrix} \dot{x}_p \\ \dot{\lambda}_p \end{bmatrix} = A \begin{bmatrix} x_p \\ \lambda_p \end{bmatrix} + E$$

$$\begin{bmatrix} x_p(0) \\ \lambda_p(0) \end{bmatrix} = \left[\begin{array}{c} x^0 \\ \hline 0 \end{array}\right]$$

THUS

$$\begin{bmatrix} x \\ \lambda \end{bmatrix} = c_1 \begin{bmatrix} x_{h1} \\ \lambda_{h1} \end{bmatrix} + c_2 \begin{bmatrix} x_{h2} \\ \lambda_{h2} \end{bmatrix}$$

$$+ \cdots + c_n \begin{bmatrix} x_{hn} \\ \lambda_{hn} \end{bmatrix} + \begin{bmatrix} x_p \\ \lambda_p \end{bmatrix}$$

CHECK B.C.

$$\begin{bmatrix} x(0) \\ \lambda(0) \end{bmatrix} = c_1 \left[\begin{array}{c} 0 \\ \vdots \\ 0 \\ \hline 1 \\ 0 \\ \vdots \\ 0 \end{array}\right] + c_2 \left[\begin{array}{c} 0 \\ \vdots \\ 0 \\ \hline 0 \\ 1 \\ \vdots \\ 0 \end{array}\right] + \cdots + \left[\begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \\ \hline 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{array}\right] c_n$$

$$= \begin{bmatrix} x_0 \\ c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \longleftarrow \text{OUR INTIAL TIME BOUNDARY CONDITIONS.}$$

HOW 'BOUT FINAL TIME:

$$\lambda(t_f) = c_1\lambda_{h_1}(t_f) + c_2\lambda_{h_2}(t_f) + \dots + c_n\lambda_{h_n}(t_f) + \lambda_p(t_f)$$

ONLY UNKNOWNS ARE $c_1 - c_n$

$$\lambda(t_f) = \underbrace{\left[\underset{\downarrow}{\lambda_{h_1}(t_f)}\ \underset{\underset{n\text{ VECTOR}}{\downarrow}}{\lambda_{h_2}(t_f)} \dots \underset{\downarrow}{\lambda_{h_n}(t_f)}\right]}_{\substack{\text{SQUARE MATRIX}\\ \text{CALL "}\Lambda_n\text{"}}} \underset{\underset{c}{\triangleright}}{\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}} + \lambda_p(t_f)$$

$$C = \Lambda_n^{-1}\left[\lambda(t_f) - \lambda_p(t_f)\right]$$

AND OUR SOLUTION IS COMPLETE.

TEST #1 SOLUTIONS

1(a). $J = \int_0^1 \phi(x, \dot{x}, t)\, dt$

$\ddot{x} + \alpha(x + \varepsilon x^3) = 0$

$\frac{\delta\phi}{\delta x} - \frac{d}{dt}\frac{\delta\phi}{\delta \dot{x}} = 0$

SOLUTION IS OBVIOUSLY:

$\phi = -\frac{1}{2}\dot{x}^2 + \alpha\left(\frac{x^2}{2} + \frac{\varepsilon}{2}x^4\right)$

b. $x(0) = x(1) = 0$

$\ddot{x} + \alpha x = 0$

$A = A\cos\sqrt{\alpha}\, t + B\sin\sqrt{\alpha}\, t$

$\lambda(0) = 0 = A$

$\Rightarrow x = B\sin\sqrt{\alpha}\, t$

$x(1) = B\sin\sqrt{\alpha} \Rightarrow \sqrt{\alpha} = n\pi$

84

SATISFY E FUNCTION:

$$E = \Phi(x, \dot{X}, t) - \Phi(x, \dot{x}, t) - (\dot{X} - \dot{x}) \frac{\delta \Phi}{\delta \dot{x}} \geq 0$$

OPTIMAL

$$\Rightarrow \Phi = \tfrac{1}{2}\dot{x}^2 + \alpha x^2/2$$

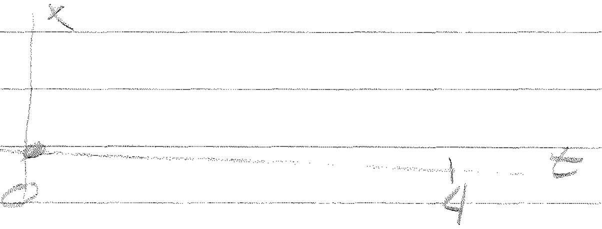
THUS, AFTER SOME PLUG

$$E = \tfrac{1}{2}\left(\dot{X} - \cancel{B}\alpha \cos\sqrt{\alpha}\, t\right)^2$$

$$> 0$$

2. $J = \int_0^4 (\dot{x}-1)^2(\dot{x}+1)^2 dt \Rightarrow \dot{x} = \pm 1$

$x(0) = 0 \quad x(4) = 2$



TURNS OUT

$$\delta^2\phi/\delta\dot{x}^2 \; \ddot{x} = 0$$

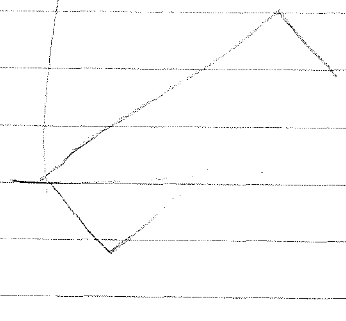
$$\Rightarrow \ddot{x} = 0 \Rightarrow x = c_1 t + c_2$$

RECALL CORNER CONDITIONS

$$\left.\frac{\delta\phi}{\delta\dot{x}}\right|_{t_-} = \left.\frac{\delta\phi}{\delta\dot{x}}\right|_{t_+}$$

$$\left.\phi - \dot{x}^T \frac{\delta\phi}{\delta\dot{x}}\right|_{t_-} = \left.\phi - \dot{x}^T \frac{\delta\phi}{\delta\dot{x}}\right|_{t_+}$$

$$\Rightarrow \dot{x} = \pm 1$$



[illegible]

$$-\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

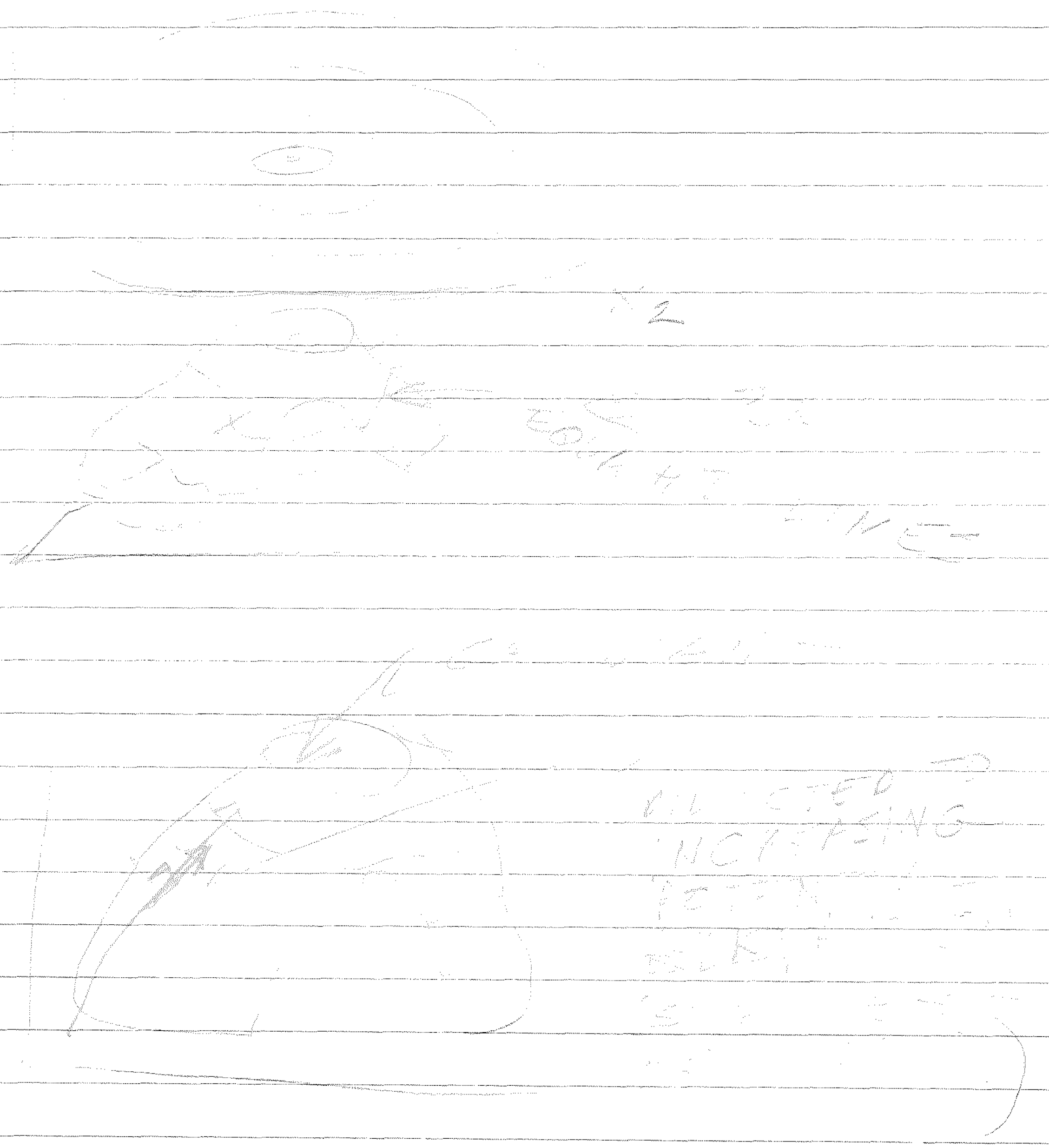
$-2p_2 + 3p$ [illegible]

[illegible] $= +2$

$p_2 = -2$ $p_1 = -2$ [illegible] $= -2$

$p_{22} = 4$ $p_2 = 2,$ $p_1 = 2$

[illegible]


2

87

STEEPEST DESCENT $= -\nabla f$

$\nabla = \delta f$

$\nabla f = \begin{bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \end{bmatrix}$

$\Xi = -\nabla f$ VECTOR

$\Xi = \frac{\partial f}{\partial x}\Big|_{x_0} \Big[\frac{\partial f}{\partial x_1}\Big|_{x_0} \quad \frac{\partial f}{\partial x_2}\Big|_{x_0} \Big]$

$\Delta X = X^{(1)} - X^{(0)} = \alpha\,\Xi$

$X^{(1)} = X^{(0)} - \alpha \nabla f\big|_{x_0}$

$f = f(x_0) + \frac{\partial f}{\partial x_1}\Big|_{x_0} \delta x_1 + \frac{\partial f}{\partial x_2}\Big|_{x_0} \delta x_2$

$\Delta f = \left(\frac{\partial f}{\partial x}\right)^T \Delta x$

$= -\alpha \left(\frac{\partial f}{\partial x}\right)^T_{x_0} \Xi\,(x^0)$

$$- \left[\frac{\delta t}{\delta t} \quad \frac{\delta t}{\epsilon_2} \quad \frac{\epsilon}{\epsilon} \right]$$

$$\sqrt{\frac{\delta t}{\epsilon_1} \quad \frac{\epsilon}{\epsilon_2}}$$

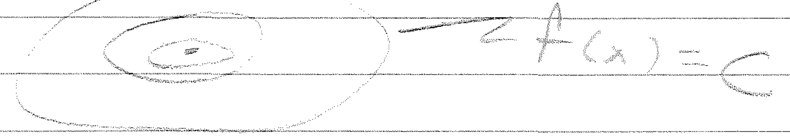
$= -\lambda$ [illegible] $\sqrt{\frac{\delta t}{\epsilon_1} \quad \frac{\epsilon}{\epsilon_2}}$

$= -\lambda \Big/ \frac{\epsilon}{\epsilon} \quad \frac{\epsilon}{\epsilon_2}$ [illegible]

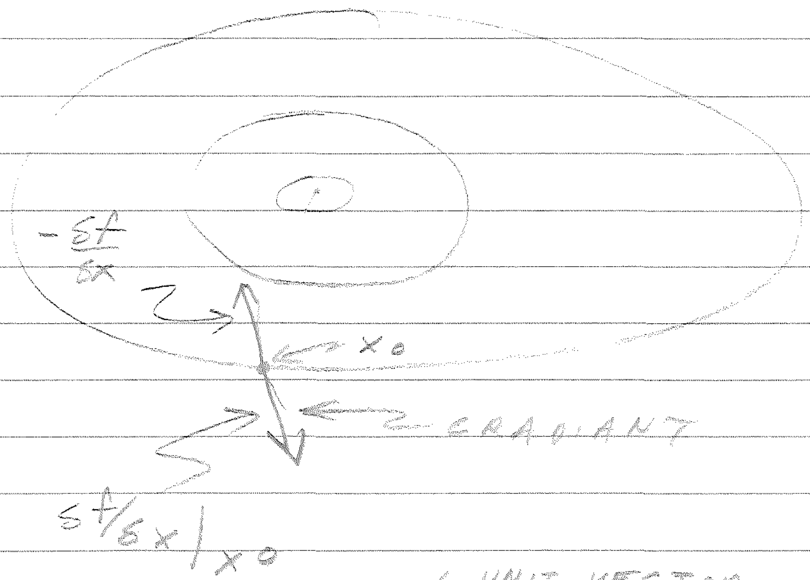
$x_1 = x_2 = U \begin{bmatrix} \\ \beta_2 \end{bmatrix}$

12/6/76 (MON)

STEEPEST DESCENT METHOD



$$\frac{\delta f}{\delta x} = 0$$



THEN

$$x^{i+1} = x^i - \tau \underline{z}(x^i)$$

UNIT VECTOR IN DIRECTION OF $\delta f/\delta x|_{x^i}$

CONSIDER

$$\dot{x} = f(x, u, t) \quad , \quad x(t_0) = x_0$$

$$J = \Theta[x(t_f), t_f] + \int_{t_0}^{t_f} \phi(x, u, t)\, dt$$

$$= \int_{t_0}^{t_f} \frac{d}{dt} \Theta\, dt + \Theta(x(t_0), t_0)$$

$$+ \int \phi\, dt$$

$$J' = \int_{t_0}^{t_f} \left(\frac{d}{dt} \Theta\, dt + \phi \right) dt$$

$$J_a = \int_{t_0}^{t_f} \left[\phi + \frac{\delta \Theta}{\delta x} \dot{x} + \frac{\delta \Theta}{\delta t} + \lambda^T (f - \dot{x}) \right] dt$$

$$\delta J_a = \left[\frac{\delta \Theta}{\delta x}\bigg|_{x(t_f)} - \lambda(t_f) \right]^T \delta x(t_f)$$

$$+ \int_{t_0}^{t_f} \left[\left\{ \dot{\lambda} + \frac{\delta \mathcal{H}}{\delta x} \right\}^T \delta x(t) \right.$$

$$+ \left(\frac{\delta \mathcal{H}}{\delta u} \right)^T \delta u(t)$$

$$\left. + (f - \dot{x})^T \delta \lambda \right] dt = 0$$

THUS

$$\begin{cases} ① & \lambda(t_f) = \frac{\delta \Theta}{\delta x}\bigg|_{x(t_f)} \\ ② & \dot{\lambda} = -\delta \mathcal{H} / \delta x \\ ③ & \delta \mathcal{H} / \delta u = 0 \\ ④ & \dot{x} = f \qquad ; \; x(t_0) = x_0 \end{cases}$$

ASSUME ①, ② AND ④ ARE SATISFIED ON i^{TH} ITERATION, BUT ③ IS <u>NOT</u>

$$\delta J_a = \int_{t_0}^{t_f} \frac{\delta \mathcal{H}^T(x, u, \lambda, t)}{\delta u} \delta u \, dt$$

$$\approx \Delta J_a$$

WE KNOW x^i, u^i, λ^i

CONSIDER

$$\delta u = U^{(i+1)}(t) - U^i(t) \implies$$

$-\delta u = \gamma\left(-{}^{\delta H}/_{\delta u}\right) \Longleftarrow$ PICK THIS

THEN

$$\Delta J_a \cong \delta J_a = -\gamma \int_{t_0}^{t_f} \left(\frac{\delta H}{\delta u}\right)^T \left(\frac{\delta H}{\delta u}\right) dt$$

$$\leq 0$$

EQUALITY ONLY IF ${}^{\delta H}/_{\delta u} = 0$

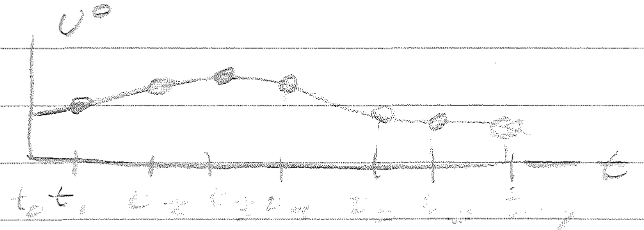
THIS GIVES u_{i+1}

SOLVE STATE & COSTATE

& REPEAT.

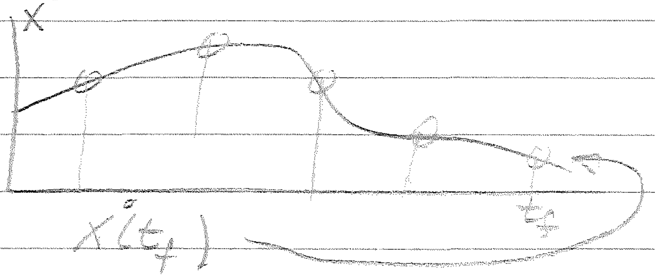
SOLVING THE WHOLE BAG

(i) SELECT $u^0(t)$



(ii) COPUTE x^0 FROM

$\dot{x}^0 = f(x^0, u^0, t) \ni x^0(t_0) = x_0$



(iii) FORM $x^0(t_f)$

CALCULATE

$$\lambda^0(t_f) = \frac{\delta\theta}{\delta x}\bigg|_{x^0(t_f)} \triangleq \lambda_f^0$$

(iv) INTEGRATE COSTATE EQ.

$$\dot{\lambda}^0 = -{}^{\delta H}/_{\delta x^0} \;\&\; \lambda^0(t_f) = \lambda_f^0$$

(INTEGRATE BASS ACKWARDS)

(v) EVALUATE

$\frac{\delta H^0}{\delta U}$

COMPUTE $\left\|\frac{\delta H^0}{\delta U}\right\|^2 \triangleq \int_{t_0}^{t_f} \left(\frac{\delta H^0}{\delta U}\right)^T \left(\frac{\delta H^0}{\delta U}\right) dt$

$\le \gamma^2$?

IS $\|\cdot\|^2 \le \gamma^2$,
WE'RE DONE,
IF IT AIN'T

(vi) CALCULATE

$$U^1 = U^0 - \tau \left(\frac{\delta H}{\delta U}\right)^0$$

START OVER AGAIN @ (i)

EXAMPLE

$$\dot{x} = -x + U \quad ; \quad x(0) = 4$$

$$J = x^2(1) + \frac{1}{2}\int_0^1 U^2 \, dt \quad ; \quad t_f = \underline{1}$$

$$H = \frac{1}{2}U^2 + \lambda(-x + 0)$$

COSTATE

$$\dot{\lambda} = -\delta H/\delta x = -(-\lambda) = \lambda$$

$$\frac{\delta H}{\delta U} = U + \lambda = 0$$

$$\lambda(t_f) = \delta\phi/\delta x \Big|_{t_f} = 2x(t_f)$$

INITIAL GUESS:

(i) PICK $u^0(t) = 1$

(ii) $\dot{x}^0 = -x^0 + u^0$; $x(0) = 4$

$= -x^0 + 1$

$\Longrightarrow x^0(t) = 3e^{-t} + 1$

(iii) $\lambda^0(t_f) = 2x(t_f) = 2(3e^{-1} + 1) = \lambda^0(1)$

(iv) $\dot{\lambda}^0 = \lambda^0$; $\lambda^0(1) = 2(3e^{-1} + 1)$

$\Rightarrow \lambda^0(t) = (\quad K \quad) e^t$

NOW $\lambda^0(1) = Ke = 2(3e^{-1} + 1)$

$\Rightarrow K = 2e^{-1}(3e^{-1} + 1)$

$\therefore \lambda^0(t) = 2e^{-1}(3e^{-1} + 1)e^t$

(v) $\left(\delta H / \delta u\right)^0 = u^0 + \lambda$

$= 1 + 2e^{-1}(3e^{-1} + 1)e^t$

(vi) $u^1 = u^0 - \tau\left(\delta H / \delta u\right)^0$

$= 1 - \tau\left[1 + 2e^{-1}(3e^{-1} + 1)e^t\right]$

$\tau = 0.1$

ETC.

DIFFERENTIAL APPROXIMATION

$\dot{x} = f(x, u, a, t)$

a IS UNKNOWN. ASSUME WE KNOW x & $\dot{x}$

WE WANNA ESTIMATE a.

PICK AN a TO MINIMIZE

$$J(a) = \frac{1}{2}\int_{t_0}^{t_f} \| \dot{x} - f(x, u, a, t) \|_R^2 \, dt$$

$$\frac{\delta J}{\delta a} = \int_{t_0}^{t_f} \frac{1}{2} \frac{d}{\delta a} \left[(\dot{x} - f(a))^T R (\dot{x} - f(a)) \right] dt$$

$$= \frac{1}{2}\int_{t_0}^{t_f} -2 \frac{\delta f}{\delta a}^T R (\dot{x} - f) \, dt = 0$$

$$= -\int_{t_0}^{t_f} \frac{\delta f}{\delta a}^T R (\dot{x} - f) \, dt = 0$$

$$\Rightarrow \int_{t_0}^{t_f} \frac{\delta f}{\delta a}^T R \dot{x} \, dt$$

$$= \int_{t_0}^{t_f} \left(\frac{\delta f}{\delta a} \right)^T f \, dt$$

EXAMPLE

$$\dot{x} = -ax^2$$

$$t \in [0, 4]$$

EX

$\dot{x} = -ax^2 \qquad t \in [0, 4]$

ASSUM $R = 1$ THUS

$$\int_0^4 (-x^2)\, \dot{x}\, dt = \int_0^4 (-x^2)(-ax^2)\, dt$$

THUS

$$a = \frac{\int_0^4 (-x^2)\, \dot{x}\, dt}{\int_0^4 x^4\, dt}$$

EX

$\dot{x} = f(x, u, t, a) \qquad x(0) = 0 \qquad t \in [0, 2]$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad a = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

$$f = \begin{bmatrix} \frac{25}{6}(\alpha x_2 + \beta x_2^3) \\ 6U - 6x_1 - 6x_2 \end{bmatrix}$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} \frac{25}{6}(\alpha x_2 + \beta x_2^3) \\ 6U - 6x_1 - 6x_2 \end{bmatrix}$$

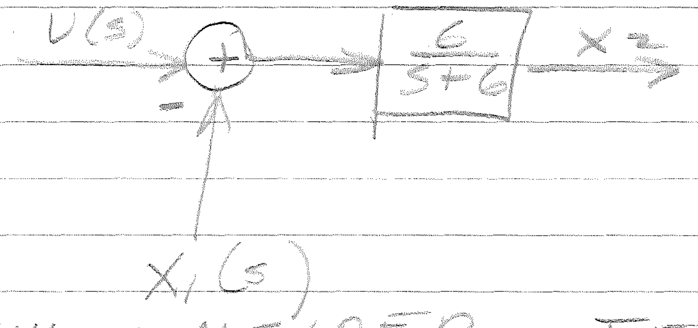
WE WANNA FIND α & β

NOW

$$s X_2 = 6U - 6X_1 - 6X_2$$

$$(s+6) X_2 = 6U - 6X_1$$

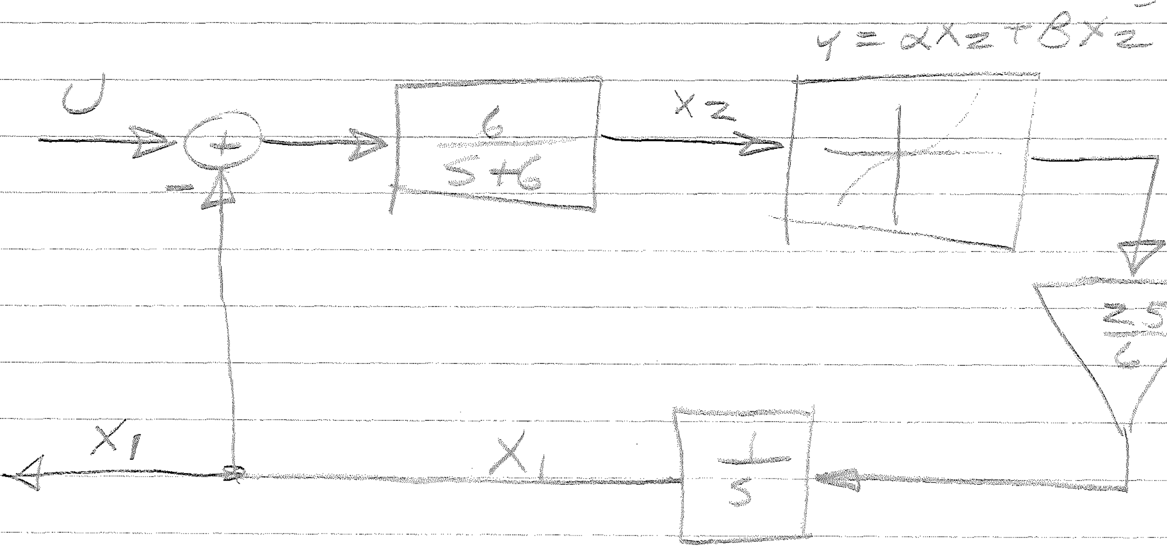
$$\Rightarrow X_2(s) = \frac{6}{s+6}\left[U(s) - X_1(s)\right]$$



NOW, CONSIDER TERM

$$\dot{X}_1 = \frac{25}{6}(\alpha X_2 + B X_2^3)$$

$$= \frac{25}{6} Y$$



NOW

$$\int_{t_0}^{t_f} \left(\frac{\delta f}{\delta q}\right)^T R \dot{x} \, dt$$

$$= \int_{t_0}^{t_f} \left(\frac{\delta f}{\delta q}\right)^T R f \, dt$$

DIFFERENTIATE WRT. q $\Longrightarrow$

$$\frac{\delta f}{\delta a} = \begin{bmatrix} \frac{\delta f_1}{\delta \alpha} & \frac{\delta f_1}{\delta B} \\ \frac{\delta f_2}{\delta \alpha} & \frac{\delta f_2}{\delta B} \end{bmatrix} = \begin{bmatrix} \frac{25}{6} x_2 & \frac{25}{6} x_2^3 \\ 0 & 0 \end{bmatrix}$$

SO, WE GET (ASSUME R = 1)

$$\int_{t_0}^{t_f} \begin{bmatrix} \frac{25}{6} x_2 & 0 \\ \frac{25}{6} x_2^3 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} dt$$

$$= \int_{t_0}^{t_f} \begin{bmatrix} \frac{25}{6} x_2 & 0 \\ \frac{25}{6} x_2^3 & 0 \end{bmatrix} \begin{bmatrix} \frac{25}{6} (\alpha x_2 + \beta x_2^3) \\ 6u - 6x_1 - 6x_2 \end{bmatrix} dt$$

OR

$$\int_{t_0}^{t_f} \begin{bmatrix} \frac{25}{6} x_2 \dot{x}_1 \\ \frac{25}{6} x_2^3 \dot{x}_1 \end{bmatrix} dt$$

$$= \int_{t_0}^{t_f} \begin{bmatrix} \left(\frac{25}{6}\right)^2 (\alpha x_2 + \beta x_2^3) x_2 \\ \left(\frac{25}{6}\right)^2 x_2^3 (\alpha x_2 + \beta x_2^3) \end{bmatrix} dt$$

GIVES

$$\int_{t_0}^{t_f} \left(\frac{25}{6}\right)^2 x_2 \, dt \, \alpha + \int_{t_0}^{t_f} \left(\frac{25}{6}\right)^2 x_2^4 \, dt \, \beta$$

$$= \int_{t_0}^{t_f} \frac{25}{6} \dot{x}_1 x_2 \, dt$$

AND

$$\int_{t_0}^{t_f} \left(\left(\frac{25}{6}\right)^2 x_2^4 \right) dt \, \alpha$$

$$+ \int_{t_0}^{t_f} \left(\frac{25}{6}\right) x_2^6 \, dt \, \beta$$

$$= \int_{t_0}^{t_f} \frac{25}{6} x_2^3 \dot{x}_1 \, dt$$

KNOWING x_1 AND x_2, WE CAN SOLVE HERE FOR α & β WHICH ARE LIKE

$$\begin{bmatrix} \int & \int \\ \int & \int \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \int \\ \int \end{bmatrix}$$

COUPLING WITH QUASILINEARIZATION

$\dot{x} = f(x, u, a, t)$ $\quad x = n \times 1$

$\quad u = m \times 1$

SINCE a IS CONST. $\quad a = r \times 1$

$\dot{a} \equiv 0$

GIVES A TOTAL OF $n+r$ EQUATIONS

ASSUME WE MEASURE x's,

$\langle c(t_i), x(t_i) \rangle = b_i \quad ; \; i = 1, 2, \ldots q$

MULTI B.C.

WE CAN, IN PRINCIPLE SOLVE THIS PROBLEM WITH QUASILINEARIZATION.

PLUG SHEET (TEST 1)

CONTROLLABILITY $[B|AB|A^2B|\ldots|A^{n-1}B]$ (RANK n)

OBSERVABILITY $[C^T|A^TC^T|A^{T^2}C^T|\ldots|A^{T^{n-1}}C^T]$

VARIATIONAL CALCULUS

$J = \int_{t_0}^{t_f} \phi(x, \dot{x}, t)dt$ ⟸ PERFORMANCE MEASURE

$\frac{\delta\phi}{\delta x} - \frac{d}{dt}\frac{\delta\phi}{\delta\dot{x}} = 0$ ⟸ EULER'S EQUATION

$\frac{\delta\phi}{\delta\dot{x}}\delta x + [\phi - \frac{\delta\phi}{\delta\dot{x}}\dot{x}]\delta t = 0$ ⟸ TRANSVERSALITY CONDITION

$[\frac{\delta\phi}{\delta\dot{x}}(\dot{\Theta} - \dot{x}) + \phi = 0$ ⟸ FOR WHEN $x(t_f) = \Theta(t_f)]$

$$\left.\begin{array}{l} \left.\frac{\delta\phi}{\delta\dot{x}}\right|_{t_-} = \left.\frac{\delta\phi}{\delta\dot{x}}\right|_{t_+} \\ \left.\phi - \dot{x}\frac{\delta\phi}{\delta\dot{x}}\right|_{t_-} = \left.\phi - \dot{x}\frac{\delta\phi}{\delta\dot{x}}\right|_{t_+} \end{array}\right\} \text{WEIERSTRASS - ERDMAN CORNER CONDITIONS}$$

VARIOUS CONSTRAINTS (LAGRANGE MULTIPLIERS)

$f(\omega, t) = 0 \Rightarrow \phi_a = \phi + \lambda^T f$ ⟸ POINT CONSTRAINT

$f(\omega, \dot{\omega}, t) = 0 \Rightarrow \phi_a = \phi + \lambda^T f$ ⟸ DIFFERENTIAL CONSTRAINT

$\Gamma_{min} < \Gamma < \Gamma_{MAX} \Rightarrow \phi_a = \phi + \lambda^T[\alpha^2 - (\Gamma_{MAX} - \Gamma)(\Gamma - \Gamma_{MIN})]$

⟸ INEQUALITY CONSTRAINT

ISOPERIMETRIC CONSTRAINT $\Rightarrow C = \int_{t_0}^{t_f} e\,dt$

$\dot{z} = e,\ z(t_0) = 0,\ z(t_f) = C,\ \phi_a = \phi + \lambda^T(\dot{z} - e)$

TURNS OUT λ = CONSTANT

THE BOLZA PROBLEM

$J = \Theta[x(t), t]\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \phi(x, u, t)dt$ ⟸ PER. MEASURE

$\dot{x} = f(x, u, t)$ ⟸ n VECTOR

$$\left.\begin{array}{l} M(t_0)x(t_0) = m_0 = 0 \\ N(t_f)x(t_f) = n_f = 0 \end{array}\right\} \text{BOUNORY CONDITIONS}$$

$H[x, u, \lambda, t] = \phi + \lambda^T f$ ⟸ HAMILTONIAN

$\circleddash = \Theta + \nu^T N$

$\frac{\delta H}{\delta\lambda} = \dot{x} = f$

$\frac{\delta H}{\delta x} = -\acute{\lambda} = \frac{\delta f}{\delta x}^T\lambda + \frac{\delta\phi}{\delta x}$

$\frac{\delta H}{\delta u} = 0$

$\lambda(t_f) = \frac{\delta\circleddash}{\delta x(t_f)}$

$H + \frac{\delta\circleddash}{\delta t_f} + \frac{\delta N^T}{\delta t_f}\nu = 0$

(THESE ASSUMING $x(t_0) = x_0$)

LINEAR REGULATOR

$\dot{x} = AX + BU$

$J = \frac{1}{2} \| X(t_f) \|_S^2 + \frac{1}{2} \int_{t_0}^{t_f} [\| x(t) \|_Q^2 + \| U \|_R^2] dt$

$U = -R^{-1} B^T \lambda \quad , \quad \lambda(t_f) = S x(t_f)$

$\lambda = PX \Rightarrow \dot{P} + PA + A^T P - PBR^{-1} B^T P + Q = 0 \Leftarrow$ RICCATI EQUATION

$P(t_f) = S$

$U = KX \Leftarrow K =$ KALMAN GAIN

BOLZA PROBLEM WITH INEQUALITY CONSTRAINT

$\dot{x} = f, \quad N|_{t_f} = 0 \quad g[x, u, t] \geq 0$

$J = \theta|_{t_f} + \int_{t_0}^{t_f} \phi \, dt$

$\dot{z}^2 = g$

$H = \phi + \lambda^T f \quad , \quad \Theta = \theta + \nu^T N$

$\Phi = H - \lambda^T \dot{x} - \Gamma^T (g - \dot{z}^2) \Leftarrow$ LAGRANGIAN

GIVES

$\frac{d}{dt} \frac{\delta \Phi}{\delta \dot{x}} - \frac{\delta \Phi}{\delta x} = 0$

$\frac{d}{dt} \frac{\delta \Phi}{\delta \dot{w}} = 0 \qquad \dot{w} = u, \ w(t_0) = 0$

$\frac{d}{dt} \delta \Phi / \delta \dot{z} = 0$

$\frac{\delta \Theta}{\delta t_f} + \frac{\delta \Theta}{\delta x}^T \dot{x} + \Phi = 0$

WEIERSTRASS E FUNCTION:

$E = \Phi(x, \dot{X}, t) - \Phi(x, \dot{x}, t) - (\dot{X} - x) \frac{\delta \Phi}{\delta \dot{x}} > 0$

PONTRYAGIN'S MAXIMUM PRINCIPLE:

$H(x, \hat{u}, \lambda, t) \leq H(x, u, \lambda, t)$

$\uparrow$ OPTIMAL CONTROL

FINAL CRAM SHEET

- $J = \int_{t_0}^{t_f} g[x(t), \dot{x}(t), t]\,dt$

$\frac{\delta g}{\delta x} - \frac{d}{dt}\frac{\delta g}{\delta \dot{x}} = 0 \Longleftarrow$ EULER'S EQUATION

$\frac{\delta g^T}{\delta \dot{x}}\delta x_f + [g - \frac{\delta g^T}{\delta \dot{x}}\dot{x}]\,\delta t_f = 0 \Longleftarrow$ TRANS. COND.

-WEIERSTRASS-ERDMAN CORNER CONDITIONS

① $\frac{\delta g}{\delta \dot{x}}\Big|_{t_1-} = \frac{\delta g}{\delta \dot{x}}\Big|_{t_1+}$

② $[g - \frac{\delta g^T}{\delta \dot{x}}\dot{x}]\Big|_{t_1-} = [g - \frac{\delta g^T}{\delta \dot{x}}\dot{x}]\Big|_{t_1+}$

-CONSTRAINED MINIMIZATION

(1) POINT: $f = 0 \Rightarrow g_a = g + \lambda^T f$

(2) ISOPERIMETRIC: $c_i = \int_{t_0}^{t_f} e_i\,dt$; $z = \int_{t_0}^{t} e_i dt$

$g_a = g + \lambda^T[e_i - \dot{z}]$; $z(t_f) = c_i$

(3) INEQUALITY: $\Gamma_{min} < \Gamma < \Gamma_{max}$; $\alpha^2 = (\Gamma_{MAX} - \Gamma)(\Gamma - \Gamma_{min})$

$g_a = g + \lambda^T[\alpha^2 - (\Gamma_{MAX} - \Gamma)(\Gamma - \Gamma_{min})]$

-WEIERSTRAUSS E FUNCTION

$E = \Phi[x, \dot{X}, t] - \Phi(x, \dot{x}, t) - (\dot{X} - \dot{x})^T \frac{\delta \Phi}{\delta \dot{x}} > 0$

$\Phi = \mathcal{H} - \lambda^T \dot{x} - \lambda^T(\text{cons.})$

- PONTRYAGIN'S MAX PRINCIPLE

$\mathcal{H}[x, \hat{u}, \lambda, t] \leq \mathcal{H}[x, u, \lambda, t]$

$\mathcal{H} = g[x, u, t] + \lambda^T \dot{x}$

$J = h(t_f, x(t_f)) + \int_{t_0}^{t_f} g(x, u, t)\,dt$

OPTIMAL $\mathcal{H} = 0$ IF $\mathcal{H}$ NOT EXPLICIT FUNC OF TIME

- BANG BANG CONTROL

$\dot{x} = f(x, t) + G(x, t)\,u(t)$; $|u| \leq 1$

$J = \Theta_f + \int_{t_0}^{t_f} [\phi(x, t) + h^T u]\,dt$

$\mathcal{H} = \phi + h^T u + \lambda^T[f + Gu]$

$\mathcal{H}[x, \hat{u}, t] \leq \mathcal{H}[x, u, t] \Rightarrow U = -sgn(h^T + \lambda^T G)$

- BANG-REST BANG CONTROL

  FOR MINIMUM FUEL PROBLEM

- SINGULAR PROBLEMS

  $\lambda = 0$ USE $\dot{\lambda}, \ddot{\lambda} = 0$ AND PLUG AWAY

- QUASILINEARIZATION

  $\dot{x}^{(i+1)} = f^{0} + \frac{\delta f}{\delta x}\Big|_0 \left[x^{(i+1)} - x^{(i)}\right] + \frac{\delta f}{\delta \lambda_0}\left[\lambda^{(i+1)} - \lambda^{(i)}\right]$

  $\dot{\lambda}^{(i+1)} = g^{0} + \dots$

  ASSUME $\lambda^{H}(0) = 1$ & $X^{H}(0) = 0$

  $\lambda^{P}(0) = 0$ & $X^{P}(0) = X_0$

  B.C. ARE: $\begin{bmatrix} X(0) \\ \lambda(0) \end{bmatrix} = \begin{bmatrix} X_0 \\ C_1 \\ \vdots \\ C_n \end{bmatrix}$

- METHOD OF STEEPEST DESCENT

  IN CALCULUS: $Y = f(x_1, x_2)$

  Z = UNIT VECTOR $= \frac{\delta Y/\delta X|_{x_0}}{\sqrt{\left(\frac{\delta Y}{\delta X_1}\right)^2_{x_0} + \left(\frac{\delta Y}{\delta X_2}\right)_0}}$

  $\Rightarrow Y^{(1)} = Y^{(0)} - Z_0 \gamma$

  FUNCTIONALS 1. ESTIMATE U^{0}

  2. COMPUTE X^{0} FROM $\dot{x}^{0} = f(x^{0}, u^{0}, t)$
  3. COMPUTE λ_f
  4. INTEGRATE COSTATE $\dot{\lambda}^{(0)} = -\frac{\delta \mathcal{H}}{\delta x}\Big|_0$
  5. COMPUTE $\delta \mathcal{H}/\delta u$ (IDEALLY = 0)

     IS $\|\cdot\|^2 \leq \gamma^2$?

  6. COMPUTE $U^{(1)} = U^{(0)} - \gamma \frac{\delta \mathcal{H}}{\delta X_0}$

- DIFFERENTIAL APPROXIMATION

  $\dot{x} = f(x, u, a, t)$

  $J(a) = \int_{t_0}^{t_f} \|\dot{x} - f\|$

  $\delta J_a = 0 \Rightarrow \int_{t_0}^{t_f} \frac{\delta f^T}{\delta a} R \dot{x}\, dt = \int_{t_0}^{t_f} \frac{\delta f^T}{\delta a} f\, dt$

William L. Everitt, *editor*

NETWORK SERIES
Robert Newcomb, *editor*

- Anner
  - *Elementary Nonlinear Electronic Circuits*
- Carlin & Giordano
  - *Network Theory: An Introduction to Reciprocal and Nonreciprocal Circuits*
- Chirlian
  - *Integrated and Active Network Analysis and Synthesis*
- Deutsch
  - *Systems Analysis Techniques*
- Herrero & Willoner
  - *Synthesis of Filters*
- Manasse, Eckert & Gray
  - *Modern Transistor Electronics Analysis and Design*
- Newcomb
  - *Active Integrated Circuit Systhesis*
- Sage
  - *Optimum Systems Control*
- Van Valkenburg
  - *Network Analysis, 2nd Ed.*

PRENTICE-HALL INTERNATIONAL, INC., *London*
PRENTICE-HALL OF AUSTRALIA, PTY. LTD., *Sydney*
PRENTICE-HALL OF CANADA, LTD., *Toronto*
PRENTICE-HALL OF INDIA PRIVATE LTD., *New Delhi*
PRENTICE-HALL OF JAPAN, INC., *Tokyo*

OPTIMUM SYSTEMS CONTROL

Andrew P. Sage

Professor and Director
Information and Control Sciences Center
Institute of Technology
Southern Methodist University

Kwongshu Chao
Aug. 31, 1970

PRENTICE-HALL, INC.
Englewood Cliffs, N.J.

2

CALCULUS OF EXTREMA AND SINGLE-STAGE DECISION PROCESSES

Many problems in modern system theory may be simply stated as extreme value problems. These can be resolved via the calculus of extrema which is the natural solution method whenever one desires to find parameter values which minimize or maximize a quantity dependent upon them. In this chapter we will consider several such problems, starting with simple scalar problems and concluding with a discussion of the vector case. The method of Lagrange multipliers will be introduced and used to solve constrained extrema problems for single-stage decision processes. A brief discussion of linear and nonlinear programming will be presented. Multistage decision processes, which can be treated by the calculus of extrema, will be reserved for a variational treatment which will result in a discrete maximum principle. Much of the work in this chapter is very basic, and a selection of only references [1] through [5] of direct interest to the systems control area is given.

2.1 Maxima and minima (scalar process)

A real function $f(x)$, defined for a scalar $x = \alpha$, has a relative maximum or a relative minimum $f(\alpha)$ for $x = \alpha$ if and only if there exists a positive real number δ such that, respectively,

$$\Delta f = f(\alpha + \Delta x) - f(\alpha) < 0 \tag{2.1-1}$$

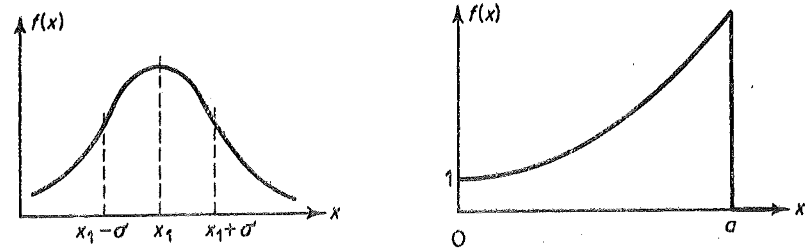
or

$$\Delta f = f(\alpha + \Delta x) - f(\alpha) > 0 \qquad (2.1\text{-}2)$$

for all $\Delta x = x - \alpha$ such that $f(\alpha + \Delta x)$ exists and $0 < |\Delta x| < \delta$. Further, if $df(x)/dx$ exists and is also continuous at $x = \alpha$, then $f(\alpha)$ can be an interior maximum or minimum only if

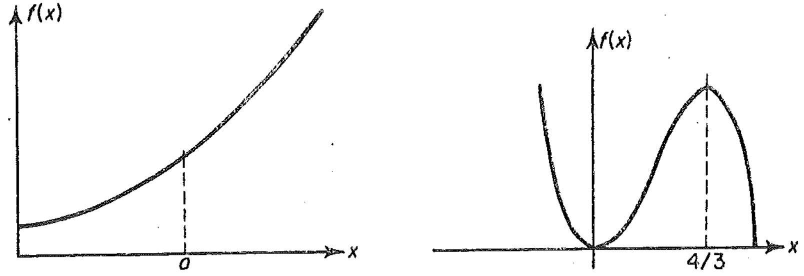
$$\left.\frac{df(x)}{dx}\right|_{x=\alpha} = 0 \qquad (2.1\text{-}3)$$

If $f(x)$ has a continuous second derivative for $x = \alpha$, the nature of the extremum at $x = \alpha$ can be determined. The following well-known procedure



(a)
$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp[-(x-x_1)^2/2\sigma^2]$
For x in the interval: $(-\infty, \infty)$
$f(x)$ has an absolute maximum at $x = x_1$.

(b)
$f(x) = e^x[u(x) - u(x-a)]$
For x in the interval: $[0, a)$
$f(x)$ has an absolute minimum at $x = 0$, and an absolute maximum at $x = a - \delta$ where δ is an arbitrarily small positive number.



(c)
$f(x) = e^x u(x)$
For x in the interval: $[0, a]$
$f(x)$ has an absolute minimum at $x = 0$, and an absolute maximum at $x = a$.
For x in the interval: $[0, +\infty]$
$f(x)$ has an absolute minimum at $x = 0$.

(d)
$f(x) = x^2(2 - x)$
For x in the interval: $(-\infty, +\infty)$
$f(x)$ has a relative minimum at $x = 0$, and a relative maximum at $x = 4/3$.

Fig. 2.1-1. Illustrations of extrema.

can be used for the determination of the extrema of a giv alar function $y = f(x)$.

1. Differentiate y with respect to x.
2. For each value of x, determine the specific values of α which satisfy the equation $dy/dx = 0$.
3. Test to see what kind of extrema the function has for each value of α thus obtained. This we can easily accomplish by the second-derivative test in which we substitute each value of α into the second derivative of y with respect to x and apply the following rule:

$$\text{If}\quad \frac{d^2y}{dx^2}\left\}\begin{array}{lll} > 0 & \text{then} & y \text{ has a relative minimum} \\ < 0 & \text{then} & y \text{ has a relative maximum} \\ = 0 & \text{then} & y \text{ has a stationary point} \end{array}\right. \qquad (2.1\text{-}4)$$

4. Evaluate the actual value of the extrema by substituting each value of α obtained into $f(x)$.

There are three different types of extrema possible. If a value of α can be found such that $f(\alpha)$ is an extremum for all x throughout its domain of definition, $f(x)$ is said to have an absolute extremum. If a value of α can be found such that $f(\alpha)$ has an extremum throughout a bounded neighborhood of x, $f(x)$ has a relative extremum at $x = \alpha$. If $f(x)$ is defined only for a limited range of values of x, and if $f(x)$ has an extremum at either boundary of x (with respect to all the values $f(x)$ has for all values of x contained within the limited range of x), then $f(x)$ has an extremum at its boundary. These different types of extrema are illustrated in Fig. 2.1-1. We will have opportunity to apply these concepts to parameter optimization of control systems in Sections 8.2 and 13.3-1.

2.2 Extrema of functions of two or more variables

The extrema-finding technique can be extended to include functions of more than one variable. Suppose $y = f(x_1, x_2, \ldots, x_n) = f(\mathbf{x})$. A procedure similar to the previous one is used, using partial derivatives instead of total derivatives. A simple example will illustrate the procedure to be followed.

Example 2.2-1

Let us consider the maximization of

$$y(\mathbf{x}) = \frac{1}{(x_1 - 1)^2 + (x_2 - 1)^2 + 1}, \qquad \mathbf{x}^T = [x_1, x_2]$$

where $\mathbf{x}^T$ is to indicate transpose of the column vector $\mathbf{x}$.† Following an extended version of the foregoing scalar procedure, we take the partial derivatives of y with respect to x_1 and x_2 and set them equal to zero to obtain:

$$\frac{\partial y}{\partial x_1} = \frac{(-1)(2x_1 - 2)}{[(x_1 - 1)^2 + (x_2 - 1)^2 + 1]^2} = 0, \qquad \alpha_1 = 1$$

$$\frac{\partial y}{\partial x_2} = \frac{(-1)(2x_2 - 2)}{[(x_1 - 1)^2 + (x_2 - 1)^2 + 1]^2} = 0, \qquad \alpha_2 = 1$$

Thus, since $\alpha_1 = \alpha_2 = 1$ are the only extrema, and since a simple computation shows that the second derivatives are nonpositive at this extrema, we see that we have a maximum at the point $\mathbf{x}^T = [1, 1]$.

Example 2.2-2

Let us now suppose that the allowable range of $\mathbf{x}$ is constrained such that $|x_1| \leq \frac{1}{2}$ and $|x_2| \leq \frac{1}{2}$. It is desired to find the value of $\mathbf{x}$ which yields a maximum for the $y = f(\mathbf{x})$ of Example 2.2-1 in the allowable or admissible range of $\mathbf{x}$. This region of state space is also shown in Fig. 2.2-1. From this figure, it is apparent that, for this simple problem, $y = f(\mathbf{x})$ has an extremum (maximum) somewhere on the boundary of the admissible range for $\mathbf{x}$, in fact precisely at $\mathbf{x}^T = [\frac{1}{2}, \frac{1}{2}]$. This is a very simple example of optimization with an inequality constraint. We will have considerably more to say about this very important type of constraint when we consider dynamic systems and the calculus of variations.

Example 2.2-3

A slightly more difficult problem arises if the allowable range of $\mathbf{x}$ is constrained such that the Euclidean norm of $\mathbf{x}$ equals one. Symbolically, this means that $\|\mathbf{x}\|^2 = \mathbf{x}^T\mathbf{x} = x_1^2 + x_2^2 + \ldots + x_n^2 = \langle \mathbf{x}, \mathbf{x} \rangle$. Since the dimension of the example that we are considering is two, the Euclidean norm squared becomes $\|\mathbf{x}\|^2 = x_1^2 + x_2^2$.

One approach to the problem is to solve for x_1 in terms of x_2, then solve for $y = f(\mathbf{x})$ in terms of x_2 alone. This will then allow us to use the standard scalar procedure. From the given constraint on the length of the Euclidean norm, we have $x_1 = (1 - x_2^2)^{1/2}$. Substituting this into the expression for $y(\mathbf{x})$ of Example 2.2-1, we find that

$$y(x_2) = \frac{1}{(\pm\sqrt{1 - x_2^2} - 1)^2 + (x_2 - 1)^2 + 1}$$

where $y(x_2)$ has the given constraint imbedded into it. The next step is to differentiate this expression with respect to the remaining variable, x_2, and set the result equal to zero. This yields two solutions. The second-derivative test shows that a maximum (which is easily shown to be an absolute maximum) occurs at $\mathbf{x}^T = [0.707, 0.707]$ and that an (absolute) minimum occurs at $\mathbf{x}^T = [-0.707, -0.707]$.

We note that, in the absence of the equality constraint, this problem has no

---

†Appendix A contains a brief presentation of vector matrix notations and vector matrix calculus.

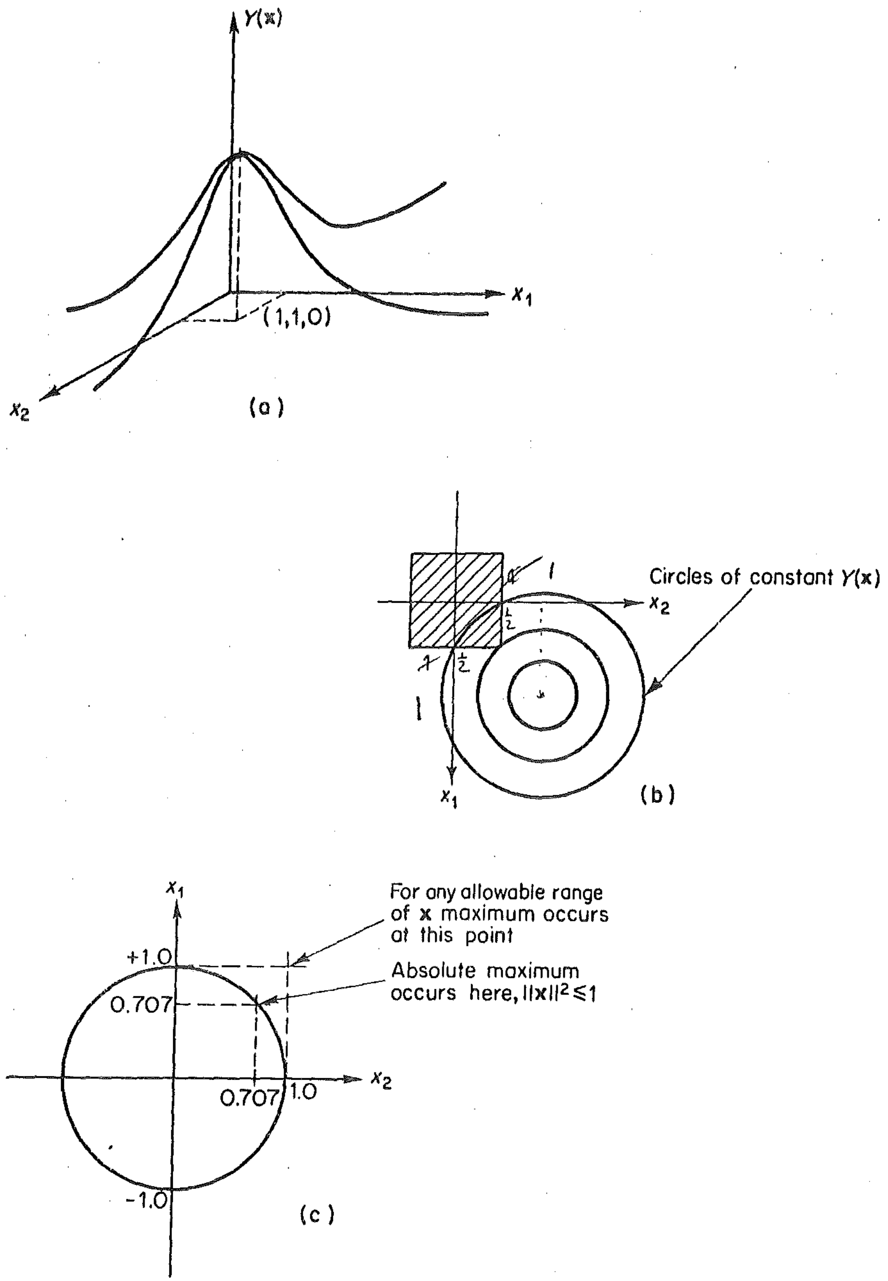


Fig. 2.2-1(a) $y(\mathbf{x}) = 1/[(x_1 - 1)^2 + (x_2 - 1)]^2$; (b) Top view of Fig. 2.2-1a showing the region defined in state space by $|x_1| \leq 1/2$; $|x_2| \leq 1/2$; (c) Top view of Fig. 2.2-1a showing the region of state space defined by $\|\mathbf{x}\|^2 \leq 1$.

two inequality constraints of this type, namely $\|\mathbf{x}\|^2 \leq 1$ and $\|\mathbf{x}\|^2 < 1$. The first constraint set is closed (and conve nce it includes the boundary $\|\mathbf{x}\|^2 = x_1^2 + x_2^2 = 1$. The second is open (and convex) since it does not include the boundary. It is generally quite difficult to work with constraints of this form. One method, satisfactory in quite a few problems, is to ignore the constraint and find the maximum (or minimum). If this turns out to be interior to the boundary of the constraint set, we have the solution. If the maximum (or minimum) occurs outside the boundary, the inequality constraint is treated as an equality constraint, and a solution is found with this constraint. Another method, to be discussed later, is to convert the inequality constraint to an equality constraint. Figure 2.2-1 illustrates salient features of these examples.

2.3 Constrained extrema problems—Lagrange multipliers

An alternate approach to extremizing a function (i.e., find those values of the independent variables which cause the dependent variable to have an extremum) with given constraints or accessory conditions is to make appropriate adjustments on the independent variable by using an adjustable multiplying parameter, commonly called a Lagrange multiplier. The procedure is to form a new function by adjoining the given constraint to the original function. This new function, then, is extremized, by means of the previously developed method. We will solve an example first by the more straightforward, but often more cumbersome, procedure and then by using the Lagrange multiplier. Considerably more justification for the Lagrange multiplier procedure will be provided in the next chapter on variational calculus.

Example 2.3-1

A tin can manufacturer wants to maximize the volume of a certain run of cans subject to the constraint that the area of tin used be a given constant. If a fixed metal thickness is assumed, a volume of tin constraint implies that the cross-sectional area is constrained.

The defining equations for this problem are:

$$\text{Volume} = V(r, l) = \pi r^2 l \tag{1}$$

$$\text{Cross-sectional area} = A(r, l) = 2\pi r^2 + 2\pi r l = A_o \tag{2}$$

Our problem is to maximize $V(r, l)$ subject to keeping $A(r, l) = A_o$, where A_o is a given constant. The same approach can be used here as in Example 2.2-3. We solve for l in terms of r (or if preferred, r in terms of l) and then express the volume as a function of r alone, noting that the constraint on the cross-sectional area is now imbedded into the expression for the volume. We then examine the first and second derivatives to discern the character and location of the extrema.

Method 1

From Eq. (2) we have

$$l = \frac{A_o - 2\pi r^2}{2\pi r} \tag{3}$$

By substituting Eq. (3) into Eq. (1), we obtain

$$V(r) = \frac{r}{2}A_o - \cancel{(2\pi r^2)} \quad \pi r^3 \tag{4}$$

We differentiate V with respect to r and set the result equal to zero to obtain

$$\frac{dV(r)}{dr} = \frac{A_o}{2} - 3\pi r^2 = 0, \qquad r = \sqrt{\frac{A_o}{6(\pi)}} \tag{5}$$

We now substitute Eq. (5) into Eq. (2) and solve for l:

$$l = \sqrt{\frac{2A_o}{3\pi}} \tag{6}$$

It is interesting to obtain the optimum length-to-radius ratio. In doing this, we see that, to get maximum volume, we make the length of the tin can equal the diameter, keeping cross-sectional area equal to a given constant.

Method 2

By using the Lagrange multiplier, we again want to extremize (maximize) the volume $V(r, l)$ subject to the constraint $A(r, l) = A_o$. First we form the adjoined function

$$V'(r, l) = V(r, l) + \lambda[A(r, l) - A_o]$$

where λ is the Lagrange multiplier. In terms of the parameters of the tin can, this expression becomes

$$V'(r, l) = \pi r^2 l + \lambda[2\pi r^2 + 2\pi r l - A_o]$$

We take the first partial derivative with respect to each of the variables and set each result equal to zero. Thus we obtain

$$\frac{\partial V'(r, l)}{\partial l} = \pi r^2 + \lambda 2\pi r = 0, \qquad r = -2\lambda$$

$$\frac{\partial V'(r, l)}{\partial r} = 2\pi r l + \lambda[4\pi r + 2\pi l] = 0, \qquad l = 2r$$

We now evaluate λ subject to given constraint, $A(r, l) = A_o$ or

$$A_o = 2\pi r^2 + 2\pi r l$$

In terms of the obtained values of r and l, this becomes

$$A_o = 2\pi(4\lambda^2) + 2\pi(-2\lambda)(-4\lambda)$$

so

$$\lambda = \pm\sqrt{\frac{A_o}{24\pi}}$$

$$r = 2\sqrt{\frac{A_0}{24\pi}}, \qquad l = 4\sqrt{\frac{A_0}{24\pi}}$$

We note that the negative square root is selected for λ to make r and l physically realizable quantities. We further note that the length-to-radius ratio is the same as obtained by the first method, as it well should be.

2.4 Vector formulation of extrema problems—single-stage decision processes

Considerable notational simplification occurs if functions of more than one variable are written in state vector notation. Thus a scalar function of several variables which is to be extremized

$$J = \theta(x_1 x_2, \ldots, x_n) \tag{2.4-1}$$

may be written as

$$J = \theta(\mathbf{x}) \tag{2.4-2}$$

where

$$\mathbf{x}^T = [x_1, x_2, \ldots, x_n] \tag{2.4-3}$$

For the majority of systems problems, it is convenient to distinguish between control vectors and state vectors. We generally desire to find a control vector, $\mathbf{u}$ or $\mathbf{u}(k)$, or $\mathbf{u}(t)$ if we have a multistage or continuous process which minimizes or maximizes some scalar index of performance of the system. This performance index will be called J. Possibly the simplest single-stage decision process with equality constraints is to minimize or maximize the scalar index of performance

$$J = \theta[\mathbf{x}, \mathbf{u}] \tag{2.4-4}$$

subject to the equality constraint

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = \mathbf{0} \tag{2.4-5}$$

where $\mathbf{x}$ is an n vector

$$\mathbf{x}^T = [x_1 x_2, \ldots, x_n] \tag{2.4-6}$$

$\mathbf{u}$ is an m vector

$$\mathbf{u}^T = [u, u_2, \ldots, u_m] \tag{2.4-7}$$

$\mathbf{f}$ is an n vector function

$$\mathbf{f}^T(\mathbf{x}, \mathbf{u}) = [f_1(\mathbf{x}, \mathbf{u}), f_2(\mathbf{x}, \mathbf{u}), \ldots, f_n(\mathbf{x}, \mathbf{u})] \tag{2.4-8}$$

The solution proceeds as follows. We adjoin Eq. (2.4-5) to Eq. (2.4-4) with a vector Lagrange multiplier in order to form a scalar quantity† which we will call $H(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda})$.

$$H(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) = \theta(\mathbf{x}, \mathbf{u}) + \boldsymbol{\lambda}^T\mathbf{f}(\mathbf{x}, \mathbf{u}) \tag{2.4-9}$$

$$\boldsymbol{\lambda}^T = [\lambda_1, \lambda_2, \ldots, \lambda_n] \tag{2.4-10}$$

We now adjust $\mathbf{x}$ and $\mathbf{u}$ such that H is a maximum or minimum. This requires

$$\frac{\partial H}{\partial \mathbf{x}} = \frac{\partial \theta}{\partial \mathbf{x}} + \frac{\partial}{\partial \mathbf{x}}\mathbf{f}^T(\mathbf{x}, \mathbf{u})\boldsymbol{\lambda} = \mathbf{0} \tag{2.4-11}$$

$$\frac{\partial H}{\partial \mathbf{u}} = \frac{\partial \theta}{\partial \mathbf{u}} + \frac{\partial}{\partial \mathbf{u}}\mathbf{f}^T(\mathbf{x}, \mathbf{u})\boldsymbol{\lambda} = \mathbf{0} \tag{2.4-12}$$

where

$$\left[\frac{\partial H}{\partial \mathbf{u}}\right]^T = \left[\frac{\partial H}{\partial u_1}, \frac{\partial H}{\partial u_2}, \ldots, \frac{\partial H}{\partial u_m}\right] \tag{2.4-13}$$

Thus $\partial H/\partial \mathbf{u}$ may be interpreted as the gradient of H with respect to $\mathbf{u}$, which is commonly designated $\nabla_{\mathbf{u}} H$. Also,

$$\frac{\partial}{\partial \mathbf{x}}\mathbf{f}^T(\mathbf{x}, \mathbf{u}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_2}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_1} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \frac{\partial f_1}{\partial x_n} & & & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \tag{2.4-14}$$

It should be noted that Eq. (2.4-14) is similar to the transpose of the Jacobian of a vector

$$[J_{\mathbf{x}}\mathbf{f}(\mathbf{x}, \mathbf{u})]^T = \begin{vmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_2}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_1} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \frac{\partial f_1}{\partial x_n} & & \mathbf{f}^T & \frac{\partial f_n}{\partial x_n} \end{vmatrix} \tag{2.4-15}$$

with at least two important differences: $\partial\mathbf{f}(\mathbf{x}, \mathbf{u})/\partial\mathbf{u}$ need not be square and is a matrix rather than a determinant. In order that J be an extremum, not only must

$$\frac{\partial H}{\partial \mathbf{x}} = 0; \qquad \frac{\partial H}{\partial \mathbf{u}} = 0 \tag{2.4-16}$$

but also the second variation of H must be greater than zero for a minimum or less than zero for a maximum (see second-derivative test, Section 2.1.)

---

†This scalar quantity, the Hamiltonian, has a number of very interesting properties that will be mentioned in later chapters.

Chapters 3, [illegible]nd 13 will provide us with considerably more information on the second variation than we present here. To see what this constraint on the second variation of H means, in terms of the necessary conditions required for making $J(\mathbf{x}, \mathbf{u})$ have an extremum, let us now formulate the second variation of $H(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda})$. The first variation of $H(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda})$ is

$$\delta H = \left(\frac{\partial H}{\partial \mathbf{x}}\right)^T \delta\mathbf{x} + \left(\frac{\partial H}{\partial \mathbf{u}}\right)^T \delta\mathbf{u} \tag{2.4-17}$$

which is the linear part of

$$\Delta H = H[\mathbf{x} + \delta\mathbf{x}, \mathbf{u} + \delta\mathbf{u}] - H[\mathbf{x}, \mathbf{u}] \tag{2.4-18}$$

To get the second variation of H, denoted $\delta^2 H$, we take the second-order part of the expansion of Eq. (2.4-18) in a Taylor series about $\delta\mathbf{u} = \mathbf{0}$, $\delta\mathbf{x} = \mathbf{0}$ to obtain

$$\begin{aligned}\delta^2 H = \frac{1}{2}\,\delta\mathbf{x}^T&\left\{\left[\frac{\partial}{\partial \mathbf{x}}\frac{\partial H}{\partial \mathbf{x}}\right]\delta\mathbf{x} + \left[\frac{\partial}{\partial \mathbf{u}}\frac{\partial H}{\partial \mathbf{x}}\right]\delta\mathbf{u}\right\}\\ + \frac{1}{2}\,\delta\mathbf{u}^T&\left\{\left[\frac{\partial}{\partial \mathbf{u}}\frac{\partial H}{\partial \mathbf{x}}\right]^T\delta\mathbf{x} + \left[\frac{\partial}{\partial \mathbf{u}}\frac{\partial H}{\partial \mathbf{u}}\delta\mathbf{u}\right]\right\}\end{aligned} \tag{2.4-19}$$

In more compact notation, this becomes

$$\delta^2 H = \frac{1}{2}[\delta\mathbf{x}^T\ \delta\mathbf{u}^T]\begin{bmatrix} \dfrac{\partial}{\partial \mathbf{x}}\dfrac{\partial H}{\partial \mathbf{x}} & \dfrac{\partial}{\partial \mathbf{u}}\dfrac{\partial H}{\partial \mathbf{x}} \\ \left[\dfrac{\partial}{\partial \mathbf{u}}\dfrac{\partial H}{\partial \mathbf{x}}\right]^T & \dfrac{\partial}{\partial \mathbf{u}}\dfrac{\partial H}{\partial \mathbf{u}} \end{bmatrix}\begin{bmatrix}\delta\mathbf{x}\\ \delta\mathbf{u}\end{bmatrix} \tag{2.4-20}$$

If we define

$$\delta\mathbf{z}^T = [\delta\mathbf{x}^T\ \delta\mathbf{u}^T], \qquad \mathbf{P} = \begin{bmatrix} \dfrac{\partial}{\partial \mathbf{x}}\dfrac{\partial H}{\partial \mathbf{x}} & \dfrac{\partial}{\partial \mathbf{u}}\dfrac{\partial H}{\partial \mathbf{x}} \\ \left[\dfrac{\partial}{\partial \mathbf{u}}\dfrac{\partial H}{\partial \mathbf{x}}\right]^T & \dfrac{\partial}{\partial \mathbf{u}}\dfrac{\partial H}{\partial \mathbf{u}} \end{bmatrix} \tag{2.4-21}$$

Eq. (2.4-20) reduces to

$$\delta^2 H = \tfrac{1}{2}\delta\mathbf{z}^T\mathbf{P}\delta\mathbf{z} = \tfrac{1}{2}\|\delta\mathbf{z}\|^2_{\mathbf{P}} \tag{2.4-22}$$

which is recognized as the standard quadratic form. A positive definite quadratic form is defined as one for which $\delta\mathbf{z}^T\mathbf{P}\,\delta\mathbf{z} > 0$ for all nonzero $\delta\mathbf{z}$. A positive semidefinite matrix, $\mathbf{P}$, is defined as one which has the property that $\delta\mathbf{z}^T\mathbf{P}\,\delta\mathbf{z} \geq 0$ for all nonzero $\delta\mathbf{z}$. In a similar fashion, negative definite and negative semidefinite quadratic forms and matrices are defined. Section 1.23 of Appendix A delineates a method which we can use to discern positive definiteness of a square matrix. Thus we can state the two necessary conditions [4] for $J(\mathbf{x},\mathbf{u})$ to have an extremum in a given interval of x for convex or concave $J(\mathbf{x},\mathbf{u})$. If $J(\mathbf{x},\mathbf{u})$ is not convex or concave, the second condition is only sufficient, and a quantity known as the bordered Hessian must be used to obtain the second necessary condition.

I. The following vectors are zero:

$$\frac{\partial H}{\partial \mathbf{x}} = \mathbf{0}; \qquad \frac{\partial H}{\partial \mathbf{u}} = \mathbf{0}$$

II. The following matrix

$$\begin{bmatrix} \dfrac{\partial}{\partial \mathbf{x}}\dfrac{\partial H}{\partial \mathbf{x}} & \dfrac{\partial}{\partial \mathbf{u}}\left(\dfrac{\partial H}{\partial \mathbf{x}}\right) \\ \left[\dfrac{\partial}{\partial \mathbf{u}}\dfrac{\partial H}{\partial \mathbf{x}}\right]^T & \dfrac{\partial}{\partial \mathbf{u}}\dfrac{\partial H}{\partial \mathbf{u}} \end{bmatrix}$$

is $\begin{cases}\text{positive semidefinite for a minimum along } \mathbf{f}(\mathbf{x}, \mathbf{u}) = \mathbf{0}\\ \text{negative semidefinite for a maximum along } \mathbf{f}(\mathbf{x}, \mathbf{u}) = \mathbf{0}\end{cases}$

A sufficient condition for a function to have a minimum (maximum) given that the first variation vanishes is that the second variation be positive (negative) where the first variation vanishes [4]. These conditions are general and need be modified only if the possibility of a singular solution exists.

Example 2.4-1

Suppose that we have a linear system represented by

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{c} = \mathbf{0}$$

and wish to find the m vector $\mathbf{u}$ which minimizes

$$J(\mathbf{x}, \mathbf{u}) = \tfrac{1}{2}\|\mathbf{u}\|^2_{\mathbf{R}} + \tfrac{1}{2}\|\mathbf{x}\|^2_{\mathbf{Q}} \qquad \frac{1}{2}\ x^T Q x$$

where $\mathbf{A}$ is an $n \times n$ matrix, $\mathbf{B}$ is an $n \times m$ matrix, $\mathbf{x}$, $\mathbf{c}$, and $\mathbf{0}$ are n vectors. $\mathbf{R}$ and $\mathbf{Q}$ are positive definite symmetric matrices of dimensionality $m \times m$ and $n \times n$.

The Hamiltonian function is formed by adjoining the cost function to the given constraint via the Lagrange multiplier technique which gives us

$$H = \tfrac{1}{2}\mathbf{u}^T\mathbf{R}\mathbf{u} + \tfrac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \boldsymbol{\lambda}^T[\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{c}]$$

In order to minimize J, it is necessary that

$$\frac{\partial H}{\partial \mathbf{x}} = \mathbf{Q}\mathbf{x} + \mathbf{A}^T\boldsymbol{\lambda} = \mathbf{0}, \qquad \frac{\partial H}{\partial \mathbf{u}} = \mathbf{R}\mathbf{u} + \mathbf{B}^T\boldsymbol{\lambda} = \mathbf{0}$$

where $\boldsymbol{\lambda}$ is to be adjusted so that the given equality constraint is satisfied, or

$$\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{c} = \mathbf{0}$$

Thus we find that $u = -R^{-1}B^T(AQ^{-1}A^T + BR^{-1}B^T)^{-1}c = -(R + B^TA^{-T}QA^{-1}B)^{-1}B^TA^{-T}QA^{-1}c$

~~$\mathbf{u} = -(\mathbf{R} + \mathbf{B}^T\mathbf{A}^{-T}\mathbf{Q}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{A}^{-T}\mathbf{Q}\mathbf{A}^{-1}\mathbf{c}$~~ In order to determine whether

is the optimum $\mathbf{u}$ vector. ~~We notice that it is necessary that the inverse of A exist in order for the u vector to exist. To check~~ if this solution does in fact cause $J(\mathbf{x}, \mathbf{u})$ to have a minimum, we find the second variation and check the necessary condition II given earlier. From Eq. (2.4-~~19~~20) and the specifications for this problem, we have

$$\delta^2 J = \frac{1}{2}[\delta\mathbf{x}^T\ \delta\mathbf{u}^T]\begin{bmatrix}\mathbf{Q} & \mathbf{0}\\ \mathbf{0} & \mathbf{R}\end{bmatrix}\begin{bmatrix}\delta\mathbf{x}\\ \delta\mathbf{u}\end{bmatrix} = \frac{1}{2}\,\delta\mathbf{x}^T\mathbf{Q}\,\delta\mathbf{x} + \frac{1}{2}\,\delta\mathbf{u}^T\mathbf{R}\,\delta\mathbf{u}$$

For $J(\mathbf{x}, \mathbf{u})$ t ve a minimum, $\delta^2 J \geq 0$, it is sufficient ~~therefore~~ Q and R ~~must~~ be non-negative definite. ~~Since this is given in the statement of the problem, the solution, if it exists, does minimize $J(\mathbf{x}, \mathbf{u})$.~~ A further requirement is obtained by noting that the first variation of $f(x,u)=0$ yields $A\delta x + B\delta u = 0$ and it is therefore only necessary, for $\delta^2 J > 0$, that $R + B^T A^{-T} Q A^{-1} B$ be positive definite.

Example 2.4-2

Suppose that we wish to minimize the cost function

$$J = \tfrac{1}{2}\|\mathbf{x}\|^2_{\mathbf{Q}}$$

subject to the constraint

$$\mathbf{x} + \mathbf{b}u + \mathbf{c} = \mathbf{0}$$

where the scalar control is bounded such that $|u| \leq 1$.

This problem can be solved without the magnitude constraint on the control with the result (from the last example)

$$u = -(\mathbf{b}^T\mathbf{Q}\mathbf{b})^{-1}\mathbf{b}^T\mathbf{Q}\mathbf{c}$$

If $|u|$ obtained from the foregoing problem is less than 1, we obtain what is called a singular solution. This is so because the H function is linear in the control variable and $\partial H/\partial u = \boldsymbol{\lambda}^T\mathbf{b} = 0$ is the equation for a stationary point which may well be a minimum. If $\mathbf{b}^T\mathbf{Q}\mathbf{b}$ is positive definite, it is at least a local minimum. If the value of u obtained is within the boundary, that value solves our problem. If the value obtained is greater in magnitude than 1, the true solution for u must be on the boundary. This type of problem is of concern in optimal control theory and will be considered in some detail for dynamic processes.

Example 2.4-3 [2]

Suppose that observations of a constant vector are taken after being corrupted with noise. Symbolically, we express this as

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{v}$$

where $\mathbf{z}$ which is composed of observed numbers is an m vector, $\mathbf{H}$ is an $m \times n$ matrix, $\mathbf{x}$ is an n vector, and $\mathbf{v}$ is an m vector representing measurement noise. It is desired to obtain the best estimate of $\mathbf{x}$, denoted $\hat{\mathbf{x}}$, such that

$$J = \tfrac{1}{2}\|\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}\|^2_{\mathbf{R}^{-1}}$$

is minimum where $\mathbf{R}$ is a symmetric positive definite matrix. We accomplish this by setting

$$\frac{\partial J}{\partial \hat{\mathbf{x}}} = \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}) = \mathbf{0}$$

Thus to obtain the best least-square error estimate of $\mathbf{x}$ we have

$$\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{z}$$

One of the simplest cases of interest occurs when we take m estimates of a scalar. In that case it is reasonable to take $\mathbf{H}$ as a unit vector of dimension m or, in other words, a column vector of 1's, and $\mathbf{R}$ as the identity matrix. For this simplest case, we have for the "best" estimate of x

$$\hat{x} = \frac{\mathbf{H}^T\mathbf{z}}{m} = \frac{1}{m}\sum_{i=1}^{m} z_i$$

which is the well-known expression for the average of a num of observations.

Another interesting case occurs when we have computed for r measurements and someone gives us an additional measurement. A great deal of effort would be involved in multiplying and inverting $\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$ if $\mathbf{H}$ is, say, a 1000 by 20 matrix. To repeat this procedure for a new 1001 by 20 matrix would probably be prohibitive of computer time, particularly if "on-line" computation is a requirement. We are thus led to seek a solution which allows us to add the new measurement without repeating the entire calculation. A method which allows us to do this is called a recursive or sequential estimation scheme. Such schemes are of considerable importance in modern system theory and will be explored in much more detail in Chapters 10 and 15.

Assume a set of measurements is represented by

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{v}$$

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \cdot \\ \cdot \\ \cdot \\ z_m \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & & & \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ h_{m1} & & \cdots & h_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_m \end{bmatrix}$$

where $\hat{\mathbf{x}}_m$ is given by $(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{z}$. Now suppose that we obtain an additional measurement such that we have

$$\begin{bmatrix} \mathbf{z} \\ \hline z_{m+1} \end{bmatrix} = \begin{bmatrix} \mathbf{H} \\ \hline \mathbf{h}^T \end{bmatrix} [\hat{\mathbf{x}}_m + \Delta\mathbf{x}] + \begin{bmatrix} \mathbf{v} \\ \hline v_{m+1} \end{bmatrix}$$

The problem now becomes one of obtaining the best estimate of $\mathbf{x}$, $\hat{\mathbf{x}}_{m+1}$, such that

$$J = \frac{1}{2}\left\| \begin{bmatrix} \mathbf{z} \\ \hline z_{m+1} \end{bmatrix} - \begin{bmatrix} \mathbf{H} \\ \hline \mathbf{h}^T \end{bmatrix} \hat{\mathbf{x}}_{m+1} \right\|^2$$

is minimum. Following a procedure similar to the previous one, we find the best estimate of $\mathbf{x}$ is

$$\hat{\mathbf{x}}_{m+1} = \left(\begin{bmatrix} \mathbf{H} \\ \hline \mathbf{h}^T \end{bmatrix}^T \begin{bmatrix} \mathbf{H} \\ \hline \mathbf{h}^T \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{H} \\ \hline \mathbf{h}^T \end{bmatrix}^T \begin{bmatrix} \mathbf{z} \\ \hline z_{m+1} \end{bmatrix}$$

where for convenience we will now assume that the matrix $\mathbf{R}$ is an identity matrix. This amounts to placing equal weight on each measurement. A recursive scheme may be developed by the use of the matrix inversion lemma [2, 3]. We recall that

$$\left\{ \begin{bmatrix} \mathbf{H} \\ \hline \mathbf{h}^T \end{bmatrix}^T \begin{bmatrix} \mathbf{H} \\ \hline \mathbf{h}^T \end{bmatrix} \right\}^{-1} = [\mathbf{H}^T\mathbf{H} + \mathbf{h}\mathbf{h}^T]^{-1}$$

If we define

$$\mathbf{P}_m^{-1} = \mathbf{H}^T\mathbf{H}, \qquad \mathbf{P}_{m+1}^{-1} = \begin{bmatrix} \mathbf{H} \\ \hline \mathbf{h}^T \end{bmatrix}^T \begin{bmatrix} \mathbf{H} \\ \hline \mathbf{h}^T \end{bmatrix} = \mathbf{P}_m^{-1} + \mathbf{h}\mathbf{h}^T$$

then the r ix inversion lemma

$$\mathbf{P}_{m+1} = \mathbf{P}_m - \mathbf{P}_m\mathbf{h}[\mathbf{h}^T\mathbf{P}_m\mathbf{h} + 1]^{-1}\mathbf{h}^T\mathbf{P}_m$$

which will be developed in Section 10.4-1 in a more general form, yields for the recursion formula

$$\begin{aligned}\hat{\mathbf{x}}_{m+1} &= \mathbf{P}_{m+1}[\mathbf{H}^T\mathbf{z} + \mathbf{h}z_{m+1}]\\ &= \mathbf{P}_m\mathbf{H}^T\mathbf{z} + \mathbf{P}_m\mathbf{h}z_{m+1} - \mathbf{P}_m\mathbf{h}[\mathbf{h}^T\mathbf{P}_m\mathbf{h} + 1]^{-1}\mathbf{h}^T\mathbf{P}_m[\mathbf{H}^T\mathbf{z} + \mathbf{h}z_{m+1}]\\ &= \hat{\mathbf{x}}_m + \mathbf{P}_m\mathbf{h}[\mathbf{h}^T\mathbf{P}_m\mathbf{h} + 1]^{-1}[z_{m+1} - \mathbf{h}^T\hat{\mathbf{x}}_m]\end{aligned}$$

Thus the new estimate is equal to the old plus a linear correction term based on the new data and the old $\mathbf{P}_m$ only. For m estimates of a scalar x with H as a unit vector of dimension m, we have

$$\mathbf{P}_m^{-1} = m, \qquad \mathbf{P}_{m+1} = \frac{1}{m+1}, \qquad \mathbf{x}_m = \frac{1}{m}\sum_{i=1}^{m} z_i$$

$$\hat{\mathbf{x}}_{m+1} = \hat{\mathbf{x}}_m + \frac{1}{m+1}[z_{m+1} - \hat{\mathbf{x}}_m] = \hat{\mathbf{x}}_m\left[\frac{m}{m+1}\right] + \frac{z_{m+1}}{m+1}$$

which is, of course, the expected answer in this simple case.

2.5 Linear and nonlinear programming

The previous section contains several examples of what are commonly called nonlinear programming problems. Basically, the nonlinear programming problem is concerned with the extremization of a continuous differentiable function of n nonnegative variables $\theta(x_1, x_2, \ldots, x_n) = \theta(\mathbf{x})$ subject to m inequality constraints $\Lambda_i(\mathbf{x}) \leq 0$, $i = 1, 2, \ldots, m$. Figure 2.5-1 illustrates some basic ideas in a nonlinear programming problem. In nonlinear programming, the θ function is called an objective function—the function to be extremized. In this book we will commonly call such functions *cost functions.*

As we have seen, ordinary calculus methods may be used to find the extremum of unconstrained functions. If ordinary calculus is applied to extremize θ, and if the resulting optimum vector x lies entirely within the constraint set $\Lambda_i \leq 0$, and if $x_i \geq 0$, then that value of x solves the optimization problem with the constraint. We have seen examples of this in Section 2.2 and Example 2.4-2. If the optimum value of x computed by extremizing θ is outside the constraint set $\mathbf{\Lambda} \leq \mathbf{0}$ then the optimum value of x lies on the boundary of the constraint set. If we knew which one of the m constraints Λ determined the optimum, then we could apply the Lagrange multiplier method and use an equality sign for that particular constraint and ignore the other constraints since the optimum x will be on the boundary of one of the known m inequality constraints. In general, we find it necessary to exploit each of the inequality constraints to determine which one of the inequality constraints to use. It is possible that more than one of the m inequality constraints will determine the optimum x as illustrated Fig. 2.5-1. We should remark that, in the typical nonlinear programming problem, the functions Λ are convex, which insures that the possible region for an optimum

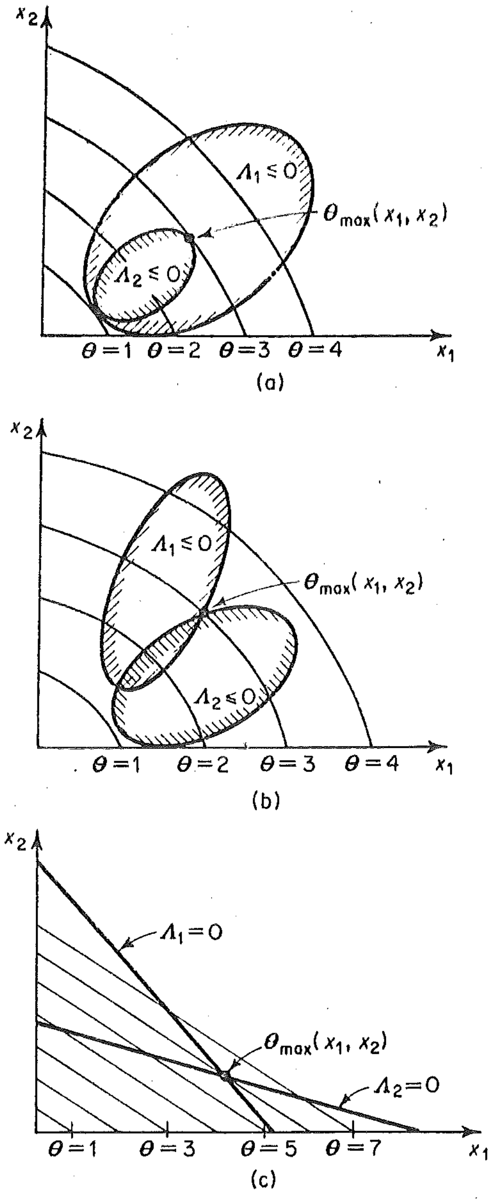


Fig. 2.5-1. Illustrations of nonlinear programming (a, b, c) and linear programming (c).

x is also cor , Also, θ is convex if minimization is required and concave if maximization is required. This requires that any local optimum is a global optimum of the cost function in the possible region of a constraint Λ_i [4].

A special case of the nonlinear programming problem is the linear programming problem which occurs when the θ and Λ functions are linear in the n vector **x**. In this case we are assured that the optimum value of **x** lies on the boundary of two or more elements of the linear constraint set $\Lambda(\mathbf{x}) \leq 0$. Clearly, the major problem is to decide which ones. This is a statement of the general linear programming problem. Of several methods available for solving the problem, the most used method appears to be the simplex method [5]. In order to use the method, certain restrictions must be applied. The variables x_i must be nonnegative, the constraints Λ_i must be linear equalities, and the cost function must be minimized by the optimum **x**.

We may transform the general problem of linear programming, that of maximizing the cost function (objective function)

$$J = \mathbf{a}^T\mathbf{x} \tag{2.5-1}$$

J is a linear function of x, i.e., J is a linear combination of the x_i's

with the m inequality constraints

$$\mathbf{Bx} \leq \mathbf{c} \tag{2.5-2}$$

into the restrictive form for the simplex method. Any number can be written as the difference of two nonnegative numbers. For instance, if x_1 has no restrictions on its sign, we may let

$$x_{n+1} - x_{n+2} = x_1, \qquad x_{n+1} \geq 0, \qquad x_{n+2} \geq 0$$

This insures the nonnegativity of the variables. Unfortunately, every substitution of this type replaces one variable (x_1) by two variables (x_{n+1} and x_{n+2}). If the original problem formulation contains inequality constraints, we convert them to equality constraints by the introduction of nonnegative slack variables. For example, if we had the constraints

$$2x_1 + 4x_2 + x_3 \geq 5, \qquad 6x_1 + x_2 + x_3 \leq 4$$

we would introduce the nonnegative variables x_4 and x_5 to obtain equalities

$$2x_1 + 4x_2 + x_3 - x_4 = 5, \qquad 6x_1 + x_2 + x_3 + x_5 = 4$$

The variables x_4 and x_5 "take up the slack" in the inequalities and are called slack variables. Again, we increase the total number of variables to be considered. The linear programming problem may now be solved by the simplex method.

Since we are to be much more concerned with optimization in dynamic systems than static optimization, we will not develop the many theorems of linear and nonlinear programming. References [4] and [5] contain thorough discussions of both of these topics. We will consider nun al methods for the optimization of single-stage decision processes in Section 13.3-1.

The extrema-finding techniques of this chapter, although quite sufficient for many different situations, will not, in general, allow the solution to many problems associated with control systems. Whereas the previously discussed techniques deal with methods for extremizing functions of one or several independent variables, in control-system design, we are typically concerned with extremizing certain types of functions whose independent variables are actually other functions. This type of function is called a *functional*. Although, as we might expect, many of the basic approaches for extremizing functionals are similar to those for extremizing functions, the end results are sometimes quite different. The solution to a given problem in extremizing a given function of one variable is, perhaps, a number associated with a coordinate point, while the analogous solution to a functional problem is a number associated with a function. The body of mathematics developed for extremizing functionals is variational calculus. This subject is at the very heart of optimal control theory and is a subject that we will explore in some detail throughout the remainder of this text.

REFERENCES

1. Korn, G.A. and T.M. Korn, *Mathematical Handbook for Engineers and Scientists.* McGraw-Hill Book Co., New York, 1962.
2. ✓ Ho, Y.C., "Method of Least Squares and Optimal Filtering Theory." Rand Memo RM 3329 PR, 1962.
3. Lee, R.C.K., *Optimal Estimation Identification and Control.* Technology Press, Cambridge, Mass., 1964.
4. Saaty, T.L., and J. Bram, *Nonlinear Mathematics.* McGraw-Hill Book Co., New York, 1964.
5. Dantzig, G.B., *Linear Programming and Extensions.* Princeton University Press, Princeton, N.J., 1963.

PROBLEMS

0 **1.** Find u such that

$$J = x^2 + u^2$$

is minimized subject to the equation

$$xu = 1$$

Use the Lagrange multiplier technique as well as the basic method.

2. Discuss the singular solution problem where **x** is a two vector.

0 **3.** Find $\hat{x}_0$ for a set of measurements where $\mathbf{z} = \mathbf{Hx}$, where

$$\mathbf{z} = \begin{bmatrix} 1.01 \\ 2.03 \\ 3.00 \\ 3.05 \\ 1.95 \\ 0.97 \end{bmatrix} \qquad \mathbf{H} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

4. Now suppose that an additional measurement

$$z_7 = 3.0; \qquad \mathbf{h}^T = [1, 1]$$

is taken. Compute $\hat{x}_7$ by the smoothing method and the matrix inversion lemma method. Compare the effort involved via each method.

5. Verify the matrix inversion lemma if

$$\mathbf{P}_{r+1}^{-1} = \mathbf{P}_r^{-1} + \mathbf{h}\mathbf{h}^T$$
$$\mathbf{P}_{r+1} = \mathbf{P}_r - \mathbf{P}_r\mathbf{h}(\mathbf{h}^T\mathbf{P}_r\mathbf{h} + 1)^{-1}\mathbf{h}^T\mathbf{P}_r$$

by showing that

$$\mathbf{P}_{r+1}^{-1}\mathbf{P}_{r+1} = \mathbf{I}$$

6. From Eqs. (2.4-14) and (2.4-17) calculate the third variation of H as given in Eq. (2.4-9).

7. Find the maximum value of

$$\theta(\mathrm{x}) = x_1^2 + x_2^2, \qquad x_1 \geq 0, \qquad x_2 \geq 0$$

subject to the inequality constraints

$$(x_1 - 4)^2 + x_2^2 \leq 1$$
$$(x_1 - 1)^2 + x_2^2 \leq 4$$

8. Find the maximum value of

$$J = x_1 + x_2, \qquad x_1 \geq 0, \qquad x_2 \geq 0$$

subject to the constraints

$$x_1 + \tfrac{1}{2}x_2 \leq 1$$
$$\tfrac{1}{2}x_1 + x_2 \leq 2$$

9. Two alternate expressions were developed for the optimum u vector of Example (2.4-1). Show that the two expressions are equivalent and that the first solution will be easier to implement computationally if the dimension of u is lower than that of x.

3

VARIATIONAL CALCULUS AND CONTINUOUS OPTIMAL CONTROL

In this chapter we will introduce the subject of the variational calculus through a derivation of the Euler-Lagrange equations and associated transversality conditions. The existence of the definite integrals defining the cost function is assumed, and it is further understood that minimizing (maximizing) functions are to be chosen from the set of all functions having continuous second derivatives on the time interval under consideration. In addition, we will assume that the integral of the cost function is at least twice continuously differentiable. Thus, this chapter will deal with most of the basic concepts necessary for solving the types of variational problems commonly classified as control-system problems. Several such examples of continuous control problems will be solved. Many of the restrictions posed here will be removed in the next chapter.

3.1 Dynamic optimization without constraints

We will now examine a functional of the simple form where t_0 and t_f are fixed

$$J\langle x\rangle = \int_{t_0}^{t_f} \phi[x(t), \dot{x}(t), t]\, dt \qquad (3.1\text{-}1)$$

Problems of remization of this functional form are sometimes called Lagrange problems. These include the Bolza problem

$$J^1\langle x\rangle = \theta[x(t), t]\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \phi[x(t), \dot{x}(t), t]\, dt \qquad (3.1\text{-}2)$$

The inclusion is apparent if Eq. (3.1-2) is rewritten in the form

$$J^1\langle x\rangle = \int_{t_0}^{t_f} \Lambda\, [x(t), \dot{x}(t), t]\, dt \qquad (3.1\text{-}3)$$

where

$$\Lambda[x(t), \dot{x}(t), t] = \phi[x(t), \dot{x}(t), t] + \frac{d}{dt}\theta[x(t), t] \qquad (3.1\text{-}4)$$

We would now like to find an $x(t)$ such that the given $J\langle x\rangle$ is extremized (i.e., maximized or minimized, depending on the given physical problem). This $x(t)$ is called an extremal, and only an extremal can cause $J\langle x\rangle$ to have an extremum. We will assume that we know the correct extremal curve, denoted $\hat{x}(t)$. Thus we can write the expression (3.1-5) for a family of curves, starting at $t = t_o$ and ending at $t = t_f$, which includes the extremal curve $\hat{x}(t)$.

$$x(t) = \hat{x}(t) + \epsilon\eta(t) \qquad (3.1\text{-}5)$$

where $\eta(t)$ is a variation in $x(t)$ and ϵ is a small number. A plot of $J\langle x\rangle$ versus ϵ for various choices of $\eta(t)$ might appear as shown in Fig. 3.1-1. It is obvious that at $\epsilon = 0$, all curves are minimum since

$$\hat{x}(t) = x(t)\,|_{\epsilon=0} \qquad (3.1\text{-}6)$$

Thus on the extremals we have

$$\frac{\partial J\langle x\rangle}{\partial\epsilon}\bigg|_{\epsilon=0} = 0 \qquad (3.1\text{-}7)$$

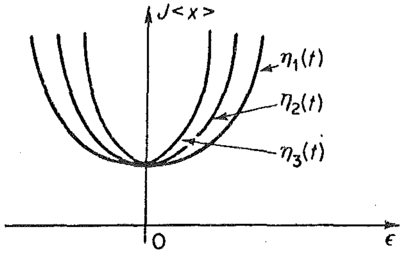


Fig. 3.1-1. Minimization problem of variational calculus.

independent of the value of $\eta(t)$ chosen. Strictly speaking, the solution obtained from Eq. (3.1-5) could cause $J\langle x\rangle$ to have a maximum or minimum or be a stationary point. The condition for a minimum is that $\partial^2 J/\partial\epsilon^2$ be positive at $\epsilon = 0$ independent of $\eta(t)$. However, in most physical problems, it is apparent that if a solution to Eq. (3.1-7) exists, it will be a solution which minimizes (maximizes) the integral, $J\langle x\rangle$, as desired. Now we can extremize Eq. (3.1-1) by using Eqs. (3.1-5) and (3.1-7). By differentiating Eq. (3.1-5) with respect to t, we obtain

$$\dot{x}(t) = \dot{\hat{x}}(t) + \epsilon\dot{\eta}(t) \qquad (3.1\text{-}8)$$

If we substitute Eqs. (3.1-5) and (3.1-8) into the given f tional (3.1-1), we then have

$$J\langle x\rangle = \int_{t_0}^{t_f} \phi[\hat{x}(t) + \epsilon\eta(t), \dot{\hat{x}}(t) + \epsilon\dot{\eta}(t), t]\, dt \qquad (3.1\text{-}9)$$

We should note that

$$\lim_{\epsilon=0} J(x) = J(\hat{x}), \qquad \lim_{\epsilon=0} x(t) = \hat{x}(t)$$

Therefore, to find the extremals of $J\langle x\rangle$ we now use Eq. (3.1-7)†

$$\frac{\partial J\langle x\rangle}{\partial\epsilon}\bigg|_{\epsilon=0} = \int_{t_0}^{t_f}\left\{\eta(t)\,\frac{\partial\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\hat{x}} + \dot{\eta}(t)\,\frac{\partial\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\dot{\hat{x}}}\right\} dt = 0$$

or

$$0 = \int_{t_0}^{t_f} \eta(t)\,\frac{\partial\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\hat{x}}\, dt + \int_{t_0}^{t_f} \dot{\eta}(t)\,\frac{\partial\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\dot{\hat{x}}}\, dt \qquad (3.1\text{-}10)$$

After simplification Eq. (3.1-10) becomes,‡

$$0 = \int_{t_0}^{t_f} \eta(t)\left[\frac{\partial\phi}{\partial\hat{x}} - \frac{d}{dt}\,\frac{\partial\phi}{\partial\dot{\hat{x}}}\right] dt + \frac{\partial\phi}{\partial\dot{\hat{x}}}\,\eta(t)\Big|_{t_0}^{t_f} \qquad (3.1\text{-}11)$$

Since Eq. (3.1-11) must equal zero independent of the value chosen for $\eta(t)$, we have

Euler Lagrange Eqn. $$\frac{\partial\phi}{\partial\hat{x}} - \frac{d}{dt}\,\frac{\partial\phi}{\partial\dot{\hat{x}}} = 0 \qquad (3.1\text{-}12)$$

transversality condition $$\frac{\partial\phi}{\partial\dot{\hat{x}}}\eta(t) = 0, \qquad \text{for} \quad t = t_o, t_f \qquad (3.1\text{-}13)$$

---

†The following is given without proof: If $u = f(x, y, z, \ldots)$ is a function of several variables, each of which is a differentiable function of $r, v, w, \ldots$, then u as a function of these new independent variables, is differentiable, and the following chain rule applies

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial r} + \cdots$$

$$\frac{\partial u}{\partial v} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial v} + \cdots$$

‡Applying the formula for integration by parts, which is

$$\int_a^b u\, dv = uv\Big|_a^b - \int_a^b v\, du$$

by letting

$$u = \frac{\partial\phi}{\partial\dot{\hat{x}}} \qquad dv = \dot{\eta}(t)\, dt$$

$$du = \frac{d}{dt}\,\frac{\partial\phi}{\partial\dot{\hat{x}}}\, dt \qquad v = n(t)$$

we have

$$\int_{t_0}^{t_f} \dot{\eta}(t)\,\frac{\partial\phi}{\partial\dot{\hat{x}}}\, dt = \eta(t)\,\frac{\partial\phi}{\partial\dot{\hat{x}}}\Big|_{t_0}^{t_f} - \int_{t_0}^{t_f} \eta(t)\,\frac{d}{dt}\,\frac{\partial\phi}{\partial\dot{\hat{x}}}\, dt$$

These relationships follow as a consequence of the following lemma.

If $x(t)$ is continuous on the closed interval $t \in [t_1, t_2]$ and if $\int_{t_1}^{t_2} x(t)\eta(t)\,dt = 0$ for every $\eta(t)$ contained in $[t_1, t_2]$ such that $\eta(t_1) = \eta(t_2) = 0$, then $x(t) = 0$ for all t in $[t_1, t_2]$. Proof of this lemma is given in reference [1].

These two very important relationships form a good foundation for solving variational problems. Equation (3.1-12) is commonly known as the Euler-Lagrange equation and Eq. (3.1-13) is the associated transversality condition. These equations specify a two-point boundary value differential equation which, when solved, determines $\hat{x}$ in terms of a known ϕ.

3.2 Remarks on transversality conditions

The various forms and uses of the transversality conditions will be covered in some detail in this chapter. We do this because these conditions are among the hardest things to correctly formulate for any variational problem, and they are generally different enough for each problem to warrant comment.

We will now examine Eq. (3.1-13) and tabulate many of the possible combinations for which this equation holds. In each case, t_o and t_f are fixed.

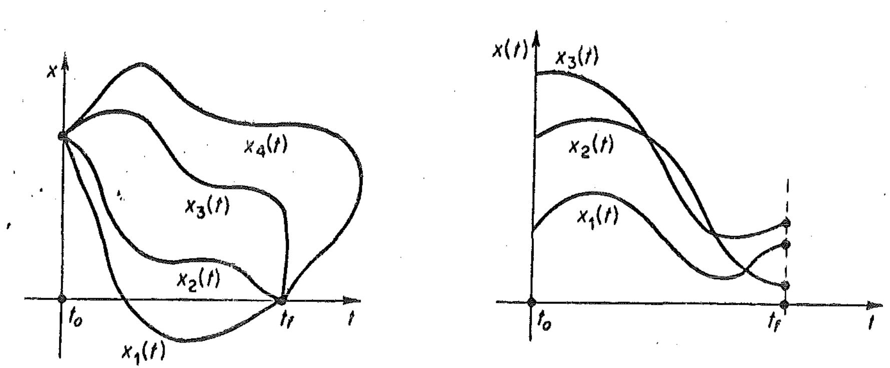
I. Fixed Beginning—Terminal Points
In this case we fix $x(t_o)$ and $x(t_f)$. Thus every admissible solution must pass through these fixed points. Therefore from Eq. (3.1-4) (3.1-5) we see that we must require that $\eta(t_o) = \eta(t_f) = 0$. In this case the correct boundary conditions are the specified $x(t_o)$ and $x(t_f)$.

II. Variable Beginning—Terminal Points
We now consider that $x(t_o)$ and $x(t_f)$ are variable or, in other words, not constrained. Therefore from Eq. (3.1-13) we have (since $\eta(t)$ can be arbitrary at the end points) $\partial\phi/\partial\dot{x} = 0$ at $t = t_o$ and $t = t_f$. When this particular situation results, the boundary conditions are called the natural boundary conditions.

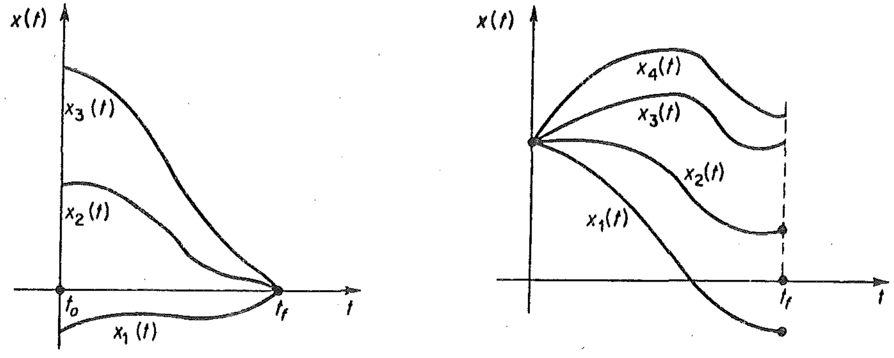
III. Variable Beginning—Fixed Terminal Points
In the case where $x(t_o)$ is variable and $x(t_f)$ is fixed, we must constrain $\eta(t_f)$ to be zero but can allow any (admissible) $\eta(t_o)$. Therefore from Eq. (3.1-13) we have the two-point boundary conditions $\partial\phi/\partial\dot{x} = 0$ at $t = t_o$, and $\eta(t_f) = 0$, which means that the other boundary condition is the specified $x(t_f)$.

IV. Fixed Beginning—Variable Terminal Points
For $x(t_o)$ fixed and $x(t_f)$ variable, a situation which often occurs in optimal control, we have from Eq. (3.1-13) that (since $\eta(t_f)$ is arbitrary) the two-point boundary conditions are the specified $x(t_o)$ and $\partial\phi/\partial\dot{x} = 0$ at $t = t_f$.

With this tabulation, the analysis of the scalar Lagrange problem (which, as previously mentioned, includes the scalar Bolza problem) is nearly complete. Figure 3.2-1 illustrates graphically the essence of this tabulation.



(a) Case I (b) Case II



(c) Case III (d) Case IV

Fig. 3.2-1. Various combinations of end conditions.

3.3 The second variation: sufficient conditions for (weak) extrema

weak means relative

Until now, in the study of extrema of functionals we have only considered a necessary condition for a functional to have a relative or weak extremum. This was, of course, the condition that the first variation vanish. In this section, we shall be briefly concerned with sufficient conditions for a function to have extrema and shall thus introduce the second variation. The next section on examples will illustrate the application of the second variation in a particularly simple case.

To establish the nature of an extremum, it is necessary to obtain $\partial^2 J/\partial\epsilon^2$ evaluated at $\epsilon = 0$ from Eq. (3.1-1) under the conditions of Eq. (3.1-5). This is

$$\left.\frac{\partial^2 J\langle x\rangle}{\partial\epsilon^2}\right|_{\epsilon=0} = \int_{t_0}\left\{\eta^2 \frac{\partial^2\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\hat{x}^2} + 2\dot{\eta}\eta \frac{\partial^2\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\hat{x}\partial\dot{\hat{x}}} + \dot{\eta}^2 \frac{\partial^2\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\dot{\hat{x}}^2}\right\} dt \tag{3.3-1}$$

Applying integration by parts and the transversality conditions [Eq. (3.1-13)] we have

does not imply

$$\left.\eta^2 \frac{\partial^2 F}{\partial y\,\partial y'}\right|_{t_0}^{t_f}$$

= 0

$$2\int_{t_0}^{t_f} \eta\dot{\eta} \frac{\partial^2\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\hat{x}\partial\dot{\hat{x}}} dt = -\int_{t_0}^{t_f}\left\{\frac{d}{dt} \frac{\partial^2\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\hat{x}\,\partial\dot{\hat{x}}}\right\}\eta^2\, dt \tag{3.3-2}$$

not necessary

No!

Thus the second variation of J becomes

$$\left.\frac{\partial^2 J\langle x\rangle}{\partial\epsilon^2}\right|_{\epsilon=0} = \int_{t_0}^{t_f}\left\{\eta^2\left[\frac{\partial^2(\hat{x}, \dot{\hat{x}}, t)}{\partial\hat{x}^2} - \frac{d}{dt} \frac{\partial^2\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\hat{x}\,\partial\dot{\hat{x}}}\right] + \dot{\eta}^2 \frac{\partial^2\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\dot{\hat{x}}^2}\right\} dt \tag{3.3-3}$$

To establish a minimum (maximum) of J, the first necessary condition is that $\partial J/\partial\epsilon = 0$ at $\epsilon = 0$ independently of the variation $\eta(t)$. The second necessary condition for a minimum (maximum) is that the second derivative of J with respect to ϵ, evaluated at $\epsilon = 0$, be equal to or greater than (equal to or less than) zero. Sufficient conditions for a weak minimum (maximum) require that the derivative be positive (negative). All of this must, of course, be true independent of the variation $\eta(t)$ and need only be true along the optimal "trajectory," $\hat{x}(t)$.

We can rewrite Eq. (3.3-1) as the quadratic form integral

$$\left.\frac{\partial^2 J(x)}{\partial\epsilon^2}\right|_{\epsilon=0} = \int_{t_0}^{t_f} [\eta(t)\dot{\eta}(t)]\begin{bmatrix}\dfrac{\partial^2\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\hat{x}^2} & \dfrac{\partial^2\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\hat{x}\,\partial\dot{\hat{x}}} \\ \dfrac{\partial^2\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\hat{x}\,\partial\dot{\hat{x}}} & \dfrac{\partial^2\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\dot{\hat{x}}^2}\end{bmatrix}\begin{bmatrix}\eta(t) \\ \dot{\eta}(t)\end{bmatrix} dt \tag{3.3-4}$$

If the matrix in this expression is at least positive (negative) semidefinite, we have certainly established a minimum (maximum). Alternately, from Eq. (3.3-3) we are assured that the second derivative is equal to or greater than zero if

$$\frac{\partial^2\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\hat{x}^2} - \frac{d}{dt}\left[\frac{\partial^2\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\hat{x}\,\partial\dot{\hat{x}}}\right] \geq 0 \tag{3.3-5}$$

and

$$\frac{\partial^2\phi(\hat{x}, \dot{\hat{x}}, t)}{\partial\dot{\hat{x}}^2} \geq 0 \tag{3.3-6}$$

For many problems in which we will have interest, the foregoing conditions are fulfilled, and we can establish necessary and sufficient conditions for a minimum. It is still possible, however, for Eq. (3.3-1) or Eq. (3.3-2) to be greater than zero even if the requirements of Eqs. (3.3-4), (3.3-5), and (3.3-6) are not satisfied, since $\eta(t)$ and $\dot{\eta}(t)$ are not independent of one another. Complete exploitation of this point is beyond the intent of this chapter. Chapters 5 and 6 of reference [1] provide an excellent and readable discussion of the necessary and sufficient conditions for a minimum. We will return again to this point in Chapter 4. We must again emphasize here that we are establishing conditions for a relative extremum, sometimes called a weak extremum, which may or may not be an absolute extremum. In Section 4.1 we will discuss some requirements for an absolute or strong extremum.

Example 3.3-1

We desire to find the curve with minimum arc length between the point $x(0) = 1$ and the line $t_f = 2$.

The first step toward solving this problem is to formulate the functional $J\langle x\rangle$. If we define the differential arc length as ds, the functional we desire to minimize is easily seen to be

$$J\langle x\rangle = \int_0^2 ds$$

with associated boundary conditions

$$x(t = 0) = 1, \qquad x(t = 2) = \text{open}$$

Noting that for a differential arc length

$$(ds)^2 = (dx)^2 + (dt)^2$$

we have

$$\frac{ds}{dt} = [1 + \dot{x}^2]^{1/2}$$

By substituting into the given cost function, we obtain

$$J\langle x\rangle = \int_0^2 [1 + \dot{x}^2]^{1/2}\, dt$$

Upon referring back to the functional defined in Eq. (3.1-1), we see that

$$\phi(x, \dot{x}, t) = [1 + \dot{x}^2]^{1/2}$$

The Euler-Lagrange equation for this problem is therefore

$$\frac{\partial\phi}{\partial\hat{x}} - \frac{d}{dt}\frac{\partial\phi}{\partial\dot{\hat{x}}} = 0$$

and thus we obtain

$$\frac{-d}{dt}\left[\frac{\dot{\hat{x}}}{(1 + \dot{\hat{x}}^2)^{1/2}}\right] = 0$$

Upon integrating, we obtain

$$\frac{\dot{\hat{x}}}{(1 + \dot{\hat{x}}^2)^{1/2}} = c = \text{constant}, \qquad \dot{\hat{x}}^2 = \frac{c^2}{1 - c^2} = a^2$$

Thus we see that the extremal curve is given by

$$\hat{x}(t) = at + b$$

Therefore, shortest distance between a point and a straight line is another straight line.

We obtain the particular solution by properly applying the transversality equation to the given boundary conditions. We note that this problem falls into situation IV, i.e., fixed beginning—variable terminal point. Thus, $x(t_o) = x(0) = 1$ and

$$\frac{\partial \phi}{\partial \dot{x}} = 0 = \frac{\dot{x}}{[1 + \dot{x}^2]^{1/2}}, \qquad \text{at} \quad t = 2$$

or $\dot{x} = 0$ at $t = 2$.

Differentiating the solution for $\dot{x}$ with respect to t, we have $\dot{x} = a$, and using the transversality conditions we obtain $a = 0$ and $b = 1$. Therefore, the extremal curve satisfying the given boundary condition and minimizing the given arc length is $x = 1$.

To mathematically demonstrate that we have obtained a minimum rather than a maximum or stationary point, it is necessary to show that the second variation, represented by Eq. (3.3-3), is greater than zero. The pertinent terms in Eq. (3.3-3) are, for this example,

$$\frac{\partial^2 \phi}{\partial x\, \partial \dot{x}} = 0, \qquad \frac{\partial^2 \phi}{\partial \dot{x}^2} = \frac{1}{(1 + \dot{x}^2)^{3/2}} \quad \therefore \frac{\partial^2 \varphi}{\partial \dot{x}} = \frac{1}{(1+0)^{3/2}} = 1 > 0$$

Since $\dot{x} = 0$ is the extremal solution, $\partial^2\phi/\partial\dot{x}^2$ is always greater than zero. Thus the second variation is greater than zero, and we have indeed established a minimum. Physically this was, of course, evident from the start.

Example 3.3-2

We desire to find the equation of the curve which minimizes the functional (boundary conditions unspecified)

$$J\langle x\rangle = \int_0^2 [\tfrac{1}{2}\dot{x}^2 + x\dot{x} + \dot{x} + x]\, dt$$

The Euler-Lagrange equation for this problem is

$$\dot{x} + 1 - \dot{x} - \ddot{x} = 0 = 1 - \ddot{x}$$

By integrating directly, we obtain the solution to this equation:

$$x(t) = \frac{t^2}{2} + C_1 t + C_2$$

To determine C_1 and C_2 we must now apply the transversality equation to the given boundary conditions. Since this is a variable beginning—terminal point problem, situation II is used, which is the natural boundary condition case.

$$\frac{\partial \phi}{\partial \dot{x}} = \dot{x} + x + 1 = 0, \qquad \text{for} \quad t = 0, 2$$

Therefore, from the solution for x and its derivative, we have

$$\frac{\partial \phi}{\partial \dot{x}} = t + C_1 + \frac{t^2}{2} + C_1 t + C_2 + 1 = 0, \qquad \text{for} \quad t = 0, 2$$

We can now solve for C_1 and C_2 from the simultaneous equations

$$C_1 + C_2 = -1, \qquad 3C_1 + C_2 = -5$$

to obtain $C_1 = -2$ and $C_2 = 1$. Therefore the extremal cur which satisfies the given boundary conditions, is

$$x(t) = \frac{t^2}{2} - 2t + 1$$

The actual value of the extremum is obtained when we substitute into the given cost function and carry out the integration to obtain $J_{\min} = \frac{4}{3}$. ? $-\frac{4}{3}$

3.4 Unspecified terminal time problems

By slightly changing the cost function given in Eq. (3.1-1) we obtain a very useful problem formulation; it is called an unspecified terminal time problem and, as will be apparent later, leads to the "minimum time" problem of optimal control. The basic problem is one of minimizing a given cost function where t_f is unspecified subject to the constraint that the final state of the system be specified by a prescribed terminal line or, in higher-dimensional problems, terminal manifold.

The cost function generally contains terms representing energy expended, distance traversed, elapsed time, and so forth, which may appear singly or in combination. The original state of the system may be specified or unspecified, and the terminal line or manifold may be time-varying or invariant.

The approach used here will be general enough so that any of the foregoing specifications can be included in the solution of a specific problem. A graphical illustration of a variable terminal time problem is given in Fig. 3.4-1. Instead of calling $J\langle x\rangle$ a functional, we will now use the systems control terminology, cost function, which for this problem will be given by

$$J\langle x\rangle = \int_{t_0}^{t_f} \phi(x, \dot{x}, t)\, dt \tag{3.4-1}$$

where t_0 is known, t_f is unspecified, and $x(t_o)$ may or may not be specified.

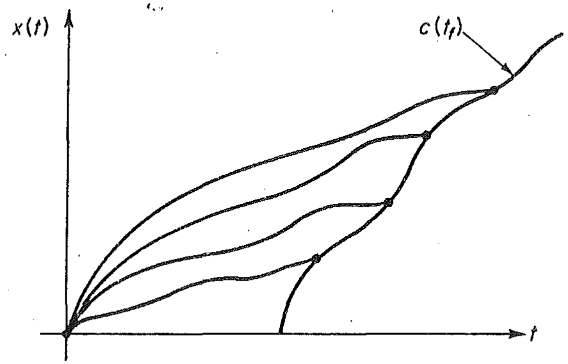


Fig. 3.4-1. Illustration of variable terminal time problem where $x(t_f) = c(t_f)$.

We should note that for the problem shown in Fig. 3.4-1 the initial state, $x(t_0)$, is specified though, in general, as previously stated, it need not be.

As before, $\hat{x}(t)$ is the required curve, here referred to as the optimal system trajectory. A family of curves, which includes the optimal trajectory $\hat{x}(t)$, starting at t_0 and ending at t_f is given by

$$x(t) = \hat{x}(t) + \epsilon\eta_x(t) \tag{3.4-2}$$

with time derivative

$$\dot{x}(t) = \dot{\hat{x}}(t) + \epsilon\dot{\eta}_x(t) \tag{3.4-3}$$

where $\eta_x(t)$ is a variation in x which depends on t.

Since the terminal time is unspecified, it must be treated as a variable and, therefore, must be examined to see if perhaps there is a final time, $\hat{t}_f$, which is optimal. We will therefore define a family of final times, one of which is the optimal final time $\hat{t}_f$:

$$t_f = \hat{t}_f + \epsilon\eta_t(t_f) \tag{3.4-4}$$

where $\eta_t(t_f)$ is a variation in t_f.

Our first step in minimizing the cost function, Eq. (3.4-1), is to substitute Eqs. (3.4-2), (3.4-3), and (3.4-4) into it, which gives us

$$J\langle x\rangle = \int_{t_0}^{\hat{t}_f+\epsilon\eta_t(t_f)} \phi[\hat{x}(t) + \epsilon\eta_x(t), \dot{\hat{x}}(t) + \epsilon\dot{\eta}_x(t), t]\,dt \tag{3.4-5}$$

We now set $\partial J/\partial\epsilon = 0$ at $\epsilon = 0$ and obtain

$$\left.\frac{\partial J}{\partial\epsilon}\right|_{\epsilon=0} = 0 = \int_{t_0}^{\hat{t}_f}\left\{\eta_x(t)\frac{\partial\phi}{\partial\hat{x}} + \dot{\eta}_x(t)\frac{\partial\phi}{\partial\dot{\hat{x}}}\right\}dt + \eta_t(\hat{t}_f)\phi[\hat{x}(\hat{t}_f), \dot{\hat{x}}(\hat{t}_f), \hat{t}_f] \tag{3.4-6}$$

Integrating a portion of Eq. (3.4-6) by parts, we obtain

$$\int_{t_0}^{\hat{t}_f}\eta_x(t)\left[\frac{\partial\phi}{\partial\hat{x}} - \frac{d}{dt}\frac{\partial\phi}{\partial\dot{\hat{x}}}\right]dt + \eta_x\left.\frac{\partial\phi}{\partial\dot{\hat{x}}}\right|_{t=t_0}^{\hat{t}_f} + \eta_t(\hat{t}_f)\phi[\hat{x}(\hat{t}_f), \dot{\hat{x}}(\hat{t}_f), \hat{t}_f] = 0 \tag{3.4-7}$$

At the terminal time, the terminal line, $C(t)$ or, in higher dimensions, terminal manifold, and the optimal trajectory $\hat{x}(t)$ intersect, as shown in Fig. 3.4-1. Therefore, using Eqs. (3.4-2) and (3.4-4), we have

$$\hat{x}[\hat{t}_f + \epsilon\eta_t(t_f)] + \epsilon\eta_x[\hat{t}_f + \epsilon\eta_t(t_f)] = C[\hat{t}_f + \epsilon\eta_t(t_f)] \tag{3.4-8}$$

We take the partial derivative of this equation with respect to ϵ and evaluate it at $\epsilon = 0$ to obtain

$$\eta_t(\hat{t}_f)\dot{\hat{x}}(\hat{t}_f) + \eta_x(\hat{t}_f) = \eta_t(\hat{t}_f)\dot{C}(\hat{t}_f) \tag{3.4-9}$$

where $\dot{\hat{x}}(t) = \partial\hat{x}/\partial t$ and $\dot{C}(t) = \partial C/\partial t$ at $t = \hat{t}_f$. Thus

$$\eta_x(\hat{t}_f) = \eta_t(\hat{t}_f)[\dot{C}(\hat{t}_f) - \dot{\hat{x}}(\hat{t}_f)] \tag{3.4-10}$$

By substituting Eq. (3.4-10) into (3.4-7), we have

$$\begin{aligned}\int_{t_0}\eta_x(t)\left[\frac{\partial\phi}{\partial\hat{x}} - \frac{d}{dt}\frac{\partial\phi}{\partial\dot{\hat{x}}}\right]dt + \eta_t(\hat{t}_f)\Big\{[C(\hat{t}_f) - \dot{\hat{x}}(\hat{t}_f)]\frac{\partial\phi[\hat{x}(\hat{t}_f), \dot{\hat{x}}(\hat{t}_f), \hat{t}_f]}{\partial\dot{\hat{x}}(\hat{t}_f)} \\ + \phi[\hat{x}(\hat{t}_f), \dot{\hat{x}}(\hat{t}_f), \hat{t}_f]\Big\} - \eta_x(t)\left.\frac{\partial\phi}{\partial\dot{\hat{x}}}\right|_{t=t_0} = 0\end{aligned} \tag{3.4-11}$$

Remembering that Eq. (3.4-11) must be identically equal to zero independent of the variations, we see that the first requirement for the solution to our problem (the second variation must also be non-positive) is that

$$\frac{\partial\phi}{\partial x} - \frac{d}{dt}\frac{\partial\phi}{\partial\dot{x}} = 0 \tag{3.4-12}$$

$$\eta_t(t)\left[(\dot{C} - \dot{x})\frac{\partial\phi}{\partial\dot{x}} + \phi\right] = 0, \qquad \text{for} \quad t = \hat{t}_f \tag{3.4-13}$$

$$\eta_x(t)\frac{\partial\phi}{\partial\dot{x}} = 0, \qquad \text{for} \quad t = t_0 \tag{3.4-14}$$

$\partial\dot{\hat{x}}$ $\hat{x}$

We recognize that Eq. (3.4-12) is the familiar Euler-Lagrange equation while Eqs. (3.4-13) and (3.4-14) comprise the transversality conditions for this problem. As before, there are four different relationships obtainable from the transversality conditions, but since they are so similar to those discussed previously, the details of these relationships are left as an exercise. We note that the ^ notation has been removed from Eqs. (3.4-12) through (3.4-14) for convenience. Let us now attempt to apply our results to a simple problem.

Example 3.4-1

We wish to minimize

$$J\langle x\rangle = \int_0^{t_f}[1 + \dot{x}^2]^{1/2}\,dt$$

with $x(0) = 1$ such that $x(t_f) = C(t_f) = 2 - t_f$.

We should recognize that the cost function is actually the arc length, which means that the distance between a point and a line is being minimized. Application of the Euler-Lagrange equation yields the optimal trajectory $x = at + b$, as in Example 3.3-1. To evaluate the arbitrary constants a and b, we make proper use of the transversality Eqs. (3.4-13) and (3.4-14). Here we specify $x(0) = 1$; thus $\eta_x(t_0) = 0$. And since t_f is unspecified, Eq. (3.4-13) becomes

$$(\dot{C} - \dot{x})\frac{\partial\phi}{\partial\dot{x}} + \phi = 0, \qquad \text{at} \quad t = t_f$$

Thus we obtain $\dot{x} = 1$ at the unspecified terminal time t_f. From the solution to the Euler-Lagrange equation and the specified initial condition, we have $x(t = 0) = 1$; so we must have $b = 1$ and $\dot{x}(t = t_f) = a = 1$. Therefore the optimal trajectory is $x(t) = t + 1$, and the final time t_f is $t_f = \frac{1}{2}$. Salient features of this problem are indicated in Fig. 3.4-2. An interesting fact here is that the optimal trajectory intersects the terminal manifold at right angles. In general,

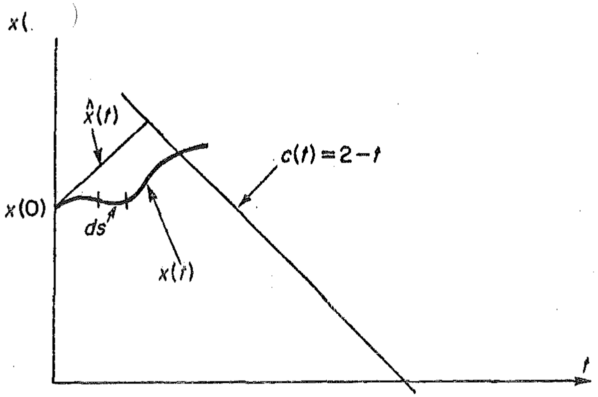


Fig. 3.4-2. Illustration of variable terminal time variable end point problem, Example (3.4-1).

the optimal trajectory will always be nontangent to the terminal manifold. This nontangency condition is, in fact, called the transversality condition.

3.5 Euler-Lagrange equations and transversality conditions—vector formulation

The previous results can be easily generalized to include scalar cost functions in n-dimensional variables via the state-space approach. That is, we desire to minimize

$$J\langle \mathbf{x} \rangle = \int_{t_o}^{t_f} \phi(\mathbf{x}, \dot{\mathbf{x}}, t)\, dt \tag{3.5-1}$$

where $\mathbf{x}$ is the system state, an n vector such that $\mathbf{x}^T = [x_1, x_2, \ldots, x_n]$. t_o, the starting time, is generally specified (it may not be); $\mathbf{x}(t_o)$ may or may not be specified; $\mathbf{x}(t_f)$ is specified by a given terminal manifold denoted $\mathbf{C}(t_f)$.† As before, the terminal time t_f does not have to be known. After following a procedure quite similar to the scalar one, we have, after setting $\partial J/\partial \epsilon$ at $\epsilon = 0$ and dropping the ^ notation the requirement that among other things

$$\int_{t_o}^{t_f} \eta^T(t)\left[\frac{\partial \phi}{\partial \mathbf{x}} - \frac{d}{dt}\frac{\partial \phi}{\partial \dot{\mathbf{x}}}\right] dt = 0 \tag{3.5-2}$$

be true independent of $\eta(t)$. This leads to the requirement that

$$\frac{\partial \phi}{\partial \mathbf{x}} - \frac{d}{dt}\frac{\partial \phi}{\partial \dot{\mathbf{x}}} = 0 \tag{3.5-3}$$

which is simply an extended version of the Euler-Lagrang quation. The associated transversality conditions are given by

$$\eta_{\mathbf{x}}^T \frac{\partial \phi}{\partial \dot{\mathbf{x}}} = 0, \qquad \text{at} \quad t = t_o \tag{3.5-4}$$

$$\eta_{\mathbf{x}}^T \frac{\partial \phi}{\partial \dot{\mathbf{x}}} + \eta_t \phi = 0, \qquad \text{at} \quad t = t_f \tag{3.5-5}$$

where η_t can be related to $\eta_{\mathbf{x}}$ by an equation obtained exactly as Eq. (3.4-8) was obtained

$$\eta_t\left[\frac{d\mathbf{x}}{dt} - \frac{d\mathbf{C}}{dt}\right] + \eta_{\mathbf{x}} = 0 \qquad \text{at } t = \hat{t}_f \tag{3.5-6}$$

Although the notation of this section may appear somewhat cumbersome, in an actual problem it is not, as the next example shows. Use of the Lagrange multiplier technique, as in the next section, will alleviate some of the burdensome notation.

? section no.: 3.7. not 3.6.

Example 3.5-1

We desire to find the transversality conditions for the minimization of

$$J = \int_{t_o}^{t_f} \phi(\mathbf{x}, \dot{\mathbf{x}}, t)\, dt$$

such that $\mathbf{x}(t_f) = \mathbf{C}(t_f)$, where $\mathbf{C}^T(t) = [c_1(t), 0, 0]$ and $\mathbf{x}^T = [x_1, x_2, x_3]$, $\mathbf{x}(t_o) = \mathbf{x}_o$, with t_o specified and t_f unspecified. The Euler-Lagrange equations are

$$\frac{\partial \phi}{\partial x_1} - \frac{d}{dt}\frac{\partial \phi}{\partial \dot{x}_1} = 0, \qquad \frac{\partial \phi}{\partial x_2} - \frac{d}{dt}\frac{\partial \phi}{\partial \dot{x}_2} = 0, \qquad \frac{\partial \phi}{\partial x_3} - \frac{d}{dt}\frac{\partial \phi}{\partial \dot{x}_3} = 0$$

with associated boundary conditions, $\mathbf{x}(t_o) = \mathbf{x}_o$, which represents the initial condition for the two-point boundary value problem, and

at t = t_f

$$(3.4\text{-}15) \longrightarrow \quad \frac{\partial \phi}{\partial \dot{x}_1} + \frac{\phi}{\dot{c}_1 - \dot{x}_1} = 0, \qquad x_2(t) = 0, \qquad \cancel{x_3(t) = 0,} \qquad \text{at} \quad t = t_f \qquad \underline{x}(t) = c(t)$$

from ∂J for $x_1(t)$ only

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Although it may seem that all unspecified terminal time problems may now be worked by mere substitution into the derived relationships, Eqs. (3.5-3) through (3.5-6), this is not the case. Many problems do not fall precisely into a form which allows direct use of our derived formulas. When this type of problem is encountered, a good procedure to follow is to derive the transversality condition for the particular problem. An example demonstrating this type of approach follows.

x_1
x_2
$\rightarrow \{ x_2$

Example 3.5-2

We wish to find the transversality conditions for the minimization of

$$J = \int_{t_o}^{t_f} \phi(\mathbf{x}, \dot{\mathbf{x}}, t)\, dt$$

such that $\|\mathbf{x}(t_f)\|^2 = 1$, where $\mathbf{x}^T = [x_1 x_2]$, with specified starting time t_o and terminal time t_f. Thus, we would like to reach the region of state-space specified

---

†In general, all the states of $\mathbf{x}(t)$ need not be specified at the terminal time. If this is in fact the case for a given problem, great care must be exercised in applying the equations derived for transversality conditions in this section. This point will again be stressed at an appropriate time in the next chapter.

by $x_1^2 + x_2^2 = 1$ specified terminal time t_f given the state at the starting time t_o, denoted by $\mathbf{x}(t_o)$.

The transversality conditions are, from Eq. (3.5-4),

$$\left(\frac{\partial\phi}{\partial\dot{\mathbf{x}}}\right)^T \eta_{\mathbf{x}} = 0 = \frac{\partial\phi}{\partial\dot{x}_1}\eta_{x_1} + \frac{\partial\phi}{\partial\dot{x}_2}\eta_{x_2}, \qquad \text{at} \quad t = t_f$$

As before, we assume that $\mathbf{x}(t) = \hat{\mathbf{x}}(t) + \epsilon\eta_{\mathbf{x}}(t)$ where x is the optimal trajectory. For this problem, this relation in component form becomes $x_1 = \hat{x}_1 + \epsilon\eta_{x_1}$ and $x_2 = \hat{x}_2 + \epsilon\eta_{x_2}$. Substituting these results into the given terminal manifold, we obtain

$$(\hat{x}_1 + \epsilon\eta_{x_1})^2 + (\hat{x}_2 + \epsilon\eta_{x_2})^2 = 1, \qquad \text{at} \quad t = t_f$$

Taking the partial derivative of the foregoing equation with respect to ϵ and then setting $\epsilon = 0$, we have

$$\hat{x}_1\eta_{x_1} + \hat{x}_2\eta_{x_2} = 0, \qquad t = t_f$$

We thus see that the specification of the terminal manifold

$$x_1^2(t_f) + x_2^2(t_f) = 1$$

leads to a linear relationship between η_{x_1} and η_{x_2} at the terminal time. If we combine this relation with the previously stated transversality condition, we obtain for one of the terminal boundary conditions

$$\frac{\partial\phi}{\partial\dot{x}_1}\frac{x_2}{x_1} - \frac{\partial\phi}{\partial\dot{x}_2} = 0, \qquad \text{at} \quad t = t_f$$

Therefore the two boundary conditions at $t = t_f$ are

$$x_1^2(t_f) + x_2^2(t_f) = 1$$

$$\frac{\partial\phi}{\partial\dot{x}_1(t_f)}\frac{x_2(t_f)}{x_1(t_f)} - \frac{\partial\phi}{\partial\dot{x}_2(t_f)} = 0$$

Thus for a given $\phi(\mathbf{x}, \dot{\mathbf{x}}, t)$, we can resolve this problem completely by solving for the optimal trajectory through the Euler-Lagrange equations and the appropriate boundary conditions which we have just obtained.

3.6 Variational notation

Much of the notation in the problems that follow can be considerably simplified if variational rather than differential notation is used. We wish to minimize (for t_0 and t_f fixed)

$$J = \int_{t_0}^{t_f} \phi(x, \dot{x}, t)\, dt \tag{3.6-1}$$

We assume, as in Section 3.1, that both $x(t)$ and $\dot{x}(t)$ are representable by a family of curves

$$x(t) = \hat{x}(t) + \epsilon\eta(t), \qquad \dot{x}(t) = \dot{\hat{x}}(t) + \epsilon\dot{\eta}(t) \tag{3.6-2}$$

where $x(t)$ is the optimal (extremal) curve and $\eta(t)$ is a variation in $x(t)$ depending upon t. We substitute Eq. (3.6-2) into Eq. (3.6-1 ıd expand $\phi(x, \dot{x}, t)$ in a Taylor series about the point $\epsilon = 0$.

$$\phi[\hat{x}(t) + \epsilon\eta(t), \dot{\hat{x}}(t) + \epsilon\dot{\eta}(t), t] = \phi(\hat{x}, \dot{\hat{x}}, t) + \frac{\partial\phi}{\partial\hat{x}}\epsilon\eta(t) + \frac{\partial\phi}{\partial\dot{\hat{x}}}\epsilon\dot{\eta}(t) + \text{H.O.T.} \tag{3.6-3}$$

where H.O.T. is used to indicate higher-order terms in $\eta(t)$ and $\dot{\eta}(t)$.

If we now let

$$\Delta J = J\langle\hat{x} + \epsilon\eta\rangle - J\langle\hat{x}\rangle$$

we can write

$$\begin{aligned} \Delta J &= \int_{t_0}^{t_f}\{\phi[\hat{x}(t) + \epsilon\eta(t), \dot{\hat{x}} + \epsilon\dot{\eta}, t] - \phi[\hat{x}(t), \dot{\hat{x}}, t]\}\, dt \\ &= \int_{t_0}^{t_f}\left\{\frac{\partial\phi}{\partial\hat{x}}\epsilon\eta(t) + \frac{\partial\phi}{\partial\dot{\hat{x}}}\epsilon\dot{\eta}(t) + \text{H.O.T.}\right\} dt \end{aligned} \tag{3.6-4}$$

Now we define the first variation of $x(t)$ and $\dot{x}(t)$ as

$$\epsilon\eta(t) = \delta x, \qquad \epsilon\dot{\eta}(t) = \delta\dot{x} \tag{3.6-5}$$

Thus

$$\Delta J = \int_{t_0}^{t_f}\left[\frac{\partial\phi}{\partial\hat{x}}\delta x + \frac{\partial\phi}{\partial\dot{\hat{x}}}\delta\dot{x} + \text{H.O.T.}\right] dt \tag{3.6-6}$$

Since the variation plays the same role in variational calculus as the differential in standard calculus, we use the property of linearity, which means that the first variation of J, δJ, the linear part of ΔJ, is

$$\delta J = \int_{t_0}^{t_f}\left[\frac{\partial\phi}{\partial\hat{x}}\delta x + \frac{\partial\phi}{\partial\dot{\hat{x}}}\delta\dot{x}\right] dt \tag{3.6-7}$$

A necessary condition for an extremum at $x(t) = \hat{x}(t)$, i.e., $\epsilon = 0$, is that the first variation of J, δJ, be zero. Applying this to Eq. (3.6-7), along with the minor simplification of integrating by parts and dropping the ^ notation, we obtain

$$\int_{t_0}^{t_f}\left[\frac{\partial\phi}{\partial x} - \frac{d}{dt}\frac{\partial\phi}{\partial\dot{x}}\right]\delta x\, dt + \frac{\partial\phi}{\partial\dot{x}}\delta x\bigg|_{t=t_0}^{t=t_f} = 0 \tag{3.6-8}$$

For Eq. (3.6-8) to equal zero independent of the variation δx, we must have

$$\frac{\partial\phi}{\partial x} - \frac{d}{dt}\frac{\partial\phi}{\partial\dot{x}} = 0 \tag{3.6-9}$$

$$\frac{\partial\phi}{\partial\dot{x}}\delta x = 0, \qquad \text{for} \quad t = t_0, t_f \tag{3.6-10}$$

We note that Eq. (3.6-9) is the Euler-Lagrange equation and Eq. (3.6-10) is its associated transversality condition.

In a similar n er, it is also easy to show that the second variation of Eq. (3.6-1), written $\delta^2 J$, is

$$\delta^2 J = \frac{1}{2}\int_{t_0}^{t_f}\left\{(\delta x)^2\left[\frac{\partial^2\phi}{\partial x^2} - \frac{d}{dt}\frac{\partial^2\phi}{\partial \dot{x}^2}\right] + (\delta\dot{x})^2\frac{\partial^2\phi}{\partial\dot{x}^2}\right\} dt \tag{3.6-11}$$

where the second variation is now defined as the quadratic part of Eq. (3.6.6) or twice Eq. (3.3-4). As previously stated, the interpretations of the second variation are that $\delta^2 J \geq 0$ implies a minimum of J and $\delta^2 J \leq 0$ implies a maximum of J. A quadratic form integral similar to Eq. (3.3-4) also follows directly.

3.7 Dynamic optimization with equality constraints—Lagrange multipliers

A constrained optimization problem may require extremizing a cost function of the form

$$J = \int_{t_0}^{t_f} \phi(\mathbf{x}, \dot{\mathbf{x}}, t)\, dt \tag{3.7-1}$$

subject to the equality constraint

$$\Lambda(\mathbf{x}, \dot{\mathbf{x}}, t) = 0 \tag{3.7-2}$$

where $\mathbf{x}^T = [x_1, x_2, \ldots, x_n]$ and $\Lambda^T = [\Lambda_1, \Lambda_2, \ldots, \Lambda_m]$ with $m \leq n$. It can be shown that the solution to this problem is the same as that obtained by extremizing

$$J' = \int_{t_0}^{t_f} [\phi(\mathbf{x}, \dot{\mathbf{x}}, t) + \boldsymbol{\lambda}^T(t)\Lambda(\mathbf{x}, \dot{\mathbf{x}}, t)]\, dt \tag{3.7-3}$$

where $\boldsymbol{\lambda}^T = [\lambda_1, \lambda_2, \ldots, \lambda_m]$ is the vector equivalent of the Lagrange multiplier discussed in Chapter 2 [4].

To illustrate the development of the Lagrange multiplier, let us consider a special case where $\mathbf{x}$ is a two vector. Suppose that we wish to minimize

$$J = \int_{t_0}^{t_f} \phi(x_1, x_2, \dot{x}_1, \dot{x}_2, t)\, dt \tag{3.7-4}$$

subject to the constraint (with fixed end points)

$$\Lambda(x_1, x_2, t) = 0 \tag{3.7-5}$$

We will use the variational notation just developed to establish a method for treating the given equality constraint. To establish a minimum, it is necessary that the first variation of Eq. (3.7-4) be zero, that is

$$\delta J = \int_{t_0}^{t_f}\left\{\delta x_1\left[\frac{\partial\phi}{\partial x_1} - \frac{d}{dt}\frac{\partial\phi}{\partial\dot{x}_1}\right] + \delta x_2\left[\frac{\partial\phi}{\partial x_2} - \frac{d}{dt}\frac{\partial\phi}{\partial\dot{x}_2}\right]\right\} dt = 0 \tag{3.7-6}$$

If δx_1 were independent of δx_2, we could simply set each term of Eq. (3.7-6) equal to 0. Since the constraint provides a dependence on x_1 and x_2, we must take the given constraint into consideration. Taking the var on of Eq. (3.7-5) we have

$$\delta\Lambda = \frac{\partial\Lambda}{\partial x_1}\delta x_1 + \frac{\partial\Lambda}{\partial x_2}\delta x_2 = 0 \tag{3.7-7}$$

It also follows that, for any $\lambda(t)$, we may multiply Eq. (3.7-7) by $\lambda(t)$ and integrate so that

$$\int_{t_0}^{t_f} \lambda(t)\left[\frac{\partial\Lambda}{\partial x_1}\delta x_1 + \frac{\partial\Lambda}{\partial x_2}\delta x_2\right] dt = 0 \tag{3.7-8}$$

If we add Eq. (3.7-6) to Eq. (3.7-8) we obtain

$$0 = \int_{t_0}^{t_f}\left\{\delta x_1\left[\frac{\partial\phi}{\partial x_1} - \frac{d}{dt}\frac{\partial\phi}{\partial\dot{x}_1} + \lambda\frac{\partial\Lambda}{\partial x_1}\right] + \delta x_2\left[\frac{\partial\phi}{\partial x_2} - \frac{d}{dt}\frac{\partial\phi}{\partial\dot{x}_2} + \lambda\frac{\partial\Lambda}{\partial x_2}\right]\right\} dt \tag{3.7-9}$$

We will now adjust λ so that the term within the first brackets under the integral is zero. It also must follow that, since δx_2 is arbitrary, the term in the second brackets under the integral is also equal to zero. It is apparent that we would have obtained the same results had we reformulated the given problem by adjoining to the cost function the constraint via a Lagrange multiplier as in Eq. (3.7-3) and used the Euler-Lagrange equations on this cost function. The resulting Euler-Lagrange equations would then be solved subject to the equality constraint of Eq. (3.7-2).

Example 3.7-1

We are given the differential system

$$\ddot{\theta} = u(t)$$

which may be interpreted as the moment of inertia of a rocket in free space, and we desire to minimize

$$J = \tfrac{1}{2}\int_0^2 (\ddot{\theta})^2\, dt$$

such that

$$\theta(t=0) = 1, \qquad \theta(t=2) = 0$$
$$\dot{\theta}(t=0) = 1, \qquad \dot{\theta}(t=2) = 0$$

To cast this problem in state space notation, we let

$$x_1(t) = \theta(t), \qquad \dot{x}_1 = x_2(t), \qquad \dot{x}_2 = u(t)$$

Now the differential system can be represented by

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)$$

where

$$\mathbf{x}^T = [x_1 \quad x_2], \qquad \mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \mathbf{b}^T = [0 \quad 1]$$

When we apply (3.7-3) ($u(t)$ is treated as another state variable, x_3), the problem becomes one of minimizing

$$J = \int_0^2 \{\tfrac{1}{2}u^2(t) + \boldsymbol{\lambda}^T(t)[\mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) - \dot{\mathbf{x}}]\}\, dt$$

$$= \int_0^2 \{\tfrac{1}{2}u^2(t) + \lambda_1(t)[x_2(t) - \dot{x}_1] + \lambda_2(t)[u(t) - \dot{x}_2]\}\, dt$$

The Euler-Lagrange equations yield

$$\dot{\lambda}_1 = 0, \qquad \dot{\lambda}_2 = -\lambda_1(t), \qquad u(t) = -\lambda_2(t)$$

The final solution is obtained by means of the given differential relationships and boundary conditions, and it is

$$x_1 = \tfrac{1}{2}t^3 - \tfrac{7}{4}t^2 + t + 1, \qquad x_2 = \tfrac{3}{2}t^2 - \tfrac{7}{2}t + 1, \qquad u = 3t - \tfrac{7}{2}$$

This system, along with a plot of the system trajectories, is shown in Fig. 3.7-1.

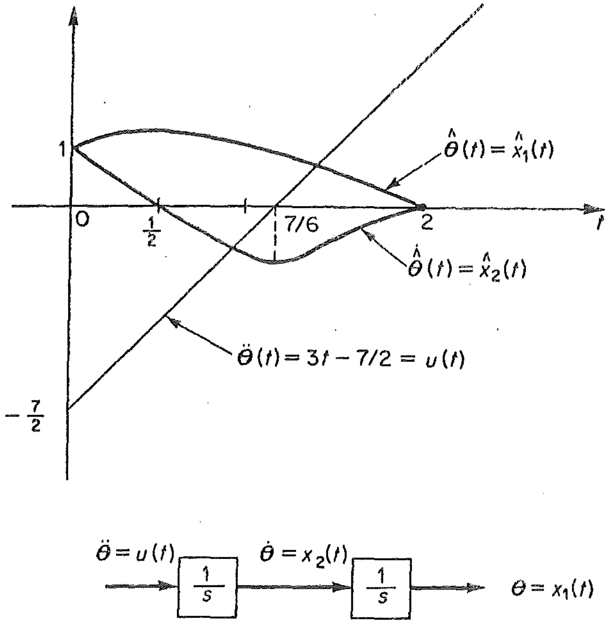


Fig. 3.7-1. Block diagram, optimal control and state variables for system of Example (3.7-1).

Example 3.7-2 Linear Servomechanism†

Suppose that we wish to minimize

$$J = \tfrac{1}{2}\int_{t_0}^{t_f} \|\mathbf{u}(t)\|^2_{\mathbf{R}(t)} + \|\mathbf{x}(t) - \mathbf{r}(t)\|^2_{\mathbf{Q}(t)}\, dt$$

for the general time-varying system specified by

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$

with $\mathbf{x}(t_0) = \mathbf{x}_0$ as the initial condition vector. $\mathbf{r}(t)$ is the desired value of the state vector $\mathbf{x}(t)$. As before, it is necessary to assume that all matrices and vectors are of compatible orders. We adjoin the differential system equality constraint to the cost function by the Lagrange multiplier to obtain

$$J' = \int_{t_0}^{t_f} \{\tfrac{1}{2}\|\mathbf{u}(t)\|^2_{\mathbf{R}(t)} + \tfrac{1}{2}\|\mathbf{x}(t) - \mathbf{r}(t)\|^2_{\mathbf{Q}(t)} + \boldsymbol{\lambda}^T(t)[\mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) - \dot{\mathbf{x}}]\}\, dt$$

The exact nature of the cost function used depends upon the particular problem being solved. Therefore $\mathbf{R}(t)$ and $\mathbf{Q}(t)$, both penalty-weighting matrices, are generally chosen with regard to the physical conditions present. We also assume that both $\mathbf{R}(t)$ and $\mathbf{Q}(t)$ are symmetric, since there is no loss in generality by doing so. The control vector, $\mathbf{u}(t)$ is treated just as if it were a state vector. Then we apply the Euler-Lagrange equations, which in this case are

$$\frac{\partial\Phi}{\partial\mathbf{x}} - \frac{d}{dt}\frac{\partial\Phi}{\partial\dot{\mathbf{x}}} = \mathbf{0}, \qquad \frac{\partial\Phi}{\partial\mathbf{u}} - \frac{d}{dt}\frac{\partial\Phi}{\partial\dot{\mathbf{u}}} = \mathbf{0}$$

where

$$\Phi = \tfrac{1}{2}\|\mathbf{u}(t)\|^2_{\mathbf{R}(t)} + \tfrac{1}{2}\|\mathbf{x}(t) - \mathbf{r}(t)\|^2_{\mathbf{Q}(t)} + \boldsymbol{\lambda}^T(t)[\mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) - \dot{\mathbf{x}}]$$

Thus

$$\frac{\partial\Phi}{\partial\mathbf{x}} = \mathbf{Q}(t)[\mathbf{x}(t) - \mathbf{r}(t)] + \mathbf{A}^T(t)\boldsymbol{\lambda}(t), \qquad \frac{\partial\Phi}{\partial\dot{\mathbf{x}}} = -\boldsymbol{\lambda}(t)$$

$$\frac{\partial\Phi}{\partial\mathbf{u}} = \mathbf{R}(t)\mathbf{u}(t) + \mathbf{B}^T(t)\boldsymbol{\lambda}(t), \qquad \frac{\partial\Phi}{\partial\dot{\mathbf{u}}} = \mathbf{0}$$

The Euler-Lagrange equations for this problem become

$$\dot{\boldsymbol{\lambda}} = -\mathbf{A}^T(t)\boldsymbol{\lambda}(t) - \mathbf{Q}(t)[\mathbf{x}(t) - \mathbf{r}(t)], \qquad \mathbf{u}(t) = -\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\boldsymbol{\lambda}(t)$$

Since $\mathbf{x}(t_f)$ is unspecified, the transversality condition at the terminal time yields $\boldsymbol{\lambda}(t_f) = \mathbf{0}$. This solution can be block-diagrammed as in Fig. 3.7-2. We note that the solution for the optimal control requires that $\mathbf{R}(t)$ have an inverse. Also, certain other requirements must be met to insure a minimum of the cost function; specifically, $\mathbf{R}(t)$ and $\mathbf{Q}(t)$ must be nonnegative definite to insure a nonnegative second variation. Thus we see that $\mathbf{R}(t)$ must be positive definite.

Although it appears that we have solved the originally stated problem, there are still some further refinements which are highly desired. Since the state of the system is specified at t_0, we are given $\mathbf{x}(t_0)$, while the adjoint operator $\boldsymbol{\lambda}(t)$ is specified at the terminal time, $\boldsymbol{\lambda}(t_f) = \mathbf{0}$. What we, in fact, have to do is solve a two-point boundary value problem (TPBVP), something which, in general, cannot always be done without recourse to electronic computers. In this particular case, since the differential equations are all linear, superposition can be invoked and a closed-form analytical solution obtained with great difficulty.

If we let $\mathbf{r}(t)$ be either a constant vector or the null vector, the foregoing problem reduces to a regulator problem. The treatment of the servomechanism problem can be made more general if we assume that indirect state observation is made available to us, that is, for the system

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$

---

†A considerably more detailed treatment of this problem will be given in Chapter 5.

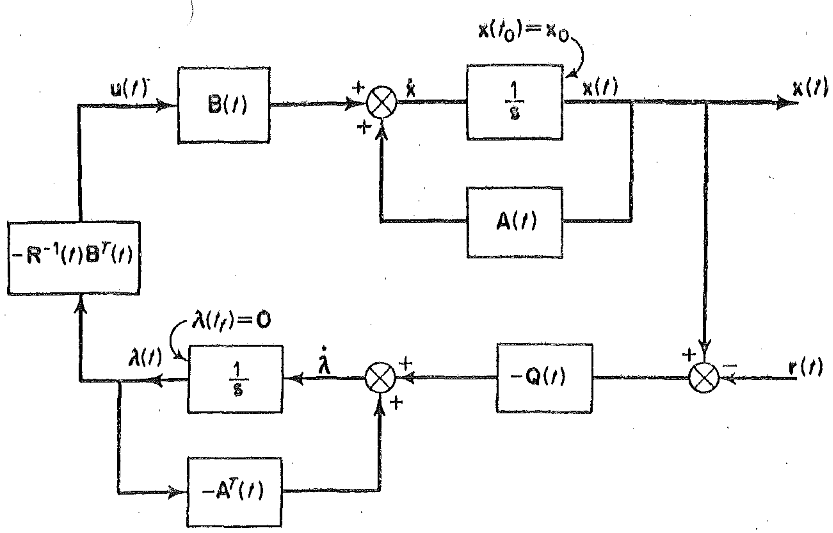


Fig. 3.7-2. Block diagram of a possible solution to the servomechanism problem.

we can obtain directly only

$$\mathbf{z}(t) = \mathbf{C}(t)\mathbf{x}(t) + \mathbf{D}(t)\mathbf{u}(t)$$

The procedure and results are quite similar to the ones obtained in this example except that requirements on observability and controllability, to be discussed in Chapter 11, are present.

To solve this two-point boundary value problem, we must require a knowledge of $\mathbf{r}(t)$ for all time in the closed interval t_o to t_f or, in shorthand notation, $\forall t \in [t_o, t_f]$. Since a two-point boundary value problem must be solved before we can determine the optimum control for this problem, it is clear that a closed-loop control has not been found. After we have formulated the Hamilton-Jacobi equations and the Pontryagin maximum principle, we will have a great deal more to say about this important problem.

3.8 Dynamic optimization with inequality constraints

In many physical problems of interest to the control engineer, there are various inequality constraints on the control vector. For example, the maximum thrust from a reaction jet is physically limited as is the maximum input reactivity in a nuclear reactor. When inequality constraints are present, it is necessary that we consider them in determining optimum system design.

Thus we are faced with minimizing a cost function of the f

$$J = \int_{t_o}^{t_f} \phi(\mathbf{x}, \dot{\mathbf{x}}, t)\, dt \tag{3.8-1}$$

with equality constraints of the form

$$\Lambda(\mathbf{x}, \dot{\mathbf{x}}, t) = 0 \tag{3.8-2}$$

and inequality constraints of the form

$$\Gamma_{\min} \leq \Gamma(\mathbf{x}, \dot{\mathbf{x}}, t) \leq \Gamma_{\max} \tag{3.8-3}$$

When the inequality constraint involves the control vector, the control vector which satisfies the constraint conditions is called an admissible control vector. One technique which is generally satisfactory for resolving the control inequality constraint problem consists of converting the inequality constraint to an equality constraint. It can be easily demonstrated that the, equations

but this is impro

$$(\Gamma_{\max\, i} - \Gamma_i)(\Gamma_i - \Gamma_{\min\, i}) = \gamma_i^2, \qquad i = 1, 2, \ldots \tag{3.8-4}$$

are equivalent to the constraints of Eq. (3.8-3), since each term on the left side of Eq. (3.8-4) must be positive, or each negative and thus have a positive product. Thus the inequality constraints have been converted to equality constraints and may be treated as such. Lagrange multipliers are then used to adjoin the equality and inequality constraints to the cost function, Eq. (3.8-1), and the Euler-Lagrange equations applied.† The technique can best be illustrated by an example.

Example 3.8-1

Let us consider the same plant dynamics as in the previous example 3.7-1.

$$\dot{x}_1 = x_2(t), \qquad \dot{x}_2 = u(t)$$

with the initial conditions $x_1(t_o) = x_o$ and $x_2(t_o) = v_o$. The problem is to find the control which maximizes $x_1(t_f)$, for fixed t_f, subject to the boundary condition equality constraint that $x_2(t_f) = v_f$ and the inequality constraint on the scalar control $u_{\min} \leq u \leq u_{\max}$. We convert the inequality constraint to an equality constraint by introducing a new variable $\alpha(t)$ and replacing the inequality constraint by

it is equivalent of minimizing $-x_1(t_f)$

$$(u - u_{\min})(u_{\max} - u) - \alpha^2 = 0$$

Thus the problem may be recast as one of minimizing $J = -x_1(t_f)$ subject to the equality constraints

$$\begin{aligned} \dot{x}_1 &= x_2(t), & x_1(t_o) &= x_o, & x_1(t_f) &= \text{open} \\ \dot{x}_2 &= u(t), & x_2(t_o) &= v_o, & x_2(t_f) &= v_f \\ (u - u_{\min})&(u_{\max} - u) - \alpha^2 = 0 \end{aligned}$$

---

†Chapters 4, 13, and 14 will consider more varied aspects, theoretical and computational, of the inequality constraint problem.

The cost function with the adjoined Lagrange multiplier becomes

$$J' = \boxed{-x_1(t_o) + \int_{t_o}^{t_f} -\dot{x}_1} + \lambda_1[x_2 - \dot{x}_1] + \lambda_2[u - \dot{x}_2] + \lambda_3[(u - u_{\min})(u_{\max} - u) - \alpha^2]\,dt$$

$-x_1(t_f) =$

The Euler-Lagrange equation

$$\frac{d}{dt}\frac{\partial\Phi}{\partial\dot{\mathbf{x}}} - \frac{\partial\Phi}{\partial\mathbf{x}} = 0, \qquad \boxed{\mathbf{x}^T = [x_1, x_2, u]} \;?\;!$$

$\underline{x}' = [x_1, x_2, u, \alpha]$

with

$$\Phi = \lambda_1[x_2 - \dot{x}_1] + \lambda_2[u - \dot{x}_2] + \lambda_3[(u - u_{\min})(u_{\max} - u) - \alpha^2] - \dot{x}_1$$

yields

$$\dot{\lambda}_1 = 0, \qquad \dot{\lambda}_2 = -\lambda_1$$
$$0 = -\lambda_2 + \lambda_3[2u - u_{\max} - u_{\min}], \qquad 0 = \alpha\lambda_3$$

Application of the natural boundary condition equation (transversality condition) to determine the single missing terminal condition on $x_1(t_f)$ yields

$$\left.\frac{\partial\Phi}{\partial\dot{x}_1}\right|_{t=t_f} = 0 = -1 - \lambda_1(t_f)$$

Thus we have arrived at the two-point boundary value problem whose solution determines the optimal state and control variables. This TPBVP is

$$\dot{x}_1 = x_2(t), \qquad x_1(t_o) = x_o$$
$$\dot{x}_2 = u(t), \qquad x_2(t_o) = v_o$$
$$\dot{\lambda}_1 = 0, \qquad \lambda_1(t_f) = -1$$
$$\dot{\lambda}_2 = -\lambda_1(t), \qquad x_2(t_f) = v_f$$
$$\alpha(t)\lambda_3(t) = 0$$
$$\lambda_2(t) = \lambda_3(t)[2u(t) - u_{\max} - u_{\min}]$$
$$\alpha^2(t) = [u(t) - u_{\min}][u_{\max} - u(t)]$$

This TPBVP is nonlinear because of the last three coupling equations above and is quite difficult to solve without recourse to a computer. In a usual version of this problem, $u_{\min} = -1$ and $u_{\max} = +1$. In that case, it is possible to show that $\alpha(t) = 0$ and

$$u(t) = -\text{sign}\,\lambda_2(t)$$

where

$$\text{sign}\,\lambda_2 = 1 \qquad \text{if} \quad \lambda_2 > 0$$
$$\text{sign}\,\lambda_2 = -1 \qquad \text{if} \quad \lambda_2 < 0$$

This does not, however, change the nonlinear nature of the two-point boundary problem. In a later chapter we will devote considerable time to various gradient methods, Newton-Raphson techniques, and other computational techniques for solving nonlinear two-point (and multipoint) boundary value problems.

REFERENCES

1. Berkovitz, L.D., "Variational Methods in Problems of Control and Programming," *Journal Math. Anal. Appl.*, Vol. 3, pp. 145–169, 1961.
2. Bliss, G.A., *Lectures on the Calculus of Variations.* University of Chicago Press, Chicago, 1963.
3. Elsgoc, L.E., *Calculus of Variations.* Pergamon Press, Ltd., New York, 1961.
4. Gelfand, I.M. and Fomin, S.V., *Calculus of Variations.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1963.
5. Kalman, R.E., "The Theory of Optimal Control and the Calculus of Variations," *Mathematical Optimization Techniques.* University of California Press, Los Angeles, California, 1963.
6. Leitman, G., *Optimal Control.* McGraw-Hill Book Company, New York, 1966.
7. Leitman, G. ed., *Optimization Techniques.* Academic Press, New York, 1962.
8. Tou, J., *Modern Control Theory.* McGraw-Hill Book Company, New York, 1964.

PROBLEMS

✓

1. A linear differential system is described by

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

where

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \mathbf{x}^T = [x_1, x_2], \qquad \mathbf{u}^T = [u_1, u_2]$$

Find $\mathbf{u}(t)$ such that

$$J = \tfrac{1}{2}\int_0^2 ||\mathbf{u}||^2\,dt$$

is minimum, given $\mathbf{x}^T(0) = [1, 1]$ and $x_1(2) = 0$.

2. Find the conditions necessary for minimizing

$$J = \theta[x(t_f)] + \int_{t_o}^{t_f} \phi(x, \dot{x}, t)\,dt$$

✓ given $x(t_o) = x_0$ and $g(x, \dot{x}, t) = 0$.

3. Use the results of Problem 2 to find the control $u(t)$, which minimizes

$$J = \frac{S}{2}x^2(2) + \tfrac{1}{2}\int_0^2 u^2\,dt$$

such that $\dot{x} = u(t)$, $x(0) = 1$.

4. A linear system is described by

$$\dot{x} = -x + u, \qquad x(0) = 1$$

It is desired to minimize

$$J = \tfrac{1}{2}\int_0^2 (x^2 + u^2)\,dt$$

✓ A feedback law is obtained if we let $u(t) = \alpha x(t)$ where $d\alpha/dt = 0$ such that α is a constant. Find the equations defining the optimum value of α.

5. Find the differential equations and associated boundary conditions whose solutions minimize

$$J = \tfrac{1}{2}\int_0^{t_f} u^2\,dt$$

for the differential system described by

$$\dot{x}_1 = -x_1 + x_2$$
$$\dot{x}_2 = u$$

with end points given by

$$x_1(0) = x_2(0) = 0$$
$$x_1^2(t_f) + x_2^2(t_f) = t_f^2 + 1$$

6. Find the value of u which minimizes (for t_f unspecified)

$$J = \int_0^{t_f} [\alpha + u^2(t) + x^2(t)]\, dt$$

for the differential system

$$\dot{x} = -x(t) + u(t), \quad x(0) = 1, \quad x(t_f) = 0$$

7. A linear second-order differential equation is described by

$$\dot{x}_1 = x_2(t), \qquad x_1(0) = 1$$
$$\dot{x}_2 = u \qquad\qquad x_2(0) = 1$$

Find, by use of the Euler-Lagrange equations and transversality conditions, the optimal control $u(t)$ which minimizes:

(a) $J = \int_0^1 u^2\, dt, \quad x_1(1) = x_2(1) = 0$ $\quad u = 15t - 10$

(b) $J = \int_0^1 u^2\, dt, \quad x_1(1) = 0$ $\quad u = 6(t-1)$

(c) $J = \int_0^{t_f} u^2\, dt, \quad x_1(t_f) = c(t_f) = -t_f^2$ $\quad u = \frac{1}{9}(1+5\sqrt{15})t - \frac{1}{3}(11+\sqrt{13})$

(Also determine t_f and $x_1(t_f)$.)

(d) $J = \int_0^{t_f} u^2\, dt, \quad x_1(t_f) = c(t_f) = -t_f^2, \quad x_2(t_f) = 0$

(e) $J = \int_0^1 \{\|\mathbf{x}\|^2 + \|\mathbf{u}\|^2\}\, dt$ $\quad u = -t^2 + 4t - 3.$

For all cases, sketch both the optimal system trajectory $\mathbf{x}(t)$ and the optimal system control $u(t)$.

8. For the fixed plant dynamics given by

$$\dot{x} = u$$

determine the optimal closed-loop system which minimizes

$$J = \tfrac{1}{2} \int_0^2 \{u^2 + (x - i)^2\}\, dt$$

where $i(t) = 1 - e^{-t}$.

9. For the fixed plant dynamics given by $\dot{x} = u(t)$, $x(0) = x_0$, determine the optimal closed-loop control which minimizes for fixed t_f

$$J = \tfrac{1}{2} s x^2(t_f) + \tfrac{1}{2} \int_0^{t_f} u^2\, dt$$

where s is an arbitrary constant. Do this by first determining the optimum open-loop control and trajectory and then let $u(t) = k(t)x(t)$.

4

THE MAXIMUM PRINCIPLE AND HAMILTON-JACOBI THEORY

In the previous chapter, we formulated many problems in the classical calculus of variations. A derivation of the Euler-Lagrange equations for both the scalar and vector cases was presented. We discussed the associated transversality conditions and some of the difficulties which we may encounter if inequality constraints are present. Several simple optimal control problems were stated and solved. In this chapter we wish to reexamine many of the problems presented in the previous chapter and obtain more general solutions for some of them. In addition, we will develop methods for handling some problems which could not be conveniently formulated by the methods in the previous chapter.

To these ends, we will present the Bolza formulation of the variational calculus using Hamiltonian methods. This will lead us into a proof of the Pontryagin maximum principle and the associated transversality conditions. We will proceed then to a development of the Hamilton-Jacobi equations, which are equivalent to Bellman's equations of continuous dynamic programming. Finally, we will give brief mention to some limitations of dynamic programming. Examples to illustrate the methods will be presented. We will reserve the next chapter for a discussion of some of the many problems which we can formulate and solve using the maximum principle.

In order to fully develop our approach to optimization theory where the termin[illegible]me is not fixed and where the control and state vectors are not necessarily smooth functions, we must consider in more detail the first variation for such problems.

4.1 Variation of functions with terminal times not fixed—the Weierstrass-Erdmann conditions

In this section, we will consider problems which arise when the terminal (or initial) time is not fixed (unspecified in the problem statement). We must reexamine our concept of a variation in order to accurately treat problems wherein the terminal (or initial) time is not fixed if we are to use the powerful concept of the first variation. We thus wish to consider the extremization of

$$J = \int_{t_0}^{t_f} \Phi[\mathbf{x}(t), \dot{\mathbf{x}}(t), t]\, dt \tag{4.1-1}$$

where all admissible trajectories are smooth and where the terminal time is not fixed. We define a variation δJ as the part of

$$\Delta J = J[\mathbf{x} + \mathbf{h}, t_f + \delta t_f] - J[\mathbf{x}, t_f] \tag{4.1-2}$$

which is linear in $\mathbf{h}$, $\dot{\mathbf{h}}$, $\delta\mathbf{x}$, $\delta\dot{\mathbf{x}}$, and δt_f. Since both $\mathbf{x}$ and t_f vary, it is appropriate to consider the variation $\delta\mathbf{x}$ as

$$\delta\mathbf{x}\,(t_f) = \mathbf{h}(t_f) + \dot{\mathbf{x}}(t_f)\,\delta t_f \tag{4.1-3}$$

For the cost function, Eq. (4.1-1), we find that

$$\Delta J = \int_{t_0}^{t_f+\delta t_f} \Phi[\mathbf{x}(t) + \mathbf{h}(t), \dot{\mathbf{x}}(t) + \dot{\mathbf{h}}(t), t]\, dt - \int_{t_0}^{t_f} \Phi[\mathbf{x}(t), \dot{\mathbf{x}}(t), t]\, dt \tag{4.1-4}$$

By taking the linear terms in this equation and performing an integration by parts, we obtain the first variation as†

$$\begin{aligned} \delta J = {} & \Phi[\mathbf{x}(t_f), \dot{\mathbf{x}}(t_f), t_f]\,\delta t_f + \mathbf{h}^T(t_f)\frac{\partial\Phi[\mathbf{x}(t_f), \dot{\mathbf{x}}(t_f), t_f]}{\partial\dot{\mathbf{x}}(t_f)} \\ & + \int_{t_0}^{t_f} \mathbf{h}^T(t)\left\{\frac{\partial\Phi}{\partial\mathbf{x}} - \frac{d}{dt}\frac{\partial\Phi}{\partial\dot{\mathbf{x}}}\right\} dt \end{aligned} \tag{4.1-5}$$

where, for convenience, we assume that $\mathbf{h}(t_o) = 0$. Using Eq. (4.1-3), the first variation becomes

$$\begin{aligned} \delta J = {} & \left\{\Phi[\mathbf{x}(t_f), \dot{\mathbf{x}}(t_f), t_f] - \dot{\mathbf{x}}^T(t_f)\frac{\partial\Phi[\mathbf{x}(t_f), \dot{\mathbf{x}}(t_f), t_f]}{\partial\dot{\mathbf{x}}(t_f)}\right\}\delta t_f \\ & + \delta\mathbf{x}^T(t_f)\frac{\partial\Phi[\mathbf{x}(t_f), \dot{\mathbf{x}}(t_f), t_f]}{\partial\dot{\mathbf{x}}(t_f)} \\ & + \int_{t_0}^{t_f} \mathbf{h}^T(t)\left\{\frac{\partial\Phi}{\partial\mathbf{x}} - \frac{d}{dt}\frac{\partial\Phi}{\partial\dot{\mathbf{x}}}\right\} dt \end{aligned} \tag{4.1-6}$$

In much of our work, it will be convenient to define a qu[illegible]ty, called the Hamiltonian, by

$$H[\mathbf{x}(t), \boldsymbol{\lambda}(t), t] = \Phi - \dot{\mathbf{x}}^T\frac{\partial\Phi}{\partial\dot{\mathbf{x}}} = \Phi + \dot{\mathbf{x}}^T\boldsymbol{\lambda} \tag{4.1-7}$$

? where the Hamiltonian is not a function of $\dot{\mathbf{x}}$. $\mathbf{x}(t)$ and $\boldsymbol{\lambda}(t)$ are called the canonical variables. In terms of the Hamiltonian, the first variation of Eq. (4.1-1), which is Eq. (4.1-6), becomes

$$\begin{aligned} \delta J = {} & -\delta\mathbf{x}^T(t_f)\boldsymbol{\lambda}(t_f) + H[\mathbf{x}(t_f), \boldsymbol{\lambda}(t_f), t_f]\,\delta t_f \\ & + \int_{t_0}^{t_f} \mathbf{h}^T(t)\left\{\frac{\partial H}{\partial\mathbf{x}} + \frac{d\boldsymbol{\lambda}}{dt}\right\} dt \end{aligned} \tag{4.1-8}$$

To establish a necessary condition for a minimum, it is necessary that the integrand in Eqs. (4.1-6) and (4.1-8) vanish and also that the transversality condition, as obtained from Eq. (4.1-8)

$$-\delta\mathbf{x}^T(t_f)\boldsymbol{\lambda}(t_f) + H[\mathbf{x}(t_f), \boldsymbol{\lambda}(t_f), t_f]\,\delta t_f = 0 \tag{4.1-9}$$

be satisfied.

Thus far in our development we have considered functions with "smooth" arcs. Let us now consider the problem of minimizing the cost function

$$J = \int_0^1 x^2(2 - \dot{x})^2\, dt$$

subject to

$$x(0) = 0, \qquad x(1) = 1$$

Physically, it is clear that the absolute minimum for J is 0 and that this is obtained for

$$\begin{aligned} x(t) &= 0 & t &\in [0, \tfrac{1}{2}] \\ x(t) &= 2t - 1 & t &\in [\tfrac{1}{2}, 1] \end{aligned}$$

which is certainly a solution to the Euler-Lagrange equation for this problem

$$x^2\ddot{x} + x\dot{x}^2 - 4x = 0$$

There is one disturbing feature about this solution, however, in that the optimum $x(t)$ has a "corner" or discontinuous first derivative which gives rise to formal difficulty since $\dot{x}$ is contained in the Euler-Lagrange equations. Certainly, though, this particular function $x(t)$ is smooth in a piecewise sense, or piecewise smooth. We will define a function as being smooth in an interval of time if it is continuous and has a continuous time derivative in the interval. A function is piecewise smooth if it is smooth except for, at most, a finite number of points. We may examine further the special requirements imposed by this "corner" by considering the Weierstrass-Erdmann conditions [1].

The Weierstrass-Erdmann corner conditions furnish us with the requirements for a solution at corners or jumps in the extremal curve. In all of our work thus far (except Section 3.8), we have considered functions defined for

---

†It is not correct to call $\mathbf{h}(t)$ the first variation if the terminal time is not fixed. This does not alter any results if differential notation, $\mathbf{x}(t) = \hat{\mathbf{x}}(t) + \epsilon\eta_x(t)$, and $t_f = \hat{t}_f + \epsilon\eta_f$ are used. It would, of course, be correct to use the symbol $\overline{\delta\mathbf{x}(t)} = \mathbf{h}(t)$, where $\overline{\delta\mathbf{x}(t)}$ is the variation in $\mathbf{x}$ only and does not include a variation in terminal time.

smooth ar nd thus have allowed only smooth solutions of the associated variational problems. The Weierstrass-Erdmann conditions extend the class of admissible arcs to include those which are only piecewise smooth. Specifically, we wish to find the function $\hat{\mathbf{x}}(t)$ among all functions $\mathbf{x}(t)$ which are continuously differentiable for $t \in [a, b]$, except at some point $c \in (a, b)$, and which satisfies prescribed boundary conditions such that the functional

$$J(\mathbf{x}) = \int_a^b \Phi[\mathbf{x}(t), \dot{\mathbf{x}}(t), t]\, dt \tag{4.1-10}$$

has an extremum. It is of course clear that, for $t \in [a, c]$ and $t \in [c, b]$, the function $\mathbf{x}(t)$ must satisfy the Euler-Lagrange equations for a minimum

$$\frac{d}{dt}\frac{\partial \Phi}{\partial \dot{\mathbf{x}}} - \frac{\partial \Phi}{\partial \mathbf{x}} = 0 \tag{4.1-11}$$

We may rewrite the cost function as a sum of two cost functions:

$$\begin{aligned} J(\mathbf{x}) &= \int_a^c \Phi[\mathbf{x}(t), \dot{\mathbf{x}}(t), t]\, dt + \int_c^b \Phi[\mathbf{x}(t), \dot{\mathbf{x}}(t), t]\, dt \\ &= J_1(\mathbf{x}) + J_2(\mathbf{x}) \end{aligned} \tag{4.1-12}$$

We may now take the first variation $\delta J_1(\mathbf{x})$ and $\delta J_2(\mathbf{x})$ separately. We assume, for the moment only, that a and b are fixed, and we require that the $\hat{\mathbf{x}}(t)$ calculated from $J_1(\mathbf{x})$ and $J_2(\mathbf{x})$ is the same at $t = c$ which is unknown. Since c is arbitrary, the first variation of $J_1(\mathbf{x})$ is

$$\begin{aligned} \delta J_1(\mathbf{x}) = {} & \boldsymbol{\delta}\mathbf{x}^T(a)\frac{\partial \Phi[\mathbf{x}(a), \dot{\mathbf{x}}(a), a]}{\partial \dot{\mathbf{x}}(a)} \\ & + \left\{\Phi[\mathbf{x}(c), \dot{\mathbf{x}}(c), c] - \dot{\mathbf{x}}^T(c)\frac{\partial \Phi[\mathbf{x}(c), \dot{\mathbf{x}}(c), c]}{\partial \dot{\mathbf{x}}(c)}\right\} \delta c \\ & + \boldsymbol{\delta}\mathbf{x}^T(c)\frac{\partial \Phi[\mathbf{x}(c), \dot{\mathbf{x}}(c), c]}{\partial \dot{\mathbf{x}}(c)} + \int_a^c \mathbf{h}^T(t)\left\{\frac{\partial \Phi}{\partial \mathbf{x}} - \frac{d}{dt}\frac{\partial \Phi}{\partial \dot{\mathbf{x}}}\right\} dt \end{aligned} \tag{4.1-13}$$

Since $\mathbf{x}(t)$ satisfies the Euler-Lagrange equations for an extremal and since $\boldsymbol{\delta}\mathbf{x}(a) = \mathbf{0}$, we have

$$\begin{aligned} \delta J_1(\mathbf{x}) = {} & \boldsymbol{\delta}\mathbf{x}^T(\tau)\frac{\partial \Phi[\mathbf{x}(\tau), \dot{\mathbf{x}}(\tau), \tau]}{\partial \dot{\mathbf{x}}(\tau)} \\ & + \left\{\Phi[\mathbf{x}(\tau), \dot{\mathbf{x}}(\tau), \tau] - \dot{\mathbf{x}}^T(\tau)\frac{\partial \Phi[\mathbf{x}(\tau), \dot{\mathbf{x}}(\tau), \tau]}{\partial \dot{\mathbf{x}}(\tau)}\right\} \delta\tau \\ & (\text{for } \tau = c - 0) \end{aligned} \tag{4.1-14}$$

In a similar fashion, we can show that the first variation for the extremal solution of $J_2(\mathbf{x})$ is

$$\begin{aligned} \delta J_2(\mathbf{x}) = {} & -\boldsymbol{\delta}\mathbf{x}^T(\tau)\frac{\partial \Phi[\mathbf{x}(\tau), \dot{\mathbf{x}}(\tau), \tau]}{\partial \dot{\mathbf{x}}(\tau)} \\ & - \left\{\Phi[\mathbf{x}(\tau), \dot{\mathbf{x}}(\tau), \tau] - \dot{\mathbf{x}}^T(\tau)\frac{\partial \Phi[\mathbf{x}(\tau), \dot{\mathbf{x}}(\tau), \tau]}{\partial \dot{\mathbf{x}}(\tau)}\right\} \delta\tau \\ & (\text{for } \tau = c + 0) \end{aligned} \tag{4.1-15}$$

In order to obtain the extremum, the extremal solution mu tisfy

$$\delta J(\mathbf{x}) = \delta J_1(\mathbf{x}) + \delta J_2(\mathbf{x}) = 0 \tag{4.1-16}$$

Thus

$$\left.\frac{\partial \Phi}{\partial \dot{\mathbf{x}}}\right|_{t=c-0} = \left.\frac{\partial \Phi}{\partial \dot{\mathbf{x}}}\right|_{t=c+0} \tag{4.1-17}$$

$$\left.\Phi - \dot{\mathbf{x}}^T\frac{\partial \Phi}{\partial \dot{\mathbf{x}}}\right|_{t=c-0} = \left.\Phi - \dot{\mathbf{x}}^T\frac{\partial \Phi}{\partial \dot{\mathbf{x}}}\right|_{t=c+0} \tag{4.1-18}$$

since $\boldsymbol{\delta}\mathbf{x}$ and δt_f are arbitrary. These requirements, Eqs. (4.1-17) and (4.1-18), are called the Weierstrass-Erdmann corner conditions and must hold at any point c where the extremal has a corner. If we use the Hamiltonian canonical variables

$$H = \Phi - \dot{\mathbf{x}}^T\frac{\partial \Phi}{\partial \dot{\mathbf{x}}} = \Phi + \boldsymbol{\lambda}^T\dot{\mathbf{x}} \tag{4.1-19}$$

$$\boldsymbol{\lambda} = -\frac{\partial \Phi}{\partial \dot{\mathbf{x}}} \tag{4.1-20}$$

we immediately see that the Weierstrass-Erdmann conditions simply require H and $\boldsymbol{\lambda}$ to be continuous on the optimum trajectory at all points where there are corners.

It is possible to generalize the Weierstrass-Erdmann corner condition in terms of the Weierstrass E function, defined as

$$E = \left\{\Phi(\mathbf{x}, \dot{\mathbf{X}}, t) - \Phi(\mathbf{x}, \dot{\mathbf{x}}, t) - (\dot{\mathbf{X}} - \dot{\mathbf{x}})^T\frac{\partial \Phi}{\partial \dot{\mathbf{x}}}\right\} \geq 0 \tag{4.1-21}$$

where $\partial\Phi/\partial\dot{\mathbf{x}}$ is evaluated at the optimum solution vector $\mathbf{x}(t)$ and $\dot{\mathbf{X}}$ is an admissible vector, one which satisfies all constraints. This provides us with necessary conditions for an extremum under constrained conditions [1, 6].†

In the next section, we will examine, among other things, minimum time problems for problems where the extremal arcs or trajectories are smooth but where the terminal time is not fixed. Thus we will need to use the expanded variational notation presented in the first part of this section. Then we will consider the important case in optimal control where the admissible control and state variables are restricted. We will then use the Weierstrass E function to develop a maximum principle. In this work we will find it necessary to interpret the vector $\mathbf{x}$ in this section as the generalized state vector, which includes the control vector.

4.2 The Bolza problem and its solution

We will introduce the Hamiltonian approach to the solution of variational problems by considering the Bolza problem of the variational calculus and

---

†Certain other conditions are also required, such as absence of conjugate points. References [1], [6], and [11] provide much elaboration on this point.

several exter[illegible]s. We shall see that the results obtained are similar in many ways to the results of the Pontryagin maximum principle which we will present in the next section. Our approach to this section will be, as before, to employ classical variational techniques.

4.2-1. Continuous optimal control problems—fixed beginning and terminal times—no inequality constraints

We are given a nonlinear differential system operating over the fixed interval $t \in [t_o, t_f]$ of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \tag{4.2-1}$$

where $\mathbf{x}(t)$, the n vector state variable, is determined by $\mathbf{u}(t)$, the m vector control variable, and the initial condition vector

$$\mathbf{x}(t_o) = \mathbf{x}_o \tag{4.2-2}$$

Actually, the statement that all components of the n-dimensional state vector are fixed at the initial time, t_o, is a bit restrictive, although it is generally true for optimal control problems. However, in the state and parameter estimation problem, not all of the components of the state vector are specified initially. Thus a more general[*] statement of the specified initial conditions is

$$\mathbf{M}(t_o)\mathbf{x}(t_o) = \mathbf{m}_o \tag{4.2-3}$$

where $\mathbf{m}_o$ is an r vector. In a similar fashion, some of the terminal states may be specified. In this case, we may find[†]

$$\mathbf{N}(t_f)\mathbf{x}(t_f) = \mathbf{n}_f \tag{4.2-4}$$

where $\mathbf{n}_f$ is a q vector, $q \leq n$.

We will return to a discussion of this point momentarily. But now we desire to determine the control $\mathbf{u}(t)$ such as to minimize

$\frac{d\theta}{dt}\,dt + \int_{t_0}^{t_f} \phi\, dt$

$$J = \theta[\mathbf{x}(t), t]\Big|_{t=t_o}^{t=t_f} + \int_{t_o}^{t_f} \phi[\mathbf{x}(t), \mathbf{u}(t), t]\, dt \tag{4.2-5}$$

We use the method of Lagrange multipliers discussed in the last chapter to adjoin the system differential equality constraint to the cost function, which gives us

$+\theta \int dt$

$$J = \theta[\mathbf{x}(t), t]\Big|_{t=t_o}^{t=t_f} + \int_{t_o}^{t_f} \{\phi[\mathbf{x}(t), \mathbf{u}(t), t] + \boldsymbol{\lambda}^T(t)[\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] - \dot{\mathbf{x}}]\}dt \tag{4.2-6}$$

ducas functional

We define a scalar function, the Hamiltonian, as

$$H[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t] = \phi[\mathbf{x}(t), \mathbf{u}(t), t] + \boldsymbol{\lambda}^T(t)\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] \tag{4.2-7}$$

---

[*] These are, of course, still not the most general statements for the initial and terminal manifold.

[†] These are, of course, not the most general statements for the terminal manifold.

Thus the cost function becomes

$$J = \theta[\mathbf{x}(t), t]\Big|_{t=t_o}^{t=t_f} + \int_{t_o}^{t_f} \{H[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t] - \boldsymbol{\lambda}^T(t)\dot{\mathbf{x}}\}\, dt \tag{4.2-8}$$

If we integrate the last term in the integrand of Eq. (4.2-8) by parts, we obtain

$$J = \{\theta[\mathbf{x}(t), t] - \underline{\boldsymbol{\lambda}^T(t)\mathbf{x}(t)}\}\Big|_{t=t_o}^{t=t_f} + \underline{\int_{t_o}^{t_f}} \{H[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t] + \underline{\dot{\boldsymbol{\lambda}}^T\mathbf{x}(t)}\}dt \tag{4.2-9}$$

We now take the first variation of J for variations in the control vector and, consequently, in the state vector about the optimal control and optimal state vector. This gives us

$$\delta J = \left\{\delta\mathbf{x}^T\underset{\overset{\shortparallel}{0}}{\left[\frac{\partial\theta}{\partial\mathbf{x}} - \boldsymbol{\lambda}\right]}\right\}\Bigg|_{t=t_o}^{t=t_f} + \int_{t_o}^{t_f} \left\{\delta\mathbf{x}^T\underset{\overset{\shortparallel}{0}}{\left[\frac{\partial H}{\delta\mathbf{x}} + \dot{\boldsymbol{\lambda}}\right]} + \delta\mathbf{u}^T\underset{\overset{\shortparallel}{0}}{\left[\frac{\partial H}{\partial\mathbf{u}}\right]}\right\} dt \tag{4.2-10}$$

A necessary condition for a minimum is that the first variation in J vanish for arbitrary variations $\delta\mathbf{x}$ and $\delta\mathbf{u}$. Thus we have as the necessary condition for a minimum the very important relations

$$\delta\mathbf{x}^T\left[\frac{\partial\theta}{\partial\mathbf{x}} - \boldsymbol{\lambda}\right] = 0, \qquad \text{for} \quad t = t_o, t_f \tag{4.2-11}$$

from (4.2-7)

$\frac{\partial H}{\partial \lambda} = f = \dot{x}$

$$\dot{\boldsymbol{\lambda}} = -\frac{\partial H}{\partial\mathbf{x}}, \qquad \underline{\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t)} = \frac{\partial H}{\partial\boldsymbol{\lambda}} \tag{4.2-12}$$

original diff. eqn.

$$\frac{\partial H}{\partial\mathbf{u}} = \mathbf{0} \tag{4.2-13}$$

Since Eqs. (4.2-3) and (4.2-4), or alternate and perhaps more general expressions for the terminal manifold, may interrelate the components of the vector variation $\delta\mathbf{x}$ at the terminal time, and since an initial manifold may interrelate the components of the vector variation $\delta\mathbf{x}$ initially, Eq. (4.2-11) is the general statement for the transversality condition for the problem treated here. For a large class of optimal control problems, the initial state of the system is specified but the terminal state is unspecified. In that case, Eq. (4.2-11) yields the transversality conditions as

at $t=t_0$ if $x(t_0)=x_0$ then $\delta x' = 0$

$$\mathbf{x}(t_o) = \mathbf{x}_o, \qquad \boldsymbol{\lambda}(t_f) = \frac{\partial\theta[\mathbf{x}(t_f), t_f]}{\partial\mathbf{x}(t_f)} \tag{4.2-14}$$

→ terminal state is not fix[illegible] i.e., δx ⇒ $\frac{\partial\theta}{\partial x}$

since $\delta\mathbf{x}(t_o) = \mathbf{0}$, $\mathbf{x}(t_o)$ is fixed, and $\delta\mathbf{x}(t_f)$ is completely arbitrary. In another broad class of problems $\mathbf{x}(t_o)$ and $\mathbf{x}(t_f)$ are fixed. In this case $\delta\mathbf{x}(t_o)$ and $\delta\mathbf{x}(t_f)$ must be zero, and $\mathbf{x}(t_o)$ and $\mathbf{x}(t_f)$ are the boundary conditions for the two-point boundary value problem. For many estimation problems, neither $\mathbf{x}(t_o)$ nor $\mathbf{x}(t_f)$ are fixed (specified). In that case, Eq. (4.2-11) yields $\boldsymbol{\lambda}(t_o) = \boldsymbol{\lambda}(t_f) = \mathbf{0}$ as the boundary conditions for the problem since $\delta\mathbf{x}(t_o)$ and $\delta\mathbf{x}(t_f)$ are arbitrary. In still another case, we might have $\mathbf{x}(t_o) = \mathbf{x}_o$, $\theta = 0$, ? ?

and $\|\mathbf{x}(t_f)\| = 1$. In this event, it is easy for us to show that the final transversality conditions are obtained if we solve the two scalar equations, each in n variables.

$$\delta\mathbf{x}^T(t_f)\mathbf{x}(t_f) = 0, \qquad \delta\mathbf{x}^T(t_f)\,\boldsymbol{\lambda}(t_f) = 0 \tag{4.2-15}$$

We now give a more general and precise interpretation to the transversality conditions. For the general case where the initial manifold is

$$\mathbf{M}[\mathbf{x}(t_o), t_o] = 0 \tag{4.2-16}$$

Λ-vector is a linear combination of the n- initial conditions' q-vector

and the terminal manifold is

$$\mathbf{N}[\mathbf{x}(t_f), t_f] = 0 \tag{4.2-17}$$

we adjoin these conditions to the θ function by means of Lagrange multipliers, $\boldsymbol{\xi}$ and $\boldsymbol{\nu}$ and obtain for the cost function

$$\begin{aligned} J = \theta[\mathbf{x}(t), t]\Big|_{t=t_o}^{t=t_f} &- \boldsymbol{\xi}^T\mathbf{M}[\mathbf{x}(t_o), t_o] + \boldsymbol{\nu}^T\mathbf{N}[\mathbf{x}(t_f), t_f] \\ &+ \int_{t_o}^{t_f} \{H[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t] - \boldsymbol{\lambda}^T(t)\dot{\mathbf{x}}\}\, dt \end{aligned} \tag{4.2-18}$$

We now apply the usual variational techniques to obtain for the transversality conditions at the initial time:

$$\boldsymbol{\lambda}(t_o) = \frac{\partial\theta}{\partial\mathbf{x}} + \left(\frac{\partial\mathbf{M}^T}{\partial\mathbf{x}}\right)\boldsymbol{\xi}, \qquad \mathbf{M}[\mathbf{x}(t), t] = 0 \qquad t = t_o \tag{4.2-19}$$

The n initial conditions are obtained from this, with r parameters to be found in Eq. (4.2-19) such that we satisfy the r conditions of Eq. (4.2-16). In a similar fashion, the terminal condition is

$$\boldsymbol{\lambda}(t_f) = \frac{\partial\theta}{\partial\mathbf{x}} + \left(\frac{\partial\mathbf{N}^T}{\partial\mathbf{x}}\right)\boldsymbol{\nu}, \qquad \mathbf{N}[\mathbf{x}(t), t] = 0, \qquad t = t_f \tag{4.2-20}$$

n terminal conditions are obtained from this with q parameters $\boldsymbol{\nu}$ found in Eq. (4.2-20) such that the q conditions of Eq. (4.2-17) are satisfied.

The n vector differential equation obtained from Eq. (4.2-12) will be called the adjoint equation. Equation (4.2-13) provides the coupling relation between the original plant dynamics, Eq. (4.2-1), and the adjoint equation, the $\dot{\boldsymbol{\lambda}}$ equation of Eq. (4.2-12). This coupling equation was obtained from

$$\delta J = \cdots + \int_{t_o}^{t_f} \left\{\delta\mathbf{u}^T \frac{\partial H}{\partial\mathbf{u}} + \cdots\right\} dt$$

this term in (4.2.10)

and it is important to note that $\delta\mathbf{u}$ must be completely arbitrary in order for us to draw the conclusion that $\partial H/\partial\mathbf{u} = \mathbf{0}$ to obtain the optimal control. For the problem posed here where the admissible control set is infinite, $\delta\mathbf{u}$ can be completely arbitrary. Where the admissible control is bounded, $\delta\mathbf{u}$ cannot be completely arbitrary, and $\partial H/\partial\mathbf{u} = \mathbf{0}$ may not be the correct requirement. We will have more to say about this later. The solution we have obtained for this problem is a special case of the Po agin maximum principle.

It is also interesting to note that, since $H = \phi + \boldsymbol{\lambda}^T\mathbf{f}$, we may compute the total derivative with respect to time as

$$\frac{dH}{dt} = \frac{\partial\phi}{\partial t} + \dot{\mathbf{x}}^T\left[\frac{\partial\phi}{\partial\mathbf{x}} + \left(\frac{\partial\mathbf{f}^T}{\partial\mathbf{x}}\right)\boldsymbol{\lambda}\right] + \dot{\mathbf{u}}^T\left[\frac{\partial\phi}{\partial\mathbf{u}} + \left(\frac{\partial\mathbf{f}^T}{\partial\mathbf{u}}\right)\boldsymbol{\lambda}\right] + \dot{\boldsymbol{\lambda}}^T\mathbf{f} + \boldsymbol{\lambda}^T\frac{\partial\mathbf{f}}{\partial t} \tag{4.2-21}$$

but from Eqs. (4.2-12) and (4.2-7), we have

$$\dot{\boldsymbol{\lambda}} = -\frac{\partial H}{\partial\mathbf{x}} = -\frac{\partial\phi}{\partial\mathbf{x}} - \left(\frac{\partial\mathbf{f}^T}{\partial\mathbf{x}}\right)\boldsymbol{\lambda} \tag{4.2-22}$$

and from Eq. (4.2-7),

$$\frac{\partial H}{\partial\mathbf{u}} = \frac{\partial\phi}{\partial\mathbf{u}} + \left(\frac{\partial\mathbf{f}^T}{\partial\mathbf{u}}\right)\boldsymbol{\lambda} \tag{4.2-23}$$

Thus, since $\dot{\mathbf{x}}^T\dot{\boldsymbol{\lambda}} = \dot{\boldsymbol{\lambda}}^T\mathbf{f}$, Eq. (4.2-21) becomes

$$\frac{dH}{dt} = \frac{\partial\phi}{\partial t} + \boldsymbol{\lambda}^T\frac{\partial\mathbf{f}}{\partial t} + \dot{\mathbf{u}}^T\frac{\partial H}{\partial\mathbf{u}} \tag{4.2-24}$$

We see that, if ϕ and $\mathbf{f}$ are not explicit functions of time, the Hamiltonian is constant along an optimal trajectory where $\partial H/\partial\mathbf{u} = \mathbf{0}$. It can be shown that this is always true along an optimal trajectory, even if we cannot require $\partial H/\partial\mathbf{u} = \mathbf{0}$. We will make use of this fact in a later development. ?

In order that J be a minimum, the second variation of J must be nonnegative along all trajectories such that Eq. (4.2-1) is satisfied. Therefore we need to compute the second variation of J in Eq. (4.2-9) and impose the requirement that the variation of Eq. (4.2-1) is zero, or that

$$\delta\dot{\mathbf{x}} - \left(\frac{\partial\mathbf{f}^{\cancel{T}}}{\partial\mathbf{x}}\right)\delta\mathbf{x} - \left(\frac{\partial\mathbf{f}^{\cancel{T}}}{\partial\mathbf{u}}\right)\delta\mathbf{u} = 0 \tag{4.2-25}$$

Applying this condition and taking the quadratic part of the Taylor series expansion of $J(\mathbf{x} + \delta\mathbf{x}, \mathbf{u} + \delta\mathbf{u}) - J(\mathbf{x}, \mathbf{u})$, Eq. (4.1-4), we have for the second variation

$$\begin{aligned} \delta^2 J = \frac{1}{2}\left[\delta\mathbf{x}^T\frac{\partial^2\theta}{\partial\mathbf{x}^2}\,\delta\mathbf{x}\right]\Bigg|_{t=t_o}^{t=t_f} \\ + \frac{1}{2}\int_{t_o}^{t_f} [\delta\mathbf{x}^T\ \delta\mathbf{u}^T] \begin{bmatrix} \dfrac{\partial^2 H}{\partial\mathbf{x}^2} & \dfrac{\partial}{\partial\mathbf{u}}\dfrac{\partial H}{\partial\mathbf{x}} \\ \left[\dfrac{\partial}{\partial\mathbf{u}}\dfrac{\partial H}{\partial\mathbf{x}}\right]^T & \dfrac{\partial^2 H}{\partial\mathbf{u}^2} \end{bmatrix} \begin{bmatrix}\delta\mathbf{x} \\ \delta\mathbf{u}\end{bmatrix} dt \end{aligned} \tag{4.2-26}$$

and this must be nonnegative for a minimum. This will be the case if the $n + m$ square matrix under the integral sign and $\partial^2\theta/\partial\mathbf{x}^2$ are nonnegative definite.

i.e., the ab does not apply

Example 4.2-1

We are given the differential system consisting of three cascaded integrators

$$\begin{aligned} \dot{x}_1 &= x_2 & x_1(0) &= 0 \\ \dot{x}_2 &= x_3 & x_2(0) &= 0 \\ \dot{x}_3 &= u & x_3(0) &= 0 \end{aligned}$$

We wish to drive the system so that we reach the terminal manifold

$$x_1^2(1) + x_2^2(1) = 1$$

such that the cost function

$$J = \tfrac{1}{2} \int_0^1 u^2 \, dt$$

is minimized. The solution to the problem proceeds as follows. We compute the Hamiltonian from Eq. (4.2-7) as

$$H = \tfrac{1}{2}u^2 + \lambda_1 x_2 + \lambda_2 x_3 + \lambda_3 u$$

and determine the coupling relation, Eq. (4.2-13),

$$\frac{\partial H}{\partial u} = 0 = u + \lambda_3$$

and the adjoint Eq. (4.2-12)

$$\dot{\lambda}_1 = -\frac{\partial H}{\partial x_1} = 0$$

$$\dot{\lambda}_2 = -\frac{\partial H}{\partial x_2} = -\lambda_1$$

$$\dot{\lambda}_3 = -\frac{\partial H}{\partial x_3} = -\lambda_2$$

From Eqs. (4.2-17) and (4.2-20) we see that the transversality condition at the terminal time is

$$x_1^2(1) + x_2^2(1) = 1$$

$$\lambda(1) = \frac{\partial \theta}{\partial \mathbf{x}} + \left(\frac{\partial \mathbf{N}^T}{\partial \mathbf{x}}\right)\nu, \qquad t = t_f$$

where

$$N[\mathbf{x}(t_f), t_f] = x_1^2(t_f) + x_2^2(t_f) - 1 = 0, \qquad t_f = 1$$

Thus

$$\lambda(1) = \begin{bmatrix} \lambda_1(1) \\ \lambda_2(1) \\ \lambda_3(1) \end{bmatrix} = \begin{bmatrix} 2x_1(1)\nu \\ 2x_2(1)\nu \\ 0 \end{bmatrix}$$

Thus the problem of finding the optimal control and associated trajectories for this example is completely resolved when we solve the two-point boundary value problem represented by

$$\begin{aligned} \dot{x}_1 &= x_2 & x_1(0) &= 0 \\ \dot{x}_2 &= x_3 & x_2(0) &= 0 \\ \dot{x}_3 &= -\lambda_3 & x_3(0) &= 0 \\ \dot{\lambda}_1 &= 0 & \lambda_1(1) &= 2x_1(1)\nu \\ \dot{\lambda}_2 &= -\lambda_1 & \lambda_2(1) &= 2x_2(1)\nu \\ \dot{\lambda}_3 &= -\lambda_2 & \lambda_3(1) &= 0 \end{aligned} \qquad \left.\begin{matrix} \\ \\ \\ \\ \\ \end{matrix}\right\} x_1^2(1) + x_2^2(1) = 1$$

Although the six first-order differential equations represented above are perfectly linear and time invariant, the solution to this problem is complicated by the nonlinear nature of the terminal conditions. We shall discover various iterative schemes for overcoming this difficulty in later chapters.

4.2-2. Continuous optimal control problems—fixed beginning and unspecified terminal times—no inequality constraints

The material of the previous subsection may be easily extended to the case where the terminal manifold equation is a function of the terminal time and the terminal time is unspecified. For convenience, we will assume that the initial time and the initial state vector are specified. Solution may then easily be obtained for the case where the initial time and initial state vector are unspecified. Therefore the problem becomes one of minimizing the cost function

$$J = \theta[\mathbf{x}(t_f), t_f] + \int_{t_o}^{t_f} \phi[\mathbf{x}(t), \mathbf{u}(t), t] \, dt \tag{4.2-27}$$

for the system described by

$$\dot{\mathbf{x}} = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t], \qquad \mathbf{x}(t_o) = \mathbf{x}_o \tag{4.2-28}$$

where t_0 is fixed and where, at the unspecified terminal time $t = t_f$, the q vector terminal manifold equation

$$\mathbf{N}[\mathbf{x}(t_f), t_f] = \mathbf{0} \tag{4.2-29}$$

is satisfied. It may be noted here that the terminal manifold line, $x(t_f) = c(t_f)$, of the previous chapter becomes here $\mathbf{N}[\mathbf{x}(t_f), t_f] = \mathbf{0}$ which is more general. We adjoin the equality constraints to the cost function via Lagrange multipliers to obtain

$$\begin{aligned} J = \theta[\mathbf{x}(t_f), t_f] + \boldsymbol{\nu}^T\mathbf{N}[\mathbf{x}(t_f), t_f] + \int_{t_o}^{t_f} \{\phi[\mathbf{x}(t), \mathbf{u}(t), t] \\ + \boldsymbol{\lambda}^T(t)[\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] - \dot{\mathbf{x}}]\} \, dt \end{aligned} \tag{4.2-30}$$

As before, we define the Hamiltonian

$$H[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t] = \phi[\mathbf{x}(t), \mathbf{u}(t), t] + \boldsymbol{\lambda}^T(t)\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t]$$

and integrate a portion of the cost function, Eq. (4.2-30), to obtain

$$\begin{aligned} J = \theta[\mathbf{x}(t_f), t_f] + \boldsymbol{\nu}^T\mathbf{N}[\mathbf{x}(t_f), t_f] - \boldsymbol{\lambda}^T(t_f)\mathbf{x}(t_f) + \boldsymbol{\lambda}^T(t_o)\mathbf{x}(t_o) \\ + \int_{t_o}^{t_f} \{H[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t] + \dot{\boldsymbol{\lambda}}^T\mathbf{x}(t)\} \, dt \end{aligned} \tag{4.2-31}$$

We agai[illegible]m the first variation by letting

$$\mathbf{x}(t) = \hat{\mathbf{x}}(t) + \mathbf{h}(t), \qquad \mathbf{u}(t) = \hat{\mathbf{u}}(t) + \delta\mathbf{u}(t), \qquad t_f = \hat{t}_f + \delta t_f \tag{4.2-32}$$

and then we form the difference $J[\mathbf{x}, \mathbf{u}, t_f] - J[\hat{\mathbf{x}}, \hat{\mathbf{u}}, \hat{t}_f]$ and retain only the linear terms. Thus we have, after dropping the ^ notation for convenience,

$$\begin{aligned}\delta J = \delta t_f &\left\{H[\mathbf{x}(t_f), \mathbf{u}(t_f), \boldsymbol{\lambda}(t_f), t_f] + \frac{\partial\Theta}{\partial t_f}\right\} \\ &+ \delta\mathbf{x}^T(t_f)\left\{\frac{\partial\Theta}{\partial\mathbf{x}} - \boldsymbol{\lambda}(t_f)\right\} \\ &+ \int_{t_o}^{t_f}\left\{\mathbf{h}^T(t)\left[\frac{\partial H}{\partial\mathbf{x}} + \dot{\boldsymbol{\lambda}}\right] + \delta\mathbf{u}^T(t)\left[\frac{\partial H}{\partial\mathbf{u}}\right]\right\} dt\end{aligned} \tag{4.2-33}$$

where

$$\Theta[\mathbf{x}(t_f), \boldsymbol{\nu}, t_f] = \theta[\mathbf{x}(t_f), t_f] + \boldsymbol{\nu}^T\mathbf{N}[\mathbf{x}(t_f), t_f] \tag{4.2-34}$$

We must set this first variation equal to zero to obtain the necessary conditions for a minimum. Therefore, the equations which determine the optimal control and state vector are

$$H = \phi[\mathbf{x}(t), \mathbf{u}(t), t] + \boldsymbol{\lambda}^T(t)\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] \tag{4.2-35}$$

$$\frac{\partial H}{\partial\boldsymbol{\lambda}} = \dot{\mathbf{x}} = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] \tag{4.2-36}$$

$$\frac{\partial H}{\partial\mathbf{x}} = -\dot{\boldsymbol{\lambda}} = \frac{\partial\mathbf{f}^T[\mathbf{x}(t), \mathbf{u}(t), t]}{\partial\mathbf{x}}\boldsymbol{\lambda}(t) + \frac{\partial\phi[\mathbf{x}(t), \mathbf{u}(t), t]}{\partial\mathbf{x}} \tag{4.2-37}$$

$$\frac{\partial H}{\partial\mathbf{u}} = 0 = \frac{\partial\phi[\mathbf{x}(t), \mathbf{u}(t), t]}{\partial\mathbf{u}} + \frac{\partial\mathbf{f}^T[\mathbf{x}(t), \mathbf{u}(t), t]}{\partial\mathbf{u}}\boldsymbol{\lambda}(t) \tag{4.2-38}$$

These represent the $2n$ differential equations for the two-point boundary value problems. The conditions at the initial time are

$$\mathbf{x}(t_o) = \mathbf{x}_o \tag{4.2-39}$$

whereas those at the final time are

$$\boldsymbol{\lambda}(t_f) = \frac{\partial\Theta}{\partial\mathbf{x}(t_f)} = \frac{\partial\theta}{\partial\mathbf{x}(t_f)} + \left[\frac{\partial\mathbf{N}^T}{\partial\mathbf{x}(t_f)}\right]\boldsymbol{\nu} \tag{4.2-40}$$

$$\mathbf{N}[\mathbf{x}(t_f), t_f] = \mathbf{0} \tag{4.2-41}$$

and

$$H[\mathbf{x}(t_f), \mathbf{u}(t_f), \boldsymbol{\lambda}(t_f), t_f] + \frac{\partial\theta}{\partial t_f} + \left(\frac{\partial\mathbf{N}^T}{\partial t_f}\right)\boldsymbol{\nu} = 0 \tag{4.2-42}$$

Equation (4.2-40) provides n conditions with q Lagrange multipliers to be determined. Equation (4.2-41) provides q equations to eliminate the Lagrange multipliers, and Eq. (4.2-42) provides the one additional equation which we must have to determine the unspecified terminal time.

Example 4.2-2

For the first-order single integration system

$$\dot{x} = u, \qquad x(0) = 1$$

we desire to find the control $u(t)$ which makes $x(t_f) = 0$, where t_f is unspecified, such as to make, for specified values of α and β,

$$J = t_f^\alpha + \tfrac{1}{2}\beta\int_0^{t_f} u^2\,dt$$

a minimum. For this problem

$$N[x(t_f), t_f] = x(t_f) = 0, \qquad \phi = \tfrac{1}{2}\beta u^2$$

$$\theta = t_f^\alpha, \qquad H = \tfrac{1}{2}\beta u^2 + \lambda u$$

The canonic equations are

$$\dot{x} = u = -\frac{\lambda}{\beta}, \qquad \dot{\lambda} = 0$$

with the boundary conditions $x(0) = 0$, $x(t_f) = 0$, where we determine the final time by solving Eq. (4.2-42) which becomes, for this example,

$$-\frac{\lambda^2(t_f)}{2\beta} + \alpha t_f^{\alpha-1} = 0$$

The solutions to the canonic equations are

$$x(t) = -\frac{\lambda(t_f)t}{\beta} + 1, \qquad \lambda(t) = \lambda(t_f)$$

But since $x(t_f) = 0$; $t_f = \beta\lambda^{-1}(t_f)$, and in the particular case where $\beta = \alpha = 1$, we can easily show from the a foregoing that $\lambda(t_f) = +(2)^{1/2}$, which determines the solution to this example. The optimum control is $u(t) = -\lambda(t) = -2^{1/2}$. The corresponding trajectory is $x(t) = 1 - 2^{1/2}t$, with $t_f = 2^{-1/2}$.

Example 4.2-3

A problem which will be of considerable interest to us later will be the "minimum time" problem. In that case

$$\theta[x(t_f), t_f] = t_f, \qquad \phi = 0$$

and we specify the optimal control and corresponding trajectory by solving Eqs. (4.2-35) through (4.2-38), which become

$$H[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t] = \boldsymbol{\lambda}^T(t)\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t]$$

$$\frac{\partial H}{\partial\boldsymbol{\lambda}} = \dot{\mathbf{x}} = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t]$$

$$\frac{\partial H}{\partial\mathbf{x}} = -\dot{\boldsymbol{\lambda}} = \frac{\partial\mathbf{f}^T[\mathbf{x}(t), \mathbf{u}(t), t]}{\partial\mathbf{x}}\boldsymbol{\lambda}(t)$$

$$\frac{\partial H}{\partial\mathbf{u}} = 0 = \frac{\partial\mathbf{f}^T[\mathbf{x}(t), \mathbf{u}(t), t]}{\partial\mathbf{u}}\boldsymbol{\lambda}(t)$$

with the boundary conditions specified by Eqs. (4.2-39) through (4.2-42)

$$\mathbf{x}(t_o) = \mathbf{x}_o$$

$$\boldsymbol{\lambda}(t_f) = \frac{\partial \mathbf{N}^T}{\partial \mathbf{x}(t_f)} \nu$$

$$\mathbf{N}[\mathbf{x}(t_f), t_f] = 0$$

$$H[\mathbf{x}(t_f), \mathbf{u}(t_f), t_f] = -1 - \left(\frac{\partial \mathbf{N}^T}{\partial t_f}\right) \nu$$

In many cases, the system is brought to rest at the unspecified time, and the terminal manifold is the origin, so that

$$\mathbf{N}[\mathbf{x}(t_f), t_f] = \mathbf{x}(t_f) = \mathbf{0}$$

Then the foregoing expressions reduce to

$$\mathbf{x}(t_o) = \mathbf{x}_o, \qquad \mathbf{x}(t_f) = \mathbf{0}$$

$$H[\mathbf{x}(t_f), \mathbf{u}(t_f), \boldsymbol{\lambda}(t_f), t_f] = -1$$

If the Hamiltonian is not an explicit function of time, Eq (4.2-24), which applies here as well, yields $dH/dt = 0$; therefore, for this minimum time problem

$$H[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t] = H[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t)] = -1$$

It should be emphasized that we are not solving the usual minimum time problem since we have imposed no inequality constraints on the control (or state) variables. An alternate version of this problem would be to consider $\theta = 0$ and $\phi = 1$. This changes the Hamiltonian for this particular problem, but it certainly does not change the optimal control and state vector, as the reader can easily verify.

4.3 The Bolza problem with control and state variable inequality constraints—the Pontryagin maximum principle

In the prior work in this chapter we treated the Bolza problem with no inequality constraints present on either the control or the state variable. We found for example that a minimum of

$$J = \theta[\mathbf{x}(t_f), t_f] + \int_{t_o}^{t_f} \phi[\mathbf{x}(t), \mathbf{u}(t), t]\, dt$$

for a system described by

$$\dot{\mathbf{x}} = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t], \qquad \mathbf{x}(t_o) = \mathbf{x}_o$$

with t_o and t_f fixed may be obtained if we define a Hamiltonian as

$$H[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t] = \phi[\mathbf{x}(t), \mathbf{u}(t), t] + \boldsymbol{\lambda}^T(t)\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t]$$

and set

$$\frac{\partial H}{\partial \boldsymbol{\lambda}} = \dot{\mathbf{x}} \qquad \mathbf{x}(t_o) = \mathbf{x}_o$$

$$\frac{\partial H}{\partial \mathbf{x}} = -\dot{\boldsymbol{\lambda}} \qquad \boldsymbol{\lambda}(t_f) = \frac{\partial \theta[\mathbf{x}(t_f), t_f]}{\partial \mathbf{x}(t_f)}$$

$$\frac{\partial H}{\partial \mathbf{u}} = 0$$

If the admissible control vector is unrestricted, then the first variation of $\mathbf{u}(t)$, $\delta\mathbf{u}(t)$, is also unrestricted, and in that part of Eq. (4.2-10) which reads

$$\int_{t_o}^{t_f} [\delta\mathbf{u}(t)]^T \left[\frac{\partial H}{\partial \mathbf{u}}\right] dt + \cdots = 0$$

we are free to set $\partial H/\partial \mathbf{u}$ equal to zero. Sections 4.1 and 4.2 describe a special case of the maximum principle where this is possible. In many problems, inequality constraints on the admissible control vector (the maximum thrust from a reaction jet is limited, for example) are present, and we must therefore take this into account if we are to determine a realistic control strategy. If $\mathbf{u}(t)$ is constrained, $\delta\mathbf{u}(t)$ may not be allowed to be completely arbitrary, and therefore we may not in general set $\partial H/\partial \mathbf{u} = 0$. Also, certain regions of the state space may be prohibited, and we must determine an optimum control such that the state $\mathbf{x}(t)$ does not enter the forbidden regions. We examined a portion of this problem in Chapter 3 and found that we could handle inequality constraints by converting them to equivalent equality constraints. In this section, we desire to find the state and control vector such that the cost function

$$J = \theta[\mathbf{x}(t_f), t_f] + \int_{t_o}^{t_f} \phi[\mathbf{x}(t), \mathbf{u}(t), t]\, dt \tag{4.3-1}$$

is minimized subject to

(a) the n differential system equality constraints

$$\dot{\mathbf{x}} = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] \tag{4.3-2}$$

(b) the q end point equality constraints ($q \leq n$) at the terminal time (which may be unspecified)

$$\mathbf{N}[\mathbf{x}(t_f), t_f] = 0 \tag{4.3-3}$$

and the initial condition equality constraint

$$\mathbf{x}(t_o) = \mathbf{x}_o \tag{4.3-4}$$

where we assume that t_o is fixed and $\mathbf{x}(t_o)$ is known. Actually, t_o does not have to be fixed and the initial condition constraint can be $\mathbf{M}[\mathbf{x}(t_o), t_o] = \mathbf{0}$, as was the case in Section 4.2. The required modifications to treat this case are small since the results are so similar to the variable end-point and variable end-time case.

(c) The r admissible control inequality constraints ($r \leq m$)

$$\mathbf{g}[\mathbf{x}(t), \mathbf{u}(t), t] \geq 0 \tag{4.3-5}$$

where we will find it necessary to impose the requirement that the ma $\partial \mathbf{g}/\partial \mathbf{u}$ be of maximum rank whenever $\mathbf{g} = \mathbf{0}$.

(d) The s inequality constraints (with no control component in the constraint) expressing the forbidden region of state space

$$\mathbf{h}[\mathbf{x}(t), t] \geq \mathbf{0} \tag{4.3-6}$$

which does not satisfy the maximum rank test in (c).

As is apparent, we have formulated a rather formidable problem in the variational calculus. We will solve the problem in such a fashion that we obtain the Pontryagin maximum principle [2, 3, 4, 5]. However, due to a slight change in the original problem statement, a more appropriate name for the result of our development would be the Pontryagin minimum principle. Our development will be patterned after that of Berkovitz who has unified many of the approaches to the optimal control problem [6, 7]. We will first consider the case where the inequalities of part (d) on the admissible regions of state space are not present and will then modify our maximum principle and associated transversality conditions to include this important case.

4.3-1. The maximum principle with control variable inequality constraints

We now wish to derive the first necessary condition for a minimum of the problem just posed, except that we will assume that there are no bounded state variables. Thus we are considering the first three of the four constraints just mentioned. Constraint (c) is very similar to the inequality constraint of Section 3.8, and we now find it desirable to expand upon that method of treating an inequality constraint.

We are given the inequality constraint

$$\mathbf{g}[\mathbf{x}(t), \mathbf{u}(t), t] \geq \mathbf{0} \tag{4.3-7}$$

We may convert this inequality constraint to an equality constraint by writing for each component of g either

$$(\dot{z}_i)^2 = g_i[\mathbf{x}(t), \mathbf{u}(t), t], \qquad z_i(t_o) = 0, \qquad i = 1, 2, \ldots, r \tag{4.3-8}$$

or

$$(y_i)^2 = g_i[\mathbf{x}(t), \mathbf{u}(t), t] \qquad i = 1, 2, \ldots, r \tag{4.3-9}$$

It is apparent that either of these two equations force g_i to be greater than or equal to zero since $(\dot{z}_i)^2$ and $(y_i)^2$ must certainly be greater than or equal to zero. This technique was apparently first proposed by Valentine [8] and extended by Berkovitz [6]. It is quite similar to the penalty function technique of Kelly [9] as we shall see in our chapter concerning the gradient and second variation methods for the computation of optimal controls. The choice between Eqs. (4.3-8) and (4.3-9) will depend largely upon the particular computer (for an analog computer, Eq. (4.3-8) is generally easier to implement than Eq. (4.3-9)) and the particular computa al algorithms used (for the quasilinearization method, Eq. (4.3-9) is considerably simpler to use than Eq. (4.3-8) and also results in less computer solution time).

Example 4.3-1

It is quite easy to see that the constraint used here includes, as a special case, that considered in Section 3.8. For example, if we require for a scalar control u, $u_{\min} \leq u \leq u_{\max}$, then we may write

$$g_1[\mathbf{x}(t), u(t), t] = u_{\max} - u \geq 0, \qquad g_2[\mathbf{x}(t), u(t), t] = u - u_{\min} \geq 0$$

and we convert these inequality constraints to equality constraints by writing

$$(y_1)^2 = u_{\max} - u, \qquad (y_2)^2 = u - u_{\min}$$

for which

$$(y_1 y_2)^2 = (u_{\max} - u)(u - u_{\min})$$

which is precisely the constraint used in Section 3.8.

For the problem at hand, we adjoin, via the Lagrange multiplier, constraints (4.3-2), (4.3-3), (4.3-4), and (4.3-5) to Eq. (4.3-1) to obtain

$$\begin{aligned} J = \theta[\mathbf{x}(t_f), t_f] + \boldsymbol{\xi}^T[\mathbf{x}(t_o)] + \boldsymbol{\nu}^T\mathbf{N}[\mathbf{x}(t_f), t_f] \\ + \int_{t_o}^{t_f} \Big\{ H[\mathbf{x}(t), \dot{\mathbf{w}}(t), \boldsymbol{\lambda}(t), t] - \boldsymbol{\lambda}^T(t)\dot{\mathbf{x}} \\ - \boldsymbol{\Gamma}^T(t)\{\mathbf{g}[\mathbf{x}(t), \dot{\mathbf{w}}(t), t] - \dot{\mathbf{z}}^2\} \Big\}\, dt \end{aligned} \tag{4.3-10}$$

where

$$(\mathbf{z}^2)^T = [z_1^2, z_2^2, z_3^2, \ldots, z_r^2] \tag{4.3-11}$$

$$H[\mathbf{x}(t), \dot{\mathbf{w}}(t), \boldsymbol{\lambda}(t), t] = \phi[\mathbf{x}(t), \dot{\mathbf{w}}(t), t] + \boldsymbol{\lambda}^T(t)\mathbf{f}[\mathbf{x}(t), \dot{\mathbf{w}}(t), t] \tag{4.3-12}$$

$$\dot{\mathbf{w}} = \mathbf{u}(t), \qquad \mathbf{w}(t_o) = \mathbf{0} \tag{4.3-13}$$

We may now apply the Euler-Lagrange equations to the above cost function or take a first variation of it in order to obtain the necessary conditions for a minimum. It is thus convenient to define a scalar function Φ, the Lagrangian, as

$$\begin{aligned} \Phi[\mathbf{x}(t), \dot{\mathbf{x}}(t), \dot{\mathbf{w}}(t), \boldsymbol{\lambda}(t), \boldsymbol{\Gamma}(t), \dot{\mathbf{z}}(t), t] = H[\mathbf{x}(t), \dot{\mathbf{w}}(t), \boldsymbol{\lambda}(t), t] \\ - \boldsymbol{\lambda}^T(t)\dot{\mathbf{x}} - \boldsymbol{\Gamma}^T(t)\{\mathbf{g}[\mathbf{x}(t), \dot{\mathbf{w}}(t), t] - \dot{\mathbf{z}}^2\} \end{aligned} \tag{4.3-14}$$

We will use the Euler-Lagrange Eqs. (3.5-3). Since there are no $\mathbf{w}(t)$ and $\mathbf{z}(t)$ terms in Eq. (4.3-14), we may write the Euler-Lagrange equations as

$$\frac{d}{dt}\frac{\partial \Phi}{\partial \dot{\mathbf{x}}} - \frac{\partial \Phi}{\partial \mathbf{x}} = \mathbf{0} \tag{4.3-15}$$

$$\frac{d}{dt}\frac{\partial \Phi}{\partial \dot{\mathbf{w}}} = \mathbf{0} \tag{4.3-16}$$

$$\frac{d}{dt}\frac{\partial \Phi}{\partial \dot{\mathbf{z}}} = \mathbf{0} \tag{4.3-17}$$

Each piecewise continuously differentiable solution of the Euler-Lagrange equations (4.3-15), (4.3-16), and (4.3-17) will be called an extremal curve or an extremal trajectory of the associated variational problem. It can be shown that the function Φ need be only piecewise smooth, and thus the Euler-Lagrange equations require that every arc of the extremal trajectory on which the first derivatives of Φ have no discontinuities be a solution of the Euler-Lagrange equations. The corner condition will answer our questions concerning what happens at possible points of discontinuity of some of the derivatives of the state or control variables. This corner condition will ensure continuity of the state and control variables by forcing $\partial\Phi/\partial\dot{\mathbf{z}}$ to be zero everywhere since it is zero at the terminal time.

The transversality conditions for this problem are obtained in the usual fashion as explained in Chapter 3 and the previous three sections. For this problem, they are easily shown to be Eqs. (4.3-3), (4.3-4), and

$$\frac{\partial\theta}{\partial t_f} + \left(\frac{\partial \mathbf{N}^T}{\partial t_f}\right)\nu + \phi - \dot{\mathbf{x}}^T\frac{\partial\Phi}{\partial\dot{\mathbf{x}}} = 0, \quad \text{for} \quad t = t_f \tag{4.3-18}$$

$$\frac{\partial\theta}{\partial \mathbf{x}} + \left(\frac{\partial \mathbf{N}^T}{\partial \mathbf{x}}\right)\nu - \boldsymbol{\lambda} = \mathbf{0}, \quad \text{for} \quad t = t_f \tag{4.3-19}$$

Also, we have for the final transversality condition

$$\delta\mathbf{z}^T(t_f)\left[\frac{\partial\Phi}{\partial\dot{\mathbf{z}}}\right] = \delta\mathbf{z}^T(t_f)\begin{bmatrix} 2\Gamma_1\dot{z}_1 \\ 2\Gamma_2\dot{z}_2 \\ \cdot \\ \cdot \\ \cdot \\ 2\Gamma_r\dot{z}_r \end{bmatrix} = 0, \qquad \text{for} \quad t = t_f$$

which allows us to write because of Eq. (4.3-17)

$$\frac{\partial\Phi}{\partial\dot{\mathbf{z}}} = \mathbf{0}, \qquad \forall t \in [t_o, t_f]$$

Since when $\Gamma_i \neq 0$, $\dot{z}_i = 0 = g_i$, and when $\dot{z}_i \neq 0$, $\Gamma_i = 0$

$$\Gamma_i\dot{z}_i = 0, \qquad i = 1, 2, \ldots, r, \qquad \forall t \in [t_o, t_f] \tag{4.3-20}$$

Also, with similar reasoning, we have

$$\frac{\partial\Phi}{\partial\dot{\mathbf{w}}} = \mathbf{0}, \qquad \forall t \in [t_o, t_f] \tag{4.3-21}$$

We shall now introduce the Hamiltonian formulation and use the Weierstrass condition to obtain the Pontryagin maximum principle. From the definition of Φ, Eq. (4.3-14), Eq. (4.3-15) yields

$$\dot{\boldsymbol{\lambda}} = -\frac{\partial H}{\partial \mathbf{x}} + \frac{\partial \mathbf{g}^T}{\partial \mathbf{x}}\boldsymbol{\Gamma} \tag{4.3-22}$$

Equation (4.3-16) with the definition of Φ, Eq. (4.3-14), gives us

$$\frac{\partial H}{\partial \dot{\mathbf{w}}} - \frac{\partial \mathbf{g}^T}{\partial \dot{\mathbf{w}}}\boldsymbol{\Gamma} = \mathbf{0} \tag{4.3-23}$$

and in a similar fashion, Eq. (4.3-17) results in

$\dot{Z}_i$

$$\Gamma_i\ddot{z}_i = 0, \qquad i = 1, 2, \ldots, r \tag{4.3-24}$$

Since Eq. (4.3-14), when solved for H, yields

$$\begin{aligned} H[\mathbf{x}, \dot{\mathbf{w}}, \boldsymbol{\lambda}, t] = \Phi[&\mathbf{x}(t), \dot{\mathbf{x}}(t), \dot{\mathbf{w}}(t), \boldsymbol{\lambda}(t), \boldsymbol{\Gamma}(t), \dot{\mathbf{z}}(t), t] \\ &+ \boldsymbol{\lambda}^T(t)\dot{\mathbf{x}} + \boldsymbol{\Gamma}^T(t)\{\mathbf{g}[\mathbf{x}(t), \dot{\mathbf{w}}(t), t] - \dot{\mathbf{z}}^2\} \end{aligned}$$

we can show that

$$H = \Phi - \dot{\mathbf{x}}^T\frac{\partial\Phi}{\partial\dot{\mathbf{x}}} - \dot{\mathbf{w}}\frac{\partial\Phi}{\partial\dot{\mathbf{w}}} - \dot{\mathbf{z}}\frac{\partial\Phi}{\partial\dot{\mathbf{z}}} \tag{4.3-25}$$

because we know that $\dot{\mathbf{z}}^2 = \mathbf{g}$, $\dot{\boldsymbol{\lambda}} = \partial\Phi/\partial\mathbf{x}$, and have just found $\partial\Phi/\partial\dot{\mathbf{w}} = \mathbf{0}$ and $\partial\Phi/\partial\dot{\mathbf{z}} = \mathbf{0}$. This is in a form for direct application of the Weierstrass condition, Eq. (4.1-21), which can be written as

$$\begin{aligned} \Phi(\mathbf{x}, \mathbf{w}, \mathbf{z}, \dot{X}, \dot{W}, \dot{Z}) - \Phi(\mathbf{x}, \mathbf{w}, \mathbf{z}, \dot{\mathbf{x}}, \dot{\mathbf{w}}, \dot{\mathbf{z}}) - (\dot{X} - \dot{\mathbf{x}})'\frac{\partial\Phi}{\partial\dot{\mathbf{x}}} \\ - (\dot{W} - \dot{\mathbf{w}})'\frac{\partial\Phi}{\partial\dot{\mathbf{w}}} - (\dot{Z} - \dot{\mathbf{z}})'\frac{\partial\Phi}{\partial\dot{\mathbf{z}}} \geq 0 \end{aligned} \tag{4.3-26}$$

where lower-case symbols indicate optimum vectors and upper-case symbols indicate admissible vectors, as before. From Eq. (4.3-25), it becomes apparent that this condition is equivalent to

$$H[\mathbf{x}, \dot{W}, \boldsymbol{\lambda}, t] \geq H[\mathbf{x}, \dot{\mathbf{w}}, \boldsymbol{\lambda}, t] \tag{4.3-27}$$

In other words, the Hamiltonian is smaller when we use the optimal control within the admissible set of controls than it is for any other control which is in this admissible set. This is the basic contribution of the maximum principle —a necessary condition for optimality is the global minimization of the Hamiltonian, H, function.

4.3-2. Summary of the maximum principle

Since our development of the maximum principle has been necessarily long, it is desirable to give a summary of the results. It is also important to note that we can successfully use the maximum principle without following each and every detail of our "proof."

We wish to minimize

$$J = \theta[\mathbf{x}(t_f), t_f] + \int_{t_o}^{t_f} \phi[\mathbf{x}(t), \mathbf{u}(t), t]\, dt \tag{4.3-28}$$

for the system described by

$$\dot{\mathbf{x}} = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] \tag{4.3-29}$$

$$\mathbf{x}(t_o) = \mathbf{x}_o, \qquad t_o \text{ fixed} \tag{4.3-30}$$

such that, at the unspecified terminal time t_f,

$$\mathbf{N}[\mathbf{x}(t_f), t_f] = \mathbf{0} \tag{4.3-31}$$

and where u [illegible]s restricted such that under this condition

$$g[u(t), t] \geq 0 \tag{4.3-32}$$

In other words, u(*t*) is not restricted in control space as a function of the state vector, x(*t*), and

$$u \in \mho \tag{4.3-33}$$

The Hamilton canonic equations, solution of which minimizes the cost function and determines the optimum state and control vectors, x(*t*) and u(*t*), may be obtained if we define a Hamiltonian

$$H[x(t), u(t), \lambda(t), t] = \phi[x(t), u(t), t] + \lambda^T(t) f[x(t), u(t), t] \tag{4.3-34}$$

and then set the Hamiltonian with $u = \hat{u}$ less than any other value of H with $u \in \mho$.

Hamiltonian be min when [illegible]ated at the [illegible]al y.

$$H[x(t), \hat{u}(t), \lambda(t), t] \leq \underset{u \in \mho}{H}[x(t), u(t), \lambda(t), t] \tag{4.3-35}$$

$$\frac{\partial H}{\partial \lambda} = \dot{x} \tag{4.3-36}$$

$$\frac{\partial H}{\partial x} = -\dot{\lambda} \tag{4.3-37}$$

subject to the two-point boundary conditions

$$x(t_o) = x_o \tag{4.3-38}$$

$$N[x(t_f), t_f] = 0 \tag{4.3-39}$$

$$\frac{\partial \theta}{\partial t_f} + \left(\frac{\partial N^T}{\partial t_f}\right)\nu + H = 0, \quad \text{at} \quad t = t_f \tag{4.3-40}$$

$$\frac{\partial \theta}{\partial x} + \left(\frac{\partial N^T}{\partial x}\right)\nu - \lambda = 0, \quad \text{at} \quad t = t_f \tag{4.3-41}$$

We frequently wish to transfer the system to the origin in minimum time so that we have

$$N[x(t_f), t_f] = 0 = x(t_f) \tag{4.3-42}$$

$$\theta[x(t_f), t_f] = t_f \tag{4.3-43}$$

$$\phi = 0 \tag{4.3-44}$$

In this particular case, the transverality conditions become

m 4.3-40

$$x(t_o) = x_o \tag{4.3-45}$$

$$x(t_f) = 0 \tag{4.3-46}$$

=1 ⇒ 1 + 0 + H = 0 at t = t_f

$$\Rightarrow H = -1, \quad \text{at} \quad t = t_f \tag{4.3-47}$$

Example 4.3-2

Let us consider briefly the time optimal control poblem for a linear time-invariant system where the length of the control vector is constrained. We wish to minimize

$$J = t_f$$

for the system

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

$$x(t_o) = x_o$$

where $u(t) \in \mho$ means $\|u(t)\| \leq 1$.

The Hamiltonian, Eq. (4.3-34), becomes

$$H[x(t), u(t), \lambda(t), t] = \lambda^T(t)[Ax(t) + Bu(t)]$$

To make H as small as possible with respect to a choice of u(*t*), we must have

$$u(t) = \frac{-B^T\lambda(t)}{\|B^T\lambda(t)\|}$$

The canonic equations become

$$\frac{\partial H}{\partial \lambda} = \dot{x} = Ax(t) + Bu(t), \qquad \frac{\partial H}{\partial x} = -\dot{\lambda} = A^T\lambda(t)$$

with the boundary conditions

$$x(t_o) = x_o, \qquad x(t_f) = 0$$

where we determine t_f by solving

$$H[x(t_f), \lambda(t_f), u(t_f)] = -1$$

But, from Eq. (4.2-24) we see that $dH/dt = 0$ since the Hamiltonian does not depend explicitly on t. Thus the above equation becomes

$$H[x(t), u(t), \lambda(t)] = -1 = \lambda^T(t)[Ax(t) + Bu(t)]$$

which is the additional relation needed to determine the terminal time.

4.3-3. The maximum principle with state (and control) variable inequality constraints

We now wish to extend the work of Section 4.3-1 to include inequality constraints on some or all of the state variables. We will represent this inequality constraint by the s vector equation

$$h[x(t), t] \geq 0 \tag{4.3-48}$$

where each component of h is assumed to be continuously differentiable in state space. There are several methods whereby we may convert Eq. (4.3-48) to an equality constraint. We may define a new variable x_{n+1} by

$$\begin{aligned} \dot{x}_{n+1} = f_{n+1} &= [h_1(x, t)]^2 H(h_1) + [h_2(x, t)]^2 H(h_2) \\ &+ \cdots + [h_s(x, t)]^2 H(h_s) \end{aligned} \tag{4.3-49}$$

where $H[h_s(x, t)]$ is a modified Heaviside step defined such that

$$H[h_s(x, t)] = \begin{cases} 0 & \text{if} \quad h_s(x, t) \geq 0 \\ K_s & \text{if} \quad h_s(x, t) < 0 \end{cases}$$

$$K_s > 0, \qquad s = 1, 2, \ldots, s \tag{4.3-50}$$

and where the initial condition is

$$x_{n+1}(t_o) = 0 \tag{4.3-51}$$

Thus we see that $x_{n+1}(t_f)$ is a direct measure of penetration of the state variable inequality constraint

$$x_{n+1}(t_f) = \int_{t_o}^{t_f} \dot{x}_{n+1}(t)\,dt = \int_{t_o}^{t_f} \{[h_1(\mathbf{x}, t)]^2 \mathrm{H}(h_1) + \cdots + [h_s(\mathbf{x}, t)]^2 \mathrm{H}(h_s)\}\,dt \tag{4.3-52}$$

We will require that the final value of $x_{n+1}(t_f)$ is zero,

$$x_{n+1}(t_f) = 0 \tag{4.3-53}$$

which will impose the restriction that we do not violate the inequality constraint. This approach is a modification by McGill [10] of a similar procedure by Kelley [9] which converts the s inequality constraint to s equality constraints of the form

$$\begin{aligned} \dot{x}_{n+1} &= [h_1(\mathbf{x}, t)]^2 \mathrm{H}(h_1), & x_{n+1}(t_o) &= 0 \\ \dot{x}_{n+2} &= [h_2(\mathbf{x}, t)]^2 \mathrm{H}(h_2), & x_{n+2}(t_o) &= 0 \\ &\;\;\vdots & &\;\;\vdots \\ \dot{x}_{n+s} &= [h_s(\mathbf{x}, t)]^2 \mathrm{H}(h_s), & x_{n+s}(t_o) &= 0 \end{aligned} \tag{4.3-54}$$

which are then added to the cost function to obtain

$$J_{\text{modified}} = J_{\text{original}} + \sum_{j=1}^{s} x_{n+j}(t_f) \tag{4.3-55}$$

The multipliers K_s are thus the penalty functions, and J_{modified} is minimized such that the constraint region is entered only slightly, if at all. If we require $x_{n+j}(t_f) = 0$ for $j = 1, 2, \ldots, s$, the constraint is of course not exceeded at all.

A slight modification of the penalty-function approach can be obtained if we define s new state variables

$$\begin{aligned} (\dot{x}_{n+1})^2 &= K_1 h_1(\mathbf{x}, t), & x_{n+1}(t_o) &= 0 \\ (\dot{x}_{n+2})^2 &= K_2 h_2(\mathbf{x}, t), & x_{n+2}(t_o) &= 0 \\ &\;\;\vdots & &\;\;\vdots \\ (\dot{x}_{n+s})^2 &= K_s h_s(\mathbf{x}, t), & x_{n+s}(t_o) &= 0 \end{aligned} \tag{4.3-56}$$

Berkovitz [7] suggests yet another method for converting the inequality constraint to an equality constraint. For the case of a scalar constraint, a variable

$$\gamma(\mathbf{x}, \eta, t) = \begin{cases} \eta^4 - h(\mathbf{x}, t) & \text{if } \eta > 0 \\ \phantom{\eta^4 -{}} h(\mathbf{x}, t) & \text{if } \eta < 0 \end{cases} \tag{4.3-57}$$

is introduced and we convert the inequality constraint $h(\mathbf{x}, t) \geq 0$ to an equality constraint by writing

$$\frac{\partial \gamma}{\partial \eta}\frac{d\eta}{dt} = \frac{\partial h}{\partial t} + \frac{\partial h}{\partial \mathbf{x}}\frac{d\mathbf{x}}{dt} \tag{4.3-58}$$

which satisfies the constraint if we have the end conditions

$$\gamma[\mathbf{x}(t_o), \eta(t_o), t_o] = \gamma[\mathbf{x}(t_f), \eta(t_f), t_f] = 0 \tag{4.3-59}$$

The Euler-Lagrange equations can, of course, be used to determine the differential equations for an extremum, and the associated transversality conditions can be used to specify the two-point boundary values. If inequality constraints on the control variables are present, we must of necessity incorporate these into our problem formulation. The Hamiltonian formulation may also be used. These methods provide us with necessary conditions only.

From Eq. (4.3-14) it follows that the Lagrangian for the problem at hand is

$$\tilde{\Phi} = \Phi + \lambda_{n+1}[f_{n+1} - \dot{x}_{n+1}]$$

$$\tilde{\Phi} = H - \boldsymbol{\lambda}^T\dot{\mathbf{x}} - \boldsymbol{\Gamma}^T[\mathbf{g} - \dot{\mathbf{z}}^2] + \lambda_{n+1}[f_{n+1} - \dot{x}_{n+1}] \tag{4.3-60}$$

where Φ is the Lagrangian for no inequality state constraint. We are using the equality constraint method of Eqs. (4.3-49) and (4.3-50). The Euler-Lagrange equations yield

$$\frac{d}{dt}\frac{\partial \Phi}{\partial \dot{\mathbf{x}}} - \frac{\partial \Phi}{\partial \mathbf{x}} - \frac{\partial f_{n+1}}{\partial \mathbf{x}}\lambda_{n+1} = 0 \tag{4.3-61}$$

$$\frac{\partial \Phi}{\partial \mathbf{u}} = \frac{d}{dt}\frac{\partial \Phi}{\partial \dot{\mathbf{w}}} = 0 \tag{4.3-62}$$

$$\frac{d}{dt}\frac{\partial \Phi}{\partial \dot{\mathbf{z}}} = 0 \tag{4.3-63}$$

which are, except for the f_{n+1} term, exactly the same as Eqs. (4.3-15), (4.3-16), and (4.3-17). Also, we see that

$$\frac{d}{dt}\lambda_{n+1}(t) = 0 \tag{4.3-64}$$

with the transversality conditions exactly as before and, in addition,

$$x_{n+1}(t_o) = x_{n+1}(t_f) = 0 \tag{4.3-65}$$

It is desirable to reinterpret these results in terms of the Hamiltonian, just as we have done for the case of control variable constraints only. We can do this easily by combining Eq. (4.3-60) with Eq. (4.3-61) and making use of the Weierstrass condition, Eq. (4.3-26), which yields

$$\dot{\boldsymbol{\lambda}} = \frac{d\boldsymbol{\lambda}(t)}{dt} = -\frac{\partial H}{\partial \mathbf{x}} - \frac{\partial f_{n+1}[\mathbf{x}(t), t]}{\partial \mathbf{x}}\lambda_{n+1} \tag{4.3-66}$$

$$\dot{\mathbf{x}} = \frac{d\mathbf{x}(t)}{dt} = \frac{\partial H}{\partial \boldsymbol{\lambda}} \tag{4.3-67}$$

$$\dot{x}_{n+1} = \frac{dx_{n+1}(t)}{dt} = f_{n+1} = [h_1(\mathbf{x}, t)]^2 \mathrm{H}(h_1) + \cdots + [h_s(\mathbf{x}, t)]^2 \mathrm{H}(h_s) \tag{4.3-68}$$

$$\dot{\lambda}_{n+1} = \frac{d\lambda_{n+1}(t)}{dt} = 0 \tag{4.3-69}$$

where

$$H[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t] = \phi[\mathbf{x}(t), \mathbf{u}(t), t] + \boldsymbol{\lambda}^T(t)\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] \quad (4.3\text{-}70)$$

$$H[\mathbf{x}(t), \hat{\mathbf{u}}(t), \boldsymbol{\lambda}(t), t] \leq H[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t]\,|_{\mathbf{u}\in\mathbf{U}} \quad (4.3\text{-}71)$$

with the two-point boundary conditions (transversality conditions)

$$\mathbf{x}(t_o) = \mathbf{x}_o \quad (4.3\text{-}72)$$

$$\mathbf{N}[\mathbf{x}(t_f), t_f] = \mathbf{0} \quad (4.3\text{-}73)$$

$$\left.\begin{aligned} &\frac{\partial\theta}{\partial t_f} + \left(\frac{\partial \mathbf{N}^T}{\partial t_f}\right)\nu + H = 0 \quad (4.3\text{-}74)\\ &\frac{\partial\theta}{\partial \mathbf{x}} + \left(\frac{\partial N^T}{\partial \mathbf{x}}\right)\nu - \boldsymbol{\lambda} = 0 \quad (4.3\text{-}75)\end{aligned}\right\} \quad \text{at} \quad t = t_f$$

$$x_{n+1}(t_o) = x_{n+1}(t_f) = 0 \quad (4.3\text{-}76)$$

These are the equations whose solutions minimize the cost function and constraints of Eqs. (4.3-28) through (4.3-33), subject to the additional constraint $\mathbf{h}[\mathbf{x}(t), t] \geq \mathbf{0}$.

Equations analogous to these could be obtained in a relatively straightforward fashion for each of the other formulations of the inequality state constraint problem presented here. Computational techniques will be used to obtain numerical solutions to problems of this type in later chapters.

Example 4.3-3

As an example of optimization with a state variable constraint, we consider the brachistochrone problem previously treated by McGill [10] and Dreyfus [11]. A particle is falling for a specified time, $t_f - t_o$, under the influence of a constant gravitational acceleration g. The particle has initial velocity $x_3(t_o) = x_{3o}$. We wish to find the path that maximizes the final value of the horizontal coordinate $x_2(t_f)$. The final value of the vertical coordinate $x_2(t_f)$ and the velocity $x_3(t_f)$ are unspecified. The path is constrained by a line $h[x_1, x_2] \geq 0$ in the x_1x_2 plane, where it is known that the unconstrained solution intersects the line. The system dynamics are described by

$$\dot{x}_1 = x_3 \cos u, \qquad x_1(t_o) = x_{1o}$$

$$\dot{x}_2 = x_3 \sin u, \qquad x_2(t_o) = x_{2o}$$

$$\dot{x}_3 = g \sin u, \qquad x_3(t_o) = x_{3o}$$

where the control u is the slope of the path. The cost function is

$$J = -x_1(t_f)$$

with no specified endpoint equality constraints, and the state vector inequality constraint

$$h(x_1x_2) = ax_1 + b - x_2 \geq 0$$

which is converted to the equality constraint

$$\dot{x}_4 = f_4 = [h(x_1, x_2)]^2\mathrm{H}(h)$$

We can easily compute the requisite nonlinear two-po [illegible] boundary value problem by direct application of the maximum principle given in this section. The equations for this TPBVP are

$$\dot{x}_1 = x_3^2\lambda_1[(\lambda_1x_3)^2 + (\lambda_2x_3 + \lambda_3g)^2]^{-1/2}, \qquad x_1(t_o) = x_{1o}$$

$$\dot{x}_2 = x_3(\lambda_2x_3 + \lambda_3g)[(\lambda_1x_3)^2 + (\lambda_2x_3 + \lambda_3g)^2]^{-1/2}, \qquad x_2(t_o) = x_{2o}$$

$$\dot{x}_3 = g(\lambda_2x_3 + \lambda_3g)[(\lambda_1x_3)^2 + (\lambda_2x_3 + \lambda_3g)^2]^{-1/2}, \qquad x_3(t_o) = x_{3o}$$

$$\dot{x}_4 = h(x_1, x_2)\mathrm{H}(h), \qquad x_4(t_o) = 0$$

$$\dot{\lambda}_1 = -2a\lambda_4h(x_1, x_2)\mathrm{H}(h), \qquad \lambda_1(t_o) = -1$$

$$\dot{\lambda}_2 = 2\lambda_4h(x_1, x_2)\mathrm{H}(h), \qquad \lambda_2(t_f) = 0$$

$$\begin{aligned}\dot{\lambda}_3 = &-\lambda_1^2x_3[(\lambda_1x_3)^2 + (\lambda_2x_3 + \lambda_3g)^2]^{-1/2}\\ &- \lambda_2(\lambda_2x_3 + \lambda_3g)[(\lambda_1x_3)^2 + (\lambda_2x_3 + \lambda_3g)^2]^{-1/2}, \qquad \lambda_3(t_f) = 0\end{aligned}$$

$$\dot{\lambda}_4 = 0, \qquad x_4(t_f) = 0$$

The solution of this set of nonlinear differential equations with the associated boundary conditions establishes the optimal trajectory and optimal control. Needless to say, this will not be an easy task. We shall examine this problem again, in Section 13.3-2, and determine a numerical solution for this optimization problem with a state variable inequality constraint.

4.4 Hamilton-Jacobi equation and continuous dynamic programming

Let us consider once more the problem of minimizing

$$J = \int_{t_o}^{t_f} \phi[\mathbf{x}(t), \mathbf{u}(t), t]\, dt \quad (4.4\text{-}1)$$

subject to the equality constraints

$$\dot{\mathbf{x}} = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t], \qquad \mathbf{x}(t_o) = \mathbf{x}_o \quad (4.4\text{-}2)$$

and the control variable inequality constraint

$$\mathbf{u}(t) \in \mathbf{U} \quad (4.4\text{-}3)$$

where $\mathbf{U}$ is a possibly infinite or semi-infinite closed interval, the admissible input set, which may depend on $\mathbf{x}(t)$ and t. Let us further assume, for the moment, that t_f is fixed and $\mathbf{x}(t_f)$ is unspecified. Suppose that we have calculated $\hat{\mathbf{u}}(t)$ and $\hat{\mathbf{x}}(t)$ to be the optimal control and trajectory. The cost function is then a function of the initial state, $\mathbf{x}(t_o)$, and the initial time, t_o, only. It is convenient to give this a special symbol such as

$$V(\mathbf{x}_o, t_o) \triangleq J(\hat{\mathbf{x}}, \hat{\mathbf{u}}) = \int_{t_o}^{t_f} \phi[\hat{\mathbf{x}}(t), \hat{\mathbf{u}}(t), t]\, dt \quad (4.4\text{-}4)$$

so that $V(\mathbf{x}_o, t_o)$ is the minimum value of the performance index when the initial system state is $\mathbf{x}_o$ and the initial time is t_o. $V(\mathbf{x}_o, t_o)$ is a function only of $\mathbf{x}_o$ and t_o since $\hat{\mathbf{x}}(t)$ and $\hat{\mathbf{u}}(t)$ are known (optimal) values for all $t \in [t_o, t_f]$.

We now consider a time Δt between t_o and t_f and rewrite the cost function, Eq. (4.4-4), as

$$V(\mathbf{x}_o, t_o) = \int_{t_o}^{t_o+\Delta t} \phi(\hat{\mathbf{x}}, \hat{\mathbf{u}}, t)\,dt + \int_{t_o+\Delta t}^{t_f} \phi(\hat{\mathbf{x}}, \hat{\mathbf{u}}, t)\,dt \qquad (4.4\text{-}5)$$
$$= J_1(\hat{\mathbf{x}}, \hat{\mathbf{u}}) + J_2(\hat{\mathbf{x}}, \hat{\mathbf{u}})$$

If we now assume that ϕ is smooth over the interval t_o to $t_o + \Delta t$ and that Δt is sufficiently small, we may rewrite the J_1 term as

alue the

$$\rightarrow \quad J_1 = \Delta t \phi[\hat{\mathbf{x}}(t_o + \alpha\Delta t), \hat{\mathbf{u}}(t_o + \alpha\Delta t), t_o + \alpha\Delta t], \qquad 0 < \alpha < 1 \qquad (4.4\text{-}6)$$

ists a

The second part of the cost function is simply

$t_o + \Delta t$

$$V_2 = V[\hat{\mathbf{x}}(t_o + \Delta t), t_o + \Delta t] = \int_{t_o+\Delta t}^{t_f} \phi[\hat{\mathbf{x}}(t), \hat{\mathbf{u}}(t), t]\,dt \qquad (4.4\text{-}7)$$

hat

This is so because of the fundamental theorem of dynamic programming which asserts that any part of an optimal trajectory is an optimal trajectory. To show that J_2 is $V[\hat{\mathbf{x}}(t_o + \Delta t), t_o + \Delta t]$, we observe that the value of J_2 depends only on the state $\hat{\mathbf{x}}(t_o + \Delta t)$ and the control $\hat{\mathbf{u}}(t)$ in the time interval from $t_o + \Delta t$ to t_f. If J_2 was greater than V_2, then there must have existed a control such that

$$J_1(\hat{\mathbf{x}}, \hat{\mathbf{u}}) + \int_{t_o+\Delta t}^{t_f} \phi[\hat{\mathbf{x}}(t), \hat{\mathbf{u}}(t), t] > V(\mathbf{x}_o, t_o) \qquad (4.4\text{-}8)$$

But this contradicts the assumption that $\hat{\mathbf{u}}(t)$ is an optimal control. However, by the definition of V_2, $J_2 \geq V_2$; thus $J_2 = V_2$.

We will now write the cost function along the optimal trajectory as

$$V(\mathbf{x}_o, t_o) = \Delta t \phi[\hat{\mathbf{x}}(t_o + \alpha\Delta t), \hat{\mathbf{u}}(t_o + \alpha\Delta t), t_o + \alpha\Delta t]$$
$$+ V[\hat{\mathbf{x}}(t_o + \Delta t), t_o + \Delta t] \qquad (4.4\text{-}9)$$

By expanding the last term in this equation in a Taylor's series about $\Delta t = 0$, we have

$$V(\mathbf{x}_o, t_o) = \Delta t \phi[\hat{\mathbf{x}}(t_o + \alpha\Delta t), \hat{\mathbf{u}}(t_o + \alpha\Delta t), t_o + \alpha\Delta t]$$
$$+ V(\mathbf{x}_o, t_o) + \left[\frac{\partial V(\mathbf{x}_o, t_o)}{\partial t_o}\right]\Delta t + \left[\frac{\partial V(\mathbf{x}_o, t_o)}{\partial \mathbf{x}_o}\right]^T \dot{\mathbf{x}}_o \Delta t \qquad (4.4\text{-}10)$$
$$+ [\Delta t]^2[\,\cdot\,] + \cdots$$

Upon taking the limit as Δt approaches zero and recalling the equality constraint of Eq. (4.4-2), we have, finally, the Hamilton-Jacobi equation

$$\frac{\partial V(\mathbf{x}_o, t_o)}{\partial t_o} + \phi[\hat{\mathbf{x}}(t_o), \hat{\mathbf{u}}(t_o), t_o] + \left[\frac{\partial V(\mathbf{x}_o, t_o)}{\partial \mathbf{x}_o}\right]^T \mathbf{f}[\hat{\mathbf{x}}(t_o), \hat{\mathbf{u}}(t_o), t_o] = 0 \qquad (4.4\text{-}11)$$

In this expression, we see that if we define

$$\boldsymbol{\lambda}(t_o) = \frac{\partial V(\mathbf{x}_o, t_o)}{\partial \mathbf{x}_o} \qquad (4.4\text{-}12)$$

we may then rewrite the Hamilton-Jacobi equation, or dropping the subscript "o" for convenience, as

$$\frac{\partial V(\mathbf{x}, t)}{\partial t} + H(\mathbf{x}, \hat{\mathbf{u}}, \boldsymbol{\lambda}, t) = 0 \qquad (4.4\text{-}13)$$

It is important for us to stress here that this Hamiltonian is the Hamiltonian evaluated (at time t_o) for the optimum control $\hat{\mathbf{u}}(t)$, since we have been assuming all along that ϕ was evaluated about the optimal control and state. Thus, yet another way for us to write the Hamilton-Jacobi equation is

$$\frac{\partial V(\mathbf{x}, t)}{\partial t} = -H\left(\mathbf{x}, \frac{\partial V}{\partial \mathbf{x}}, t\right) \qquad (4.4\text{-}14)$$

where

$$H\left(\mathbf{x}, \frac{\partial V}{\partial \mathbf{x}}, t\right) = \text{Min}_{\mathbf{u} \in \mathbf{U}}\, H\left[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t) = \frac{\partial V(\mathbf{x}, t)}{\partial \mathbf{x}}, t\right] \qquad (4.4\text{-}15)$$

When t_f is fixed and $\mathbf{x}(t_f)$ is unspecified, it is an easy matter for us to show from Eq. (4.4-4) that the initial condition for the Hamilton-Jacobi equation is

$$V[\mathbf{x}(t_f), t_f] = 0 \qquad (4.4\text{-}16)$$

If we had obtained the Hamilton-Jacobi equation for the cost function

$$J = \theta[\mathbf{x}(t_f), t_f] + \int_{t_o}^{t_f} \phi[\mathbf{x}(t), \mathbf{u}(t), t]\,dt \qquad (4.4\text{-}17)$$

we would have obtained the same Hamilton-Jacobi equation (4.4-13) with the initial condition (at the terminal time)

$$V[\mathbf{x}(t_f), t_f] = \theta[\mathbf{x}(t_f), t_f] \qquad (4.4\text{-}18)$$

Needless to say, the Hamilton-Jacobi equation cannot be easily solved in general. However, when it can, $\mathbf{u}(t)$ is determined as a function of $\mathbf{x}(t)$, or in other words, we find a feedback control law which is highly desirable. The Hamilton-Jacobi partial differential equation is equivalent to the functional equation of dynamic programming or Bellman's equation [11,12,13]. It is sometimes called the Hamilton-Jacobi-Bellman equation [14].

Example 4.4-1

Let us consider the linear constant differential system described by

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t), \qquad \mathbf{x}(0) = \mathbf{x}_o$$

where $\mathbf{A}$ is an $n \times n$ matrix and $\mathbf{b}$ is an n vector. Any $u(t)$ is assumed to be admissible. We wish to find $u(t)$ as a function of $\mathbf{x}(t)$ such that

$$J = \tfrac{1}{2}\int_0^\infty [\mathbf{x}^T\mathbf{Q}\mathbf{x} + ru^2]\,dt$$

We now consider a time Δt between t_o and t_f and rewrite the cost function, Eq. (4.4-4), as

$$\begin{aligned} V(\mathbf{x}_o, t_o) &= \int_{t_o}^{t_o+\Delta t} \phi(\hat{\mathbf{x}}, \hat{\mathbf{u}}, t)\,dt + \int_{t_o+\Delta t}^{t_f} \phi(\hat{\mathbf{x}}, \hat{\mathbf{u}}, t)\,dt \\ &= J_1(\hat{\mathbf{x}}, \hat{\mathbf{u}}) + J_2(\hat{\mathbf{x}}, \hat{\mathbf{u}}) \end{aligned} \tag{4.4-5}$$

If we now assume that ϕ is smooth over the interval t_o to $t_o + \Delta t$ and that Δt is sufficiently small, we may rewrite the J_1 term as

$$J_1 = \Delta t \phi[\hat{\mathbf{x}}(t_o + \alpha\Delta t), \hat{\mathbf{u}}(t_o + \alpha\Delta t), t_o + \alpha\Delta t], \qquad 0 < \alpha < 1 \tag{4.4-6}$$

The second part of the cost function is simply

$$V_2 = V[\hat{\mathbf{x}}(t_o + \Delta t), t_o + \Delta t] = \int_{t_o+\Delta t}^{t_f} \phi[\hat{\mathbf{x}}(t), \hat{\mathbf{u}}(t), t]\,dt \tag{4.4-7}$$

This is so because of the fundamental theorem of dynamic programming which asserts that any part of an optimal trajectory is an optimal trajectory. To show that J_2 is $V[\hat{\mathbf{x}}(t_o + \Delta t), t_o + \Delta t]$, we observe that the value of J_2 depends only on the state $\hat{\mathbf{x}}(t_o + \Delta t)$ and the control $\hat{\mathbf{u}}(t)$ in the time interval from $t_o + \Delta t$ to t_f. If J_2 was greater than V_2, then there must have existed a control such that

$$J_1(\hat{\mathbf{x}}, \hat{\mathbf{u}}) + \int_{t_o+\Delta t}^{t_f} \phi[\hat{\mathbf{x}}(t), \hat{\mathbf{u}}(t), t] > V(\mathbf{x}_o, t_o) \tag{4.4-8}$$

But this contradicts the assumption that $\hat{\mathbf{u}}(t)$ is an optimal control. However, by the definition of V_2, $J_2 \geq V_2$; thus $J_2 = V_2$.

We will now write the cost function along the optimal trajectory as

$$\begin{aligned} V(\mathbf{x}_o, t_o) = \Delta t \phi[\hat{\mathbf{x}}(t_o + \alpha\Delta t), \hat{\mathbf{u}}(t_o + \alpha\Delta t), t_o + \alpha\Delta t] \\ + V[\hat{\mathbf{x}}(t_o + \Delta t), t_o + \Delta t] \end{aligned} \tag{4.4-9}$$

By expanding the last term in this equation in a Taylor's series about $\Delta t = 0$, we have

$$\begin{aligned} V(\mathbf{x}_o, t_o) = \Delta t \phi[&\hat{\mathbf{x}}(t_o + \alpha\Delta t), \hat{\mathbf{u}}(t_o + \alpha\Delta t), t_o + \alpha\Delta t] \\ &+ V(\mathbf{x}_o, t_o) + \left[\frac{\partial V(\mathbf{x}_o, t_o)}{\partial t_o}\right]\Delta t + \left[\frac{\partial V(\mathbf{x}_o, t_o)}{\partial \mathbf{x}_o}\right]^T \dot{\mathbf{x}}_o \Delta t \\ &+ [\Delta t]^2[\,\cdot\,] + \cdots \end{aligned} \tag{4.4-10}$$

Upon taking the limit as Δt approaches zero and recalling the equality constraint of Eq. (4.4-2), we have, finally, the Hamilton-Jacobi equation

$$\frac{\partial V(\mathbf{x}_o, t_o)}{\partial t_o} + \phi[\hat{\mathbf{x}}(t_o), \hat{\mathbf{u}}(t_o), t_o] + \left[\frac{\partial V(\mathbf{x}_o, t_o)}{\partial \mathbf{x}_o}\right]^T \mathbf{f}[\hat{\mathbf{x}}(t_o), \hat{\mathbf{u}}(t_o), t_o] = 0 \tag{4.4-11}$$

In this expression, we see that if we define

$$\boldsymbol{\lambda}(t_o) = \frac{\partial V(\mathbf{x}_o, t_o)}{\partial \mathbf{x}_o} \tag{4.4-12}$$

we may then rewrite the Hamilton-Jacobi equation, [illegible] dropping the subscript "*o*" for convenience, as

$$\frac{\partial V(\mathbf{x}, t)}{\partial t} + H(\mathbf{x}, \hat{\mathbf{u}}, \boldsymbol{\lambda}, t) = 0 \tag{4.4-13}$$

It is important for us to stress here that this Hamiltonian is the Hamiltonian evaluated (at time t_o) for the optimum control $\hat{\mathbf{u}}(t)$, since we have been assuming all along that ϕ was evaluated about the optimal control and state. Thus, yet another way for us to write the Hamilton-Jacobi equation is

$$\frac{\partial V(\mathbf{x}, t)}{\partial t} = -H\left(\mathbf{x}, \frac{\partial V}{\partial \mathbf{x}}, t\right) \tag{4.4-14}$$

where

$$H\left(\mathbf{x}, \frac{\partial V}{\partial \mathbf{x}}, t\right) = \mathrm{Min}_{\mathbf{u} \in \mathbf{U}}\, H\left[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t) = \frac{\partial V(\mathbf{x}, t)}{\partial \mathbf{x}}, t\right] \tag{4.4-15}$$

When t_f is fixed and $\mathbf{x}(t_f)$ is unspecified, it is an easy matter for us to show from Eq. (4.4-4) that the initial condition for the Hamilton-Jacobi equation is

$$V[\mathbf{x}(t_f), t_f] = 0 \tag{4.4-16}$$

If we had obtained the Hamilton-Jacobi equation for the cost function

$$J = \theta[\mathbf{x}(t_f), t_f] + \int_{t_o}^{t_f} \phi[\mathbf{x}(t), \mathbf{u}(t), t]\,dt \tag{4.4-17}$$

we would have obtained the same Hamilton-Jacobi equation (4.4-13) with the initial condition (at the terminal time)

$$V[\mathbf{x}(t_f), t_f] = \theta[\mathbf{x}(t_f), t_f] \tag{4.4-18}$$

Needless to say, the Hamilton-Jacobi equation cannot be easily solved in general. However, when it can, $\mathbf{u}(t)$ is determined as a function of $\mathbf{x}(t)$, or in other words, we find a feedback control law which is highly desirable. The Hamilton-Jacobi partial differential equation is equivalent to the functional equation of dynamic programming or Bellman's equation [11,12,13]. It is sometimes called the Hamilton-Jacobi-Bellman equation [14].

Example 4.4-1

Let us consider the linear constant differential system described by

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t), \qquad \mathbf{x}(0) = \mathbf{x}_o$$

where $\mathbf{A}$ is an $n \times n$ matrix and $\mathbf{b}$ is an n vector. Any $u(t)$ is assumed to be admissible. We wish to find $u(t)$ as a function of $\mathbf{x}(t)$ such that

$$J = \tfrac{1}{2}\int_0^\infty [\mathbf{x}^T\mathbf{Q}\mathbf{x} + ru^2]\,dt$$

is a minimur) is a positive constant semidefinite matrix, and r is positive. The Hamiltonian for the problem is

$$H(\mathbf{x}, u, \boldsymbol{\lambda}, t) = \tfrac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \tfrac{1}{2}ru^2 + \boldsymbol{\lambda}^T\mathbf{A}\mathbf{x} + \boldsymbol{\lambda}^T\mathbf{b}u$$

We need to find the control u which minimizes the Hamiltonian. This is

$$\frac{\partial H}{\partial u} = 0 = ru + \mathbf{b}^T\boldsymbol{\lambda}$$

so

$$u = -\mathbf{b}^T\boldsymbol{\lambda}r^{-1}$$

and the Hamiltonian becomes

$$H(\mathbf{x}, \boldsymbol{\lambda}, t) = \tfrac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \boldsymbol{\lambda}^T\mathbf{A}\mathbf{x} - \tfrac{1}{2}\boldsymbol{\lambda}^T\mathbf{b}\mathbf{b}^T\boldsymbol{\lambda}r^{-1}$$

Since the system and the $\mathbf{Q}$ and r terms are time invariant and since the optimization is for a process of infinite duration, it follows that $V(\mathbf{x}, t)$ will depend only upon the initial state $\mathbf{x}$. This implies that

$$\frac{\partial V(\mathbf{x}, t)}{\partial t} = 0$$

Therefore, since $\boldsymbol{\lambda} = \partial V/\partial \mathbf{x}$, the Hamilton-Jacobi equation becomes

$$\frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \left(\frac{\partial V}{\partial \mathbf{x}}\right)^T\mathbf{A}\mathbf{x} - \frac{1}{2}\left[\left(\frac{\partial V}{\partial \mathbf{x}}\right)^T\mathbf{b}\right]^2 r^{-1} = 0$$

If we assume a solution

$$V(\mathbf{x}, t) = \tfrac{1}{2}\mathbf{x}^T\mathbf{P}\mathbf{x}$$

we see that

$$\frac{\partial V}{\partial \mathbf{x}} = \mathbf{P}\mathbf{x}$$

and the Hamilton-Jacobi equation may be written as

$$\mathbf{x}^T[\tfrac{1}{2}\mathbf{Q} + \tfrac{1}{2}\mathbf{P}\mathbf{A} + \tfrac{1}{2}\mathbf{A}^T\mathbf{P} - \tfrac{1}{2}\mathbf{P}\mathbf{b}\mathbf{b}^T\mathbf{P}r^{-1}]\,\mathbf{x} = 0$$

which says that, for any nonzero $\mathbf{x}(t)$, the matrix $\mathbf{P}$ must satisfy the $n(n + 1)/2$ algebraic equations (the $\mathbf{P}$ matrix is symmetric)

$$\mathbf{Q} + \mathbf{P}\mathbf{A} + \mathbf{A}^T\mathbf{P} - \mathbf{P}\mathbf{b}\mathbf{b}^T\mathbf{P}r^{-1} = 0$$

This equation is solved for $\mathbf{P}$, and then the control is computed from

$$u = -\mathbf{b}^T\boldsymbol{\lambda}r^{-1} = -\mathbf{b}^Tr^{-1}\left(\frac{\partial V}{\partial \mathbf{x}}\right) = -\mathbf{b}^T\mathbf{P}\mathbf{x}r^{-1}$$

If we further consider the system

$$\dot{x}_1 = x_2, \qquad x_1(0) = x_{10}$$
$$\dot{x}_2 = u, \qquad x_2(0) = x_{20},$$

and the cost function

$$J = \tfrac{1}{2}\int_0^\infty (4x_1^2 + u^2)\,dt$$

it is easy for us to show that the optimum control is given by

$$u = -2x_1 - 2x_2$$

Example 4.4-2

Consider the system

$$\dot{x} = -x^3 + u, \qquad x(0) = x_0$$

with cost function

$$J = \tfrac{1}{2}\int_0^{t_f} (x^2 + u^2)\,dt$$

where it is desired to determine the optimal feedback control. We accomplish this by forming the Hamiltonian

$$H(x, u, \lambda, t) = \tfrac{1}{2}x^2 + \tfrac{1}{2}u^2 + \lambda u - \lambda x^3$$

We then set $\partial H/\partial u = 0$ and note that $\lambda = \partial V/\partial x$ to obtain $u = -\lambda$; then

$$H\left(x, \frac{\partial V}{\partial x}\right) = \frac{1}{2}x^2 - \frac{1}{2}\left[\frac{\partial V(x, t)}{\partial x}\right]^2 - \left[\frac{\partial V(x, t)}{\partial x}\right]x^3$$

The Hamilton-Jacobi equation is

$$\frac{\partial V(x, t)}{\partial t} - \frac{1}{2}\left[\frac{\partial V(x, t)}{\partial x}\right]^2 - \left[\frac{\partial V(x, t)}{\partial x}\right]x^3 + \tfrac{1}{2}x^2 = 0$$

with $V[x(t_f), t_f] = 0$.

If the optimization interval is infinite, then $\partial V/\partial t = 0$, and we need to solve the differential equation

$$\left[\frac{dV(x)}{dx}\right]^2 + 2\left[\frac{dV(x)}{dx}\right]x^3 - x^2 = 0$$

with $V(0) = 0$ as the initial condition. We may approximate the solution to this ordinary differential equation by a series expansion

$$V(x) = p_0 + p_1x + \frac{1}{2}p_2x^2 + \frac{1}{3!}p_3x^3 + \frac{1}{4!}p_4x^4 + \cdots$$

If we terminate the series after the fourth-order term, substitute the assumed solution into the differential equation, and equate like powers of x (up to x^4), we obtain $p_0 = p_1 = p_3 = 0, p_2 = 1, p_4 = -6$. Thus the approximate closed-loop control is

$$u = -\lambda = -\frac{dV}{dx} = -x + x^3$$

We naturally may question the stability of the approximate control. However, with u as obtained, the system differential equation becomes

$$\dot{x} = -x^3 + u = -x$$

which is certainly stable.

A similar procedure to this could have been used to obtain an approximate solution to the nonlinear partial differential equation that is the Hamilton-Jacobi equation for this example. In this case, the p's would be functions of time, and we would obtain matrix Riccati-type equations [15]. This approach has many attractive features. In particular, only initial condition problems need be solved. However, there are two disadvantages: There is no assurance of system stability; the number of matrix Riccati differential equations which must be solved increases gre with the order of the differential system and the order of the polynomial i. or the approximate solution to $V(\mathbf{x}, t)$. If an expansion in x to the N order is used for an n vector differential system, the number of distinct Riccati-type differential equations is

$$E = \sum_{j=1}^{N} \frac{(n-1+j)!}{(n-1)!\,j!}$$

for an assumed solution of the form

$$V(\mathbf{x}, t) = \sum_{j=1}^{n} p_j x_j + \tfrac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} p_{jk} x_j x_k + \tfrac{1}{6} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{l=1}^{n} p_{jkl} x_j x_k x_l + \cdots$$

If, for example, the solution to a four-vector differential system is approximated by terms up to and including the fourth power in x, we need to solve sixty-nine differential equations to obtain the closed-loop control.

Our discussion of the second variation technique, the invariant imbedding procedure, and specific optimal control using the quasilinearization approach will point out many interesting interconnections with the approach used to obtain the solution to this example.

In our development thus far, we have assumed that the terminal time, t_f, is fixed. It is possible to remove this restriction with the result that the Hamilton-Jacobi equation (4.4-13), (4.4-14), or (4.4-15) is still applicable. The initial condition for the Hamilton-Jacobi equation is still Eq. (4.4-18) and, in addition, the terminal time is determined by the relation

$$H\left(\mathbf{x}, \frac{\partial V}{\partial \mathbf{x}}, t\right) + 1 = 0, \qquad \text{at} \quad t = t_f \tag{4.4-19}$$

which holds if the problem is a minimum time problem such that

$$V(\mathbf{x}, t) = t_f - t \tag{4.4-20}$$

If, further, the differential system is time invariant, the Hamiltonian is equal to -1 at all times along the optimal trajectory.

We may formally obtain the Pontryagin maximum principle by taking appropriate partial derivatives of the Hamilton-Jacobi equations (Problem 9). However, the resulting maximum principle is not applicable to as broad a class of problems as is possible. The reason for this is that it is necessary that $V(\mathbf{x}, t)$ be smooth or, in other words, twice continuously differentiable with respect to x in order to obtain the Hamilton canonic equations of the maximum principle. We shall illustrate these difficulties with a simple example.

Example 4.4-3

A second-order example will now be discussed to illustrate that the assumption of the differentiability of $V(\mathbf{x}, t)$ does not hold in some of the simplest cases. We will consider the problem of transferring the system represented by the differential equations

$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = u$$

from an initial state $\mathbf{x}_0$ to the origin in minimum time. W sume that the admissible set for the scalar control is described by $|u(t)| \leq 1$.

This problem can be solved by the Pontryagin maximum principle. In the time optimal problem

$$J = \int_{t_0}^{t_f} (1)\, dt$$

Therefore, the Hamiltonian is

$$H[\mathbf{x}, u, \lambda, t] = 1 + \lambda_1 x_2 + \lambda_2 u$$

The adjoint equations are

$$\dot{\lambda}_1 = 0, \qquad \dot{\lambda}_2 = -\lambda_1$$

The solutions to these equations are

$$\lambda_1 = C_1, \qquad \lambda_2 = C_2 - C_1 t$$

where C_j is the initial condition on λ_j. The control which minimizes the Hamiltonian subject to $|u| \leq 1$ is

$$u = -\text{sign}\,\lambda_2 = -\text{sign}\,(C_2 - C_1 t)$$

The initial conditions C_1 and C_2 are not arbitrary but must be such that $\mathbf{x}(t_f) = 0$ since it is desired to transfer the system $\mathbf{x}_0$ to the origin in minimum time. When $u = +1$, the solution to the differential system equation is

$$x_2 = t + x_2(0)$$

$$x_1 = \frac{t^2}{2} + x_2(0)t + x_1(0)$$

If t is eliminated from the foregoing, we obtain

$$x_1 = \frac{x_2^2}{2} + x_1(0) - \frac{x_2^2(0)}{2}$$

When $u = -1$, the solution to the differential system equations is

$$x_2 = -t + x_2'(0)$$

$$x_1 = \frac{-t^2}{2} + x_2'(0)t + x_1'(0)$$

and if t is eliminated in the foregoing, we obtain

$$x_1 = \frac{-x_2^2}{2} + x_1'(0) + \frac{x_2'^2(0)}{2}$$

By determining the constants C_1 and C_2 in terms of x_1 and x_2, it is a straightforward task for us to show that the control law is

$$u = -\text{sign}[x_1(t) + \tfrac{1}{2}x_2(t)\,|x_2(t)|]$$

These equations represent the optimal control and trajectories for $u = -1$ and $u = +1$, respectively, and they indicate that these trajectories are segments of parabolas. Figure 4.4-1 is a plot of some of these parabolas.

The segment of the parabola which is not an optimal trajectory has been represented by a broken line. The optimal control can be determined from Fig. 4.4-1 and a knowledge of the state of the system. The curve AOB represents the switching curve. When x lies below AOB, $u = +1$ until the system state reaches the curve AO, at which time the control switches to -1. If x lies above AOB, $u = -1$ until it reaches BO, where it switches to $+1$.

The optimal transition time $T(\mathbf{x})$, which is the cost function J or $V(\mathbf{x}, t)$, can be determined from the solutions for x_1 and x_2. Figure 4.4-2 is a plot of $T(\mathbf{x})$, the minimum time to transfer to the origin for the case in which the initial x_2 is held constant ($x_{2_0} = -2$), and x_{1_0} is varied about the switching line.

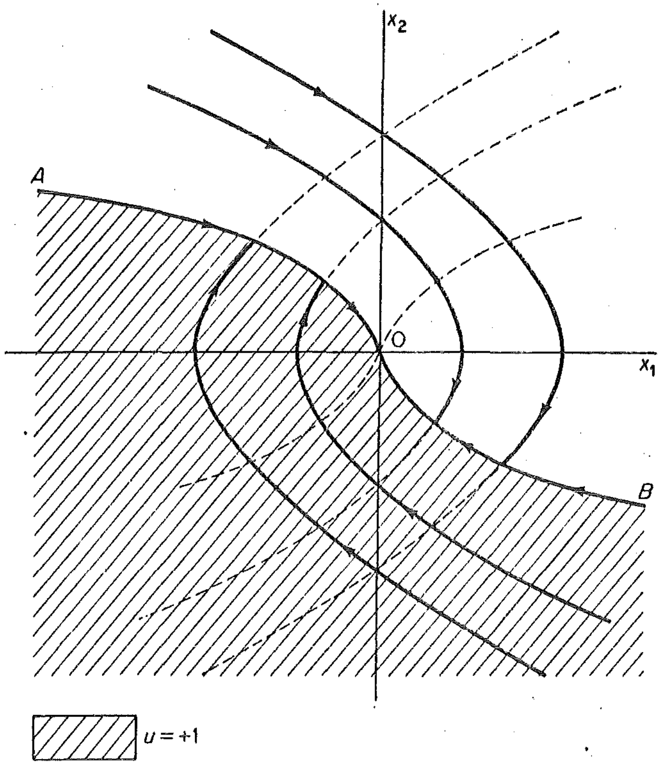


Fig. 4.4-1 Switching curve and trajectories for minimum time Example (4.4-3).

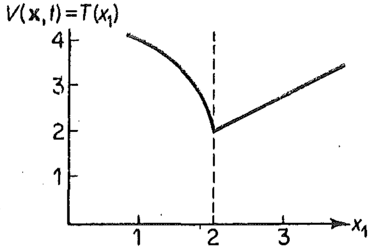


Fig. 4.4-2 Minimum time to origin for fixed x_{2_0} Example (4.4-3).

From the graph it can be seen that $\partial T(\mathbf{x})/\partial x_1$ has a discontinuity at the switching curve. It can be shown analytically that $\partial T(\mathbf{x})/\partial x_1$ "blows up" as x_1 approaches $+2$ from the left. Hence the Hamilton-Jacobi equation would not be applicable in examples of this type. This example indicates the loss of generality which results from deriving the maximum principle from the Hamilton-Jacobi-Bellman equations.

REFERENCES

1. Gelfand, I.M. and Fomin, S.V., *Calculus of Variations.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.
2. Pontryagin, L.S., et al., *The Mathematical Theory of Optimal Processes.* Wiley, New York. 1962.
3. Rozonoér, L.I., "L.S. Pontryagin's Maximum Principle in the Theory of Optimum Systems I." *Automatica i Telemekhanika*, Vol. 20, No. 10, pp. 1320–1334, October, 1959.
4. Rozonoér, L.I., "L.S. Pontryagin's Maximum Principle in the Theory of Optimum Systems II." *Automatica i Telemekhanika*, Vol. 20, No. 11, pp. 1441–1458, November, 1959.
5. Rozonoér, L.I., "L.S. Pontryagin's Maximum Principle in the Theory of Optimum Systems III." *Automatica i Telemekhanika*, Vol. 20, No. 12, pp. 1561–1578, December, 1959.
6. Berkovitz, L.D., "Variational Methods in Problems of Control and Programming." *J. Math. Anal. Appl.*, Vol. 3, pp. 145–169, 1961.
7. Berkovitz, L.D., "On Control Problems with Bounded State Variables." *J. Math. Anal. Appl.*, Vol. 5, pp. 488–498, 1962.
8. Valentine, F.A., "The Problem of Lagrange with Differential Inequalities as Added Side Conditions." *Contributions to the Calculus of Variations 1933–1937*, University of Chicago Press, Chicago, Illinois, 1937.
9. Kelley, H.J., "Methods of Gradients." *Optimization Techniques*, G. Leitman, Ed., Ch. 6, Academic Press, New York, 1962.
10. McGill, R., "Optimal Control, Inequality State Constraints, and the Generalized Newton-Raphson Algorithm." *J. SIAM on Control*, Vol. 3, pp. 291–298, 1965.
11. Dreyfus, S.E., *Dynamic Programming and the Calculus of Variations.* Academic Press, New York, 1965.
12. Kalman, R.E., "Contributions to the Theory of Optimal Control." *Bol. Soc. Mat. Mex.*, Vol. 5, pp. 102–119, 1960.
13. *Mathematical Optimization Techniques*, R. Bellman, Ed. University of California Press, Berkeley, Calif., 1963.
14. Bryson, A.E., Jr., "Optimal Programming and Control." *Proceedings IBM Scientific Computing Symposium on Control Theory and Applications*, IBM Publication 320-1939, 1966.
15. Merriam, C.W. III, *Optimization Theory and the Design of Feedback Control Systems.* McGraw-Hill Book Company, New York, 1964.

PROBLEMS

1. Find the TPBVP which, when solved, yields the control, $u(t)$, and trajectory, $x(t)$, which minimize

$$J = \tfrac{1}{2}\int_0^1 (x^2 + u^2)\,dt$$

for the system

✓

$$\dot{x} = -x^3 + u, \qquad x(0) = 1$$

2. Find the control and trajectory which transfers the system

$$\dot{x}_1 = x_2, \qquad x_1(0) = 0$$
$$\dot{x}_2 = u, \qquad x_2(0) = 0$$

to the line

$$x_1(1) + x_2(1) = 1$$

such that

$$J = \tfrac{1}{2}\int_0^1 u^2(t)\,dt$$

is minimized.

3. Find the control and trajectory which transfers the system

$$\dot{x} = -x + u$$

from $x(0) = 10$ to $x(1) = 0$ such that

$$J = \tfrac{1}{2}\int_0^1 (\ddot{u})^2\,dt$$

✓ is minimized.

4. Find the control and trajectory which minimizes

$$J = \tfrac{1}{2}\int_0^4 x^2(t)\,dt$$

subject to the inequality constraint $|u(t)| \leq 1$ for the system $\dot{x} = u$ such that $x(0) = 1$, $x(4) = 1$.

5. Determine the Weierstrass-Erdmann corner conditions for the minimization of the cost function

$$J = \int_0^1 x^2(2 - \dot{x})^2\,dt$$

6. What is the Weierstrass E function for the cost function of Problem 5?

7. For the system

$$\dot{x}_1 = x_2, \qquad x_1(0) = 10$$
$$\dot{x}_2 = u, \qquad x_2(0) = 0$$

find the control and trajectory which minimizes

$$J = t_f^2 + \tfrac{1}{2}\int_0^{t_f} u^2\,dt$$

if the desired final state is:
(a) $x_1(t_f) = x_2(t_f) = 0$.
(b) $x_1(t_f) = 0, \qquad x_2(t_f) =$ unspecified.

$= \frac{1}{2}x_1^2 + \frac{1}{2}(\frac{\partial V}{\partial x_1}$

8. ✓ Develop a second- and fourth-order approximation to th lution of the Hamilton-Jacobi equation to find the closed-loop control which minimizes

$$J = \tfrac{1}{2}\int_{t_0}^{t_f} (x_1^2 + u^2)\,dt \qquad \frac{\partial V(x,t_0,t_f)}{\partial t} + \frac{1}{2}x_1^2 - \frac{1}{2}\left(\frac{\partial V}{\partial x_2}\right)^2$$

for the system

$$\dot{x}_1 = x_2 + x_2^3 \qquad + \frac{\partial V}{\partial x_2}(x_1 - x_2) + \frac{\partial V}{\partial x_1}(x_2 + x$$
$$\dot{x}_2 = x_1 - x_2 + u \qquad \text{Since } t_f \to \infty \Rightarrow \frac{\partial V}{\partial t} =$$

Compute and compare the actual numerical results when t_f is infinite.

9. Derive the Pontryagin maximum principle from the Hamilton-Jacobi equation by calculating $(d/dt)(\partial V/\partial \mathrm{x})$ and $\partial V/\partial \lambda$ as outlined in Section 4.4. Observe the differentiability requirement on $V(\mathrm{x}, t)$.

Now assume a 2nd order

$$V(x) = P_0 + \sum_{i=1}^{2} P_i x$$
$$+ \sum_{i,j=1}^{2}\sum P_{ij} x_i x_j$$

10. ✓ Find the control vector which minimizes

$$J = \tfrac{1}{2}\int_0^1 (x^2 + u_1^2 + u_2^2)\,dt$$

for the system described by

$$\dot{x} = u_1 + u_2, \qquad x(0) = 1 \qquad \text{plug in and find } u$$

Use the maximum principle and the Hamilton-Jacobi equations to find the optimum control vector.

coefficients.

11. Set up the differential equations and boundary conditions to minimize for t_f unspecified

$$J = \int_0^{t_f} u^2 dt + t_f x_2(t_f)$$

subject to the constraints
a) $\dot{x}_1 = x_2$, $\dot{x}_2 = x_3$, $\dot{x}_3 = u$
b) $\mathrm{x}(0) = \mathbf{0}$
c) $|u| \leq 1; |x_3| \leq 10$
✓ d) $x_1(t_f) = t_f^2$, $x_2(t_f) = x_3^2(t_f)$

12. Set up the equations and boundary conditions to optimize the system

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = x_3, \quad \dot{x}_3 = u$$

for the performance index with t_f unspecified

$$J = \int_0^{t_f} x_1^2 dt + t_f^2 x_2(t_f)$$

subject to all of the following constraints
a) $\mathrm{x}^T(0) = [1, 0, 0]$
b) $x_1(t_f) = x_2(t_f)$
c) $x_3(t_f) = 0$
d) $|u| \leq 1$
e) $\int_0^{t_f} u^2 dt = 1$

13. Find the Hamilton-Jacobi equation for the system

$$\dot{x}_1 = x_2 \qquad \dot{x}_2 = -x_2 - x_1^2 + u$$

if the performance index is

$$J = \int_0^{t_f} (x_1^2 + u^2)\, dt$$

14. Show that the solution of the Hamilton-Jacobi equation for the system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{u}, \quad \mathbf{A}^T + \mathbf{A} = \mathbf{0}, \quad \|\mathbf{u}\| \leq 1$$

and the cost function

$$J = \int_0^{t_f} dt = t_f$$

is

$$V(\mathbf{x}) = \|\mathbf{x}\|$$

What is the optimal control?

15. Find the optimal control to minimize

$$J = \int_0^{t_f} dt$$

for the system

$$\dot{x} = -x + u,$$

when

$$x(0) = 1, \quad x(t_f) = 0$$

$$|u| \leq 1 + |x|$$

5

OPTIMUM SYSTEMS CONTROL EXAMPLES

In this chapter, we will illustrate some, but certainly by no means all, or even a majority, of the optimal control problems for which closed-form analytic solutions have been obtained. The problems we will solve in this chapter are very important in their own right and illustrate the use of the maximum principle for problems in which closed-form analytic solutions may be obtained. Specifically, we will discuss the linear regulator problem, the first solution of which was due to Kalman [1, 2, 3, 4]. We then discuss the minimum time problem which has been considered by Pontryagin [5], Bellman [6], LaSalle [7], and many others [8 through 13].

A characteristic of some minimum time problems is the possibility of a singular solution. The possibility of singular solutions is well-recognized in the variational calculus literature and has been extensively discussed for control problems by Johnson [14, 15, 16] and others. Minimum fuel problems for linear differential systems are then discussed. A variety of authors, but notably Athans, have discussed various aspects of minimum fuel problems including the possibility of singular solutions [17 through 20]. Finally, the minimum time, minimum fuel, and minimum energy control of self-adjoint systems are discussed. It is certainly true that the self-adjoint assumption, coupled with the need for as many control inputs as state variables, seriously restricts the practical usefulness of the solutions, particularly for high-order systems. However, the relative ease with which the control can be computed makes this an excellent example for a relatively thorough analysis.

Many other optimal control problems are solved in this book other than the ones in this chapter. Discrete and distributed parameter problems are reserved for the next two chapters. Chapter 11 discusses several optimal control problems with regard to observability and controllability. Nonlinear problems, which include the majority of optimal control problems, are discussed in Chapters 13, 14, and 15. The literature in this area is very extensive. For an excellent survey of many other problems plus a lengthy bibliography, we refer to the survey papers of Paiewonsky [22] and Athans [23].

5.1 The linear regulator

We will now study a particular control problem which has as its solution a linear feedback control law. It occurs where we have a linear differential system

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{u}, \qquad \mathbf{x}(t_o) = \mathbf{x}_o \tag{5.1-1}$$

and wish to find the control which minimizes the cost function (for t_f fixed)

$$J = \tfrac{1}{2}\mathbf{x}^T(t_f)\mathbf{S}\mathbf{x}(t_f) + \tfrac{1}{2}\int_{t_o}^{t_f} [\mathbf{x}^T(t)\mathbf{Q}(t)\mathbf{x}(t) + \mathbf{u}^T(t)\mathbf{R}(t)\mathbf{u}(t)]\,dt \tag{5.1-2}$$

Clearly, there is no loss of generality in assuming $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{S}$ to be symmetric. We may obtain the solution to this problem via the maximum principle or the Hamilton-Jacobi equation. Here, we will use the former method. The Hamiltonian is

$$H[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t] = \tfrac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \tfrac{1}{2}\mathbf{u}^T\mathbf{R}\mathbf{u} + \boldsymbol{\lambda}^T\mathbf{A}\mathbf{x} + \boldsymbol{\lambda}^T\mathbf{B}\mathbf{u} \tag{5.1-3}$$

Application of the maximum principle requires that, for an optimum control,

$$\frac{\partial H}{\partial \mathbf{u}} = \mathbf{0} = \mathbf{R}(t)\mathbf{u}(t) + \mathbf{B}^T(t)\boldsymbol{\lambda}(t) \tag{5.1-4}$$

and

$$\frac{\partial H}{\partial \mathbf{x}} = -\dot{\boldsymbol{\lambda}} = \mathbf{Q}(t)\mathbf{x}(t) + \mathbf{A}^T(t)\boldsymbol{\lambda}(t) \tag{5.1-5}$$

with the terminal condition

$$\boldsymbol{\lambda}(t_f) = \frac{\partial \theta}{\partial \mathbf{x}(t_f)} = \mathbf{S}\mathbf{x}(t_f) \tag{5.1-6}$$

Thus we require that

$$\mathbf{u}(t) = -\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\boldsymbol{\lambda}(t) \tag{5.1-7}$$

and we shall inquire whether we may convert this to a closed-loop control by assuming that the solution for the adjoint is similar to Eq. (5.1-6)

$$\boldsymbol{\lambda}(t) = \mathbf{P}(t)\mathbf{x}(t) \tag{5.1-8}$$

If we substitute this relation into Eqs. (5.1-1) and (5.1-7), we see that we must require

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}(t) - \mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{P}(t)\mathbf{x}(t) \tag{5.1-9}$$

Also, from Eqs. (5.1-8) and (5.1-5) we require

$$\dot{\boldsymbol{\lambda}} = \dot{\mathbf{P}}\mathbf{x}(t) + \mathbf{P}(t)\dot{\mathbf{x}} = -\mathbf{Q}(t)\mathbf{x}(t) - \mathbf{A}^T(t)\mathbf{P}(t)\mathbf{x}(t) \tag{5.1-10}$$

By combining Eqs. (5.1-9) and (5.1-10) we have

$$[\dot{\mathbf{P}} + \mathbf{P}(t)\mathbf{A}(t) + \mathbf{A}^T(t)\mathbf{P}(t) - \mathbf{P}(t)\mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{P}(t) + \mathbf{Q}(t)]\mathbf{x}(t) = \mathbf{0} \tag{5.1-11}$$

Since this must hold for all nonzero $\mathbf{x}(t)$, the term premultiplying $\mathbf{x}(t)$ must be zero. Thus the $\mathbf{P}$ matrix, which we see is an $n \times n$ symmetric matrix and which has $n(n + 1)/2$ different terms, must satisfy the matrix Riccati equation — which, as we shall see later, must be positive definite —

$$\dot{\mathbf{P}} = -\mathbf{P}(t)\mathbf{A}(t) - \mathbf{A}^T(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{P}(t) - \mathbf{Q}(t) \tag{5.1-12}$$

with a terminal condition given by Eqs. (5.1-6) and (5.1-8)

$$\mathbf{P}(t_f) = \mathbf{S} \tag{5.1-13}$$

Thus we may solve the matrix Riccati equation backward in time from t_f to t_o, store the matrix

$$\mathbf{K}(t) = -\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{P}(t) \tag{5.1-14}$$

and then obtain a closed-loop control from

$$\mathbf{u}(t) = +\mathbf{K}(t)\mathbf{x}(t) \tag{5.1-15}$$

It is important to note that all components of the state vector must be accessible. We will remove this restriction in Chapter 11 when we discuss the ideal observer. A block diagram for accomplishing this solution to the regulator problem is shown in Fig. 5.1-1. If we compute the second variation, we find that

$$\delta^2 J = \tfrac{1}{2}\,\delta\mathbf{x}^T(t_f)\mathbf{S}\,\delta\mathbf{x}(t_f) + \tfrac{1}{2}\int_{t_o}^{t_f} [\delta\mathbf{x}^T(t)\mathbf{Q}(t)\,\delta\mathbf{x}(t) + \delta\mathbf{u}^T(t)\mathbf{R}(t)\,\delta\mathbf{u}(t)]\,dt \tag{5.1-16}$$

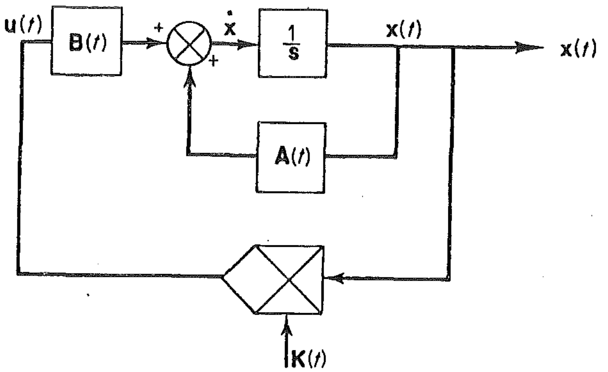


Fig. 5.1-1 Optimum linear closed-loop regulator.

Thus, **Q** and **S** must be at least positive semidefinite in order to establish the sufficient condition for a minimum. In addition, we know from Eq. (5.1-7) that **R** must have an inverse; therefore, it is sufficient that **R** be positive definite and that **Q** and **S** be at least positive semidefinite.

In some cases it may turn out that certain elements of the **S** matrix are large enough to give computational difficulties. In this case, it is possible and perhaps desirable to obtain an inverse Riccati differential equation; we let

$$\mathbf{P}(t)\mathbf{P}^{-1}(t) = \mathbf{I} \tag{5.1-17}$$

and, by differentiating, we obtain

$$\dot{\mathbf{P}}\mathbf{P}^{-1}(t) + \mathbf{P}(t)\dot{\mathbf{P}}^{-1} = 0 \tag{5.1-18}$$

such that we obtain an "inverse" matrix Riccati equation

$$\dot{\mathbf{P}}^{-1} = \mathbf{A}(t)\mathbf{P}^{-1}(t) + \mathbf{P}^{-1}(t)\mathbf{A}^T(t) - \mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}^T(t) + \mathbf{P}^{-1}(t)\mathbf{Q}(t)\mathbf{P}^{-1}(t) \tag{5.1-19}$$

with

$$\mathbf{P}^{-1}(t_f) = \mathbf{S}^{-1} \tag{5.1-20}$$

In this way, for example, it is possible to solve the Riccati equation such that $\mathbf{S}^{-1} = [0]$, the null matrix, which will require that each and every component of the state vector approach the origin as the time approaches the terminal time. The "gains" $\mathbf{K}(t)$, or at least some components of them, become infinite at the terminal time in this case. It is also necessary to assume certain controllability requirements here, as we shall see in Chapter 11.

It is possible to write the nonlinear $n \times n$ matrix Riccati equation with a terminal condition as a $2n$ vector linear differential equation with two-point boundary conditions. We will use this approach, in part, to solve a Riccati equation associated with a filtering problem in Chapter 9. Our discussion of the second variation method in Chapter 13 will also make use of a Riccati transformation.

Example 5.1-1

Consider the scalar system

$$\dot{x} = -\tfrac{1}{2}x(t) + u(t), \qquad x(t_0) = x_0$$

with the cost function

$$J = \tfrac{1}{2}sx^2(t_f) + \tfrac{1}{2}\int_{t_0}^{t_f} [2x^2(t) + u^2(t)]\,dt$$

The Riccati equation, Eq. (5.1-12), becomes

$$\dot{p} = p + p^2 - 2, \qquad p(t_f) = s$$

which has a solution that we may write as either

$$p(t) = -0.5 + 1.5 \tanh(-1.5t + \xi_1)$$

or

$$p(t) = -0.5 + 1.5 \coth(-1.5t + \xi_2)$$

where ξ_1 and ξ_2 are adjusted such that $p(t_f) = s$.

For example, if

(a) $s = 0$, $t_f = 1$, then $\xi_1 = 1.845$ radians, which gives

$$K(t) = -R^{-1}B^TP = 0.5 - 1.5 \tanh(-1.5t + 1.845)$$

Since $s = 0$, we are not particularly weighting the state at the final time, and the "gain" (and control) goes to zero at the final time.

(b) $s = 10$, $t_f = 10$, then $\xi_2 = 15.1425$ radians. In this case we are applying a great weight to the error at $t = t_f$, and the gain becomes large (-10) at the terminal time.

(c) $s = \infty$, the Riccati equation cannot be solved directly since it has an infinite initial condition. The inverse Riccati equation can be solved with zero terminal condition to give

$$K^{-1}(t) \equiv [0.25 + 0.75 \tanh(-1.5t + 1.5t_f - 0.346)]$$

As t_f becomes infinite, it is easy to show that $K(t)$ becomes unity and, as is expected, the feedback gain becomes constant. Figure 5.1-2 illustrates $K(t)$, the "Kalman gains" as they are sometimes called, for these three cases for this particular problem.

Example 5.1-2

Let us consider the optimum closed-loop control for a nuclear reactor system. Specifically, we wish to consider a very simple reactor model with zero temperature feedback. Only one group of delayed neutrons will be used.

The reactor kinetics are described by the equations

$$\dot{n} = \frac{(\rho - \beta)n}{\Lambda} + \lambda c, \qquad \dot{c} = \frac{\beta n}{\Lambda} - \lambda c$$

where the neutron density, n, and the precursor concentration, c, are the state variables, and the reactivity ρ is the control variable. The system has the initial

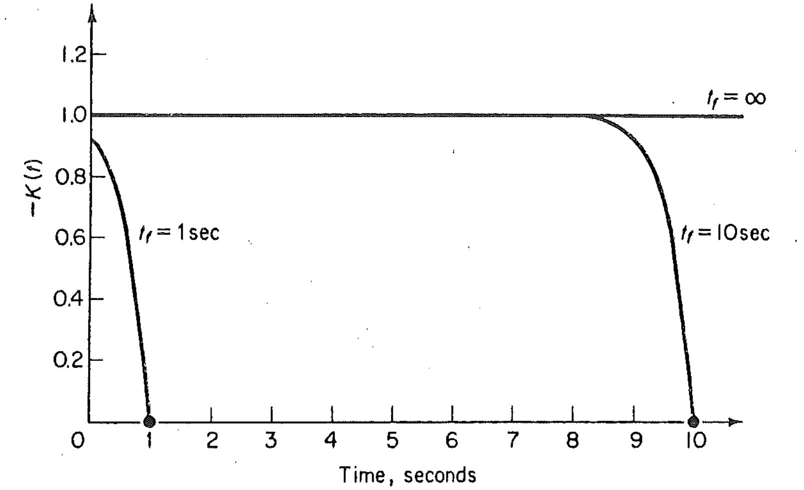


Fig. 5.1-2a (-1) times Kalman gain for controller, $s = 0$.

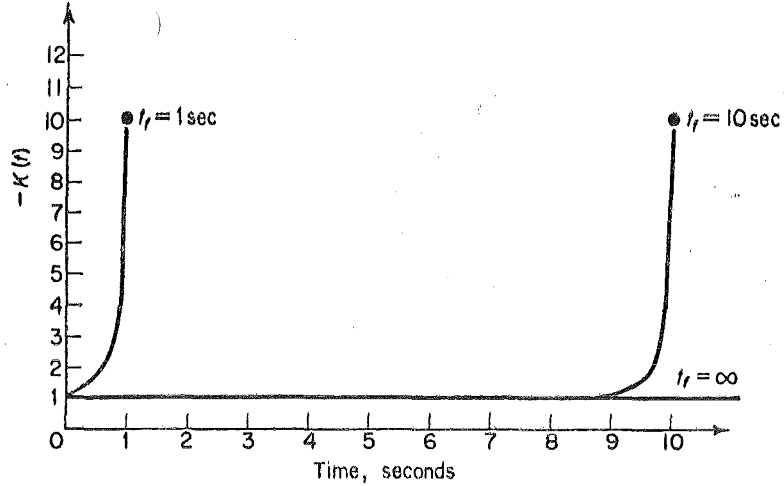


Fig. 5.1-2b (−1) times Kalman gain for controller, $s = 10$.

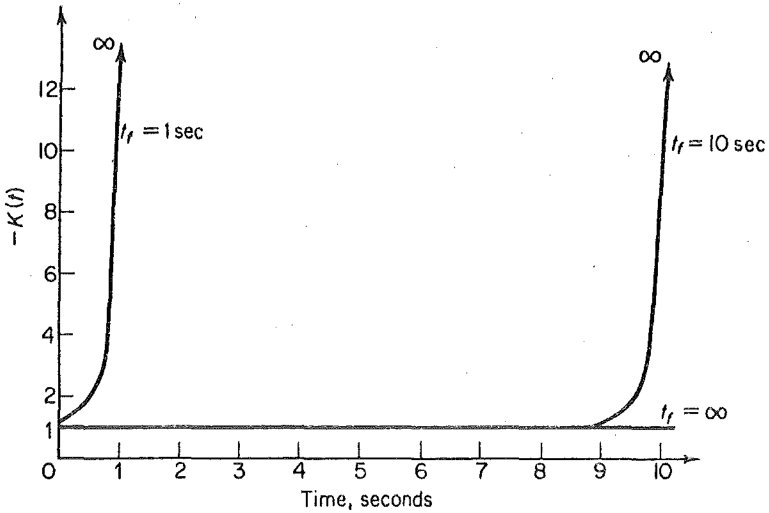


Fig. 5.1-2c (−1) times Kalman gain for controller, $s = \infty$.

conditions $n(0) = n_o$ and $c(0) = c_o$. β, Λ and λ are constants, the average fraction of precursors formed, effective neutron lifetime, and precursor decay constant.

The problem is to increase the power from the initial state n_o to a terminal state dn_o, where d is some constant greater than 1.0. The performance index for the system is

$$J_1 = \tfrac{1}{2}\int_0^{t_f} \dot{\rho}^2 \, dt$$

The control variable therefore becomes $\dot{\rho}$, and ρ, in effect, thus becomes a state variable. The kinetics equations may then be rewritten as

$$\dot{n} = \frac{(\rho - \beta)n}{\Lambda} + \lambda c$$

$$\dot{c} = \frac{\beta n}{\Lambda} - \lambda c$$

$$\dot{\rho} = u$$

where u is the control variable. Chapter 14 on quasilinearization indicates how the nonlinear two-point boundary value problem resulting from the use of optimal control theory may be used to obtain the optimum control and trajectory, which are shown in Fig. 5.1-3, for the following system parameters

$\lambda = 0.1\ \text{sec}^{-1}$

$\Lambda = 10^{-3}$ sec

$n_o = 10$ kW

$d = 5$

$\beta = 0.0064$

$t_f = 0.5$ sec

We will now develop a method of feedback control about the optimal trajectory which minimizes a cost function J_2; it will be quadratic in deviation from the nominal (optimal for J_1) trajectory and control.

Having formulated a model for the nuclear reactor system and determined the optimal trajectories, we now desire to determine the linearized system coefficient matrix about the optimal trajectory. The deviations of the state and control variables about the optimal or nominal trajectories are expressed by

$$n = n_n(t) + \Delta n(t), \qquad c = c_n(t) + \Delta c(t)$$

$$\rho = \rho_n(t) + \Delta\rho(t), \qquad u = u_n(t) + \Delta u(t)$$

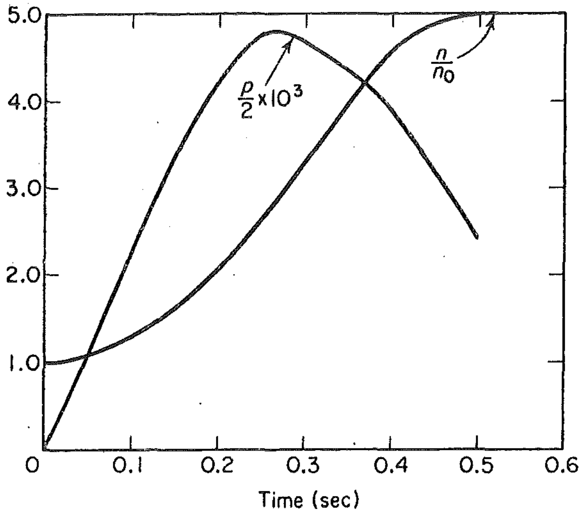


Fig. 5.1-3 Optimal control (reactivity) and trajectory (flux density) for Example (5.1-2).

The state vect

$$\Delta \mathbf{x}^T(t) = [\Delta n(t), \Delta c(t), \Delta \rho(t)]$$

The linearized model becomes

$$\Delta \dot{\mathbf{x}} = \begin{bmatrix} a_{11}(t) & \lambda & a_{13}(t) \\ \dfrac{\beta}{\Lambda} & -\lambda & 0 \\ 0 & 0 & 0 \end{bmatrix} \Delta \mathbf{x}(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \Delta u$$
$$= \mathbf{A}(t)\Delta \mathbf{x}(t) + \mathbf{b}(t)\Delta u(t)$$

where

$$a_{11}(t) = \frac{\rho_n(t) - \beta}{\Lambda}, \qquad a_{13}(t) = \frac{n_n(t)}{\Lambda}$$

To complete our design of the closed-loop controller, we must evaluate A(t) and b(t) about the optimum or nominal trajectories, select the R, Q, and S matrices, and solve the associated Riccati equation. The nominal trajectory, control, and time-varying gains are then stored and used to complete the closed-loop controller design.

The choice of the R, Q, and S matrices to minimize

$$J_2 = \tfrac{1}{2}\Delta \mathbf{x}^T(t_f)\mathbf{S}\Delta \mathbf{x}(t_f) + \tfrac{1}{2}\int_{t_0}^{t_f} [\Delta \mathbf{x}^T(t)\mathbf{Q}(t)\Delta \mathbf{x}(t) + r(t)\Delta u^2(t)]\, dt$$

is somewhat arbitrary and can perhaps best be done here by experimentation. We can accomplish this only after we have obtained a knowledge of possible disturbances which may drive the system off of the nominal trajectory. Let us assume that we will use

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 10^4 \end{bmatrix}, \qquad \mathbf{S} = 0, \qquad r = 1$$

In Chapter 13 the second variation and neighboring optimal methods of control-law computation will lead us to a method for choosing the proper weighting matrices for a variety of cases, in particular, for relating J_1 and J_2.

The control, $\Delta u(t)$, is computed from

$$\Delta u(t) = -\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{P}(t)\Delta \mathbf{x}(t)$$
$$= -[p_{31}(t)\Delta n(t) + p_{32}(t)\Delta c(t) + p_{33}(t)\Delta \rho(t)]$$

where it is necessary to solve the 3 × 3 matrix Riccati equation, having six different first-order differential equations, to obtain P(t). Figure (5.1-4) illustrates the Kalman gains $-\mathbf{K}^T(t) = [p_{31}(t), p_{32}(t), p_{33}(t)]$ for this example. Figure (5.1-5) indicates how the complete closed-loop controller is obtained. It is interesting to note that, in an actual physical problem, the precursor concentration is not measurable, and therefore we need to add an "observer" of this particular state variable. We also need to discuss many more aspects of this problem such as disturbances and parameter variations. We will postpone further consideration of these important questions until we establish some foundation in state and parameter estimation and optimal adaptive control. We have, in this example, illustrated how a basically nonlinear problem may be linearized, and a linear time-varying closed-loop controller obtained, if a nominal trajectory is known. Since this can be accomplished for a variety of problems, we see that the linear regulator problem is indeed an important one.

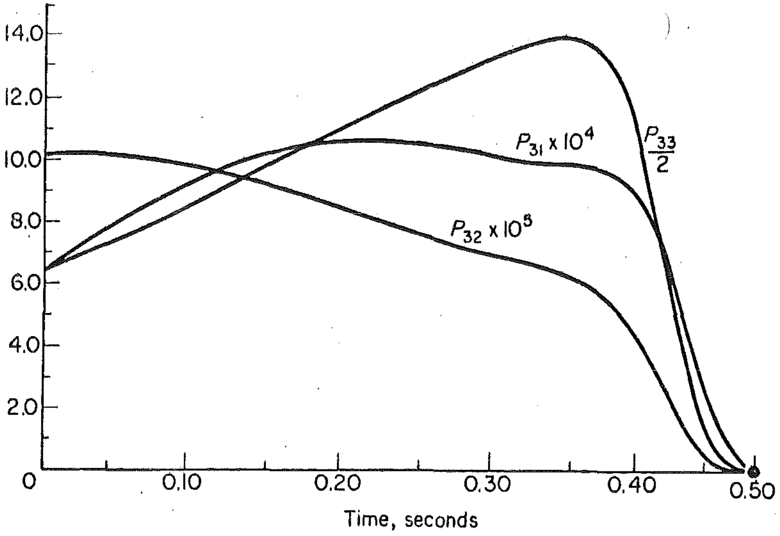


Fig. 5.1-4 Kalman gains for Example (5.1-2).

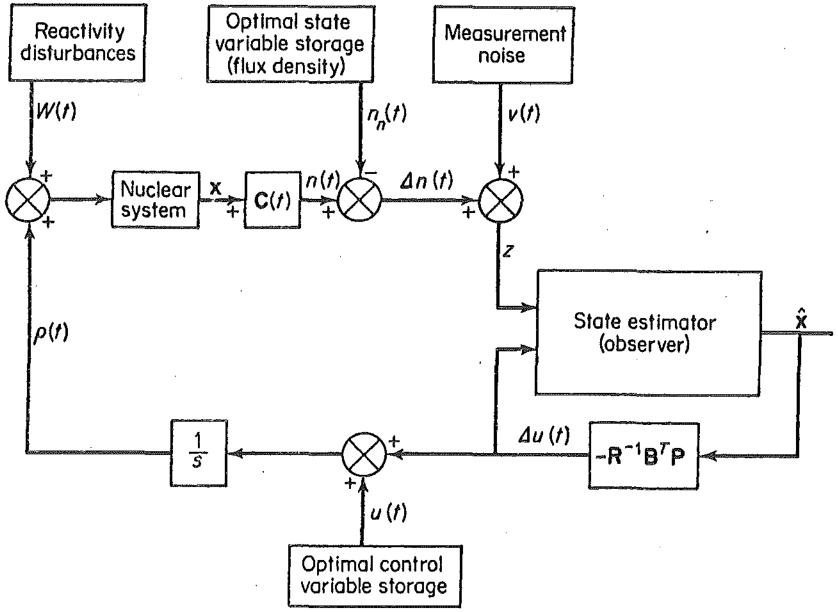


Fig. 5.1-5 Structure of controller for Example (5.1-2).

5.2 The ar servomechanism

The linear regulator problem considered in the preceding section can be generalized in several ways. We can assume that we desire to find the control in such a way as to cause the output to track or follow a desired output state, $\boldsymbol{\eta}(t)$. We may also assume that there is a forcing function (not the control) for the system differential equations. Therefore, we will consider the minimization of

$$J = \tfrac{1}{2}\|\boldsymbol{\eta}(t_f) - \mathbf{z}(t_f)\|_{\mathbf{S}}^2 + \tfrac{1}{2}\int_{t_o}^{t_f} [\|\boldsymbol{\eta}(t) - \mathbf{z}(t)\|_{\mathbf{Q}(t)}^2 + \|\mathbf{u}(t)\|_{\mathbf{R}(t)}^2]\, dt \quad (5.2\text{-}1)$$

for the system which contains an input or plant noise vector $\mathbf{w}(t)$

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{w}(t), \qquad \mathbf{x}(t_o) = \mathbf{x}_o \quad (5.2\text{-}2)$$

$$\mathbf{z}(t) = \mathbf{C}(t)\mathbf{x}(t) \quad (5.2\text{-}3)$$

The requirements on the various matrices are the same as in the preceding section. We proceed in exactly the same fashion as for the regulator problem. The Hamiltonian is, from Eq. (4.3-34),

$$\begin{aligned} H(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, t) = {} & \tfrac{1}{2}\|\boldsymbol{\eta}(t) - \mathbf{C}(t)\mathbf{x}(t)\|_{\mathbf{Q}(t)}^2 + \tfrac{1}{2}\|\mathbf{u}(t)\|_{\mathbf{R}(t)}^2 \\ & + \boldsymbol{\lambda}^T(t)[\mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{w}(t)] \end{aligned} \quad (5.2\text{-}4)$$

We employ the maximum principle and set $\partial H/\partial \mathbf{u} = \mathbf{0}$ to obtain

$$\mathbf{u}(t) = -\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\boldsymbol{\lambda}(t) \quad (5.2\text{-}5)$$

and

$$\frac{\partial H}{\partial \mathbf{x}} = -\dot{\boldsymbol{\lambda}} = \mathbf{C}^T(t)\mathbf{Q}(t)[\mathbf{C}(t)\mathbf{x}(t) - \boldsymbol{\eta}(t)] + \mathbf{A}^T(t)\boldsymbol{\lambda}(t) \quad (5.2\text{-}6)$$

with the terminal condition

$$\boldsymbol{\lambda}(t_f) = \mathbf{C}^T(t_f)\mathbf{S}[\mathbf{C}(t_f)\mathbf{x}(t_f) - \boldsymbol{\eta}(t_f)] \quad (5.2\text{-}7)$$

In order to attempt to determine a closed-loop control, we assume

$$\boldsymbol{\lambda}(t) = \mathbf{P}(t)\mathbf{x}(t) - \boldsymbol{\xi}(t) \quad (5.2\text{-}8)$$

We substitute this relation into the canonic equations and determine the requirements for a solution. By a procedure analogous to that of the preceding section, we easily obtain the following requirements

$$\begin{aligned} \dot{\mathbf{P}} = {} & -\mathbf{P}(t)\mathbf{A}(t) - \mathbf{A}^T(t)\mathbf{P}(t) \\ & + \mathbf{P}(t)\mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{P}(t) - \mathbf{C}^T(t)\mathbf{Q}(t)\mathbf{C}(t) \end{aligned} \quad (5.2\text{-}9)$$

$$\mathbf{P}(t_f) = \mathbf{C}^T(t_f)\mathbf{S}\mathbf{C}(t_f) \quad (5.2\text{-}10)$$

and

$$\dot{\boldsymbol{\xi}} = -[\mathbf{A}(t) - \mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{P}(t)]^T\boldsymbol{\xi} + \mathbf{P}(t)\mathbf{w}(t) - \mathbf{C}^T(t)\mathbf{Q}(t)\boldsymbol{\eta}(t) \quad (5.2\text{-}11)$$

$$\boldsymbol{\xi}(t_f) = \mathbf{C}^T(t_f)\mathbf{S}\boldsymbol{\eta}(t_f) \quad (5.2\text{-}12)$$

Thus we see that the linear servomechanism problem omposed of two parts: a linear regulator part, plus a prefilter to determine the optimal driving function from the desired value, $\boldsymbol{\eta}(t)$, of the system output. The optimum control law is linear and is obtained from Eq. (5.2-5) as

$$\mathbf{u}(t) = -\mathbf{R}^{-1}(t)\mathbf{B}^T(t)[\mathbf{P}(t)\mathbf{x}(t) - \boldsymbol{\xi}(t)] \quad (5.2\text{-}13)$$

Unfortunately, the optimal control is, in practice, often computationally unrealizable because it involves $\boldsymbol{\xi}(t)$ which must be solved backward from t_f to t_o and, therefore, requires a knowledge of $\boldsymbol{\eta}(t)$ and $\mathbf{w}(t)$ for all time $t \in [t_o, t_f]$. This is quite often not known at the initial time t_o.

Example 5.2-1

Let us consider the minimization of the cost function

$$J = \tfrac{1}{2}\int_0^{t_f} [(x_1 - \eta_1)^2 + u^2]\, dt$$

for the system described by

$$\dot{x}_1 = x_2, \qquad x_1(0) = x_{10}$$
$$\dot{x}_2 = u, \qquad x_2(0) = x_{20}$$

We first use Eqs. (5.2-9) and (5.2-10) to obtain the Riccati equation for this example

$$\dot{p}_{11} = p_{12}^2 - 1, \qquad p_{11}(t_f) = 0$$
$$\dot{p}_{12} = -p_{11} + p_{12}p_{22}, \qquad p_{12}(t_f) = 0$$
$$\dot{p}_{22} = -2p_{12} + p_{22}^2, \qquad p_{22}(t_f) = 0$$

If we allow t_f to become infinite, we obtain the solution $p_{11} = p_{22} = \sqrt{2}$, $p_{12} = 1$. Thus we have for the closed-loop control

$$u = -\mathbf{R}^{-1}\mathbf{B}^T[\mathbf{P}\mathbf{x} - \boldsymbol{\xi}] = -x_1 - \sqrt{2}\, x_2 + \xi_2$$

where we must determine $\boldsymbol{\xi}$ by solving Eqs. (5.2-11) and (5.2-12) which become for this example

$$\dot{\xi}_1 = \xi_1 - \eta_1, \qquad \xi_1(t_f) = 0$$
$$\dot{\xi}_2 = -\xi_1 + \sqrt{2}\,\xi_2, \qquad \xi_2(t_f) = 0$$

If $\eta_1 = \alpha$, a constant, for t greater than zero, we are justified in obtaining the equilibrium solution for the $\dot{\boldsymbol{\xi}}$ equation if $t_f = \infty$ by setting $\dot{\boldsymbol{\xi}} = \mathbf{0}$ to obtain $\xi_2 = 0.707\xi_1 = \eta_1 = \alpha$. If $\eta_1 = 1 - e^{-t}$, we will then find by a simple limiting process that for $t_f = \infty$,

$$\xi_2(t) = 1 + \frac{1}{2 + \sqrt{2}}e^{-t}, \qquad t \geq 0$$

We may realize this solution as shown in Figure (5.2-1).

We note that if $\mathbf{w}(t) = \boldsymbol{\eta}(t) = \mathbf{0}$, or for that matter, any vector constant in time, the servomechanism problem reduces to a regulator problem except that it is an "output" regulator problem rather than a "state" regulator problem because of the presence of the output matrix $\mathbf{C}(t)$. It is not necessary

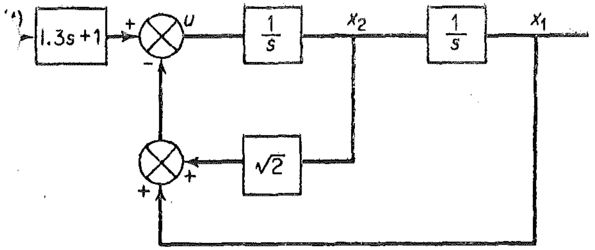


Fig. 5.2-1 Block diagram of optimum servomechanism for Example (5.2-1).

for the system to be controllable in order to find a solution to the regulator problem. The only exception to this is in the limiting cases where S becomes infinite or where t_f becomes infinite. It is, however, necessary that the system be observable in order for a solution to the output regulator problem to exist. We will expand considerably on these ideas when we consider controllability, observability, and the reachable zone problem in Chapter 11.

It is possible to give a frequency-domain interpretation to the regulator and servomechanism problem for the infinite time interval case for a constant system. We will present this method, due to Kalman, in Chapter 9 where the duality concept will allow us to treat both the estimation and the control problems.

5.3 Bang bang control and minimum time problems

Maximum effort control problems have become increasingly important in a variety of applications. It is natural that we ask under what circumstances optimal controls will always be maximum effort, or *bang bang*. To do this, we will restrict each component of the control vector, $\mathbf{u}(t)$, to some bounded interval. Let us consider the nonlinear differential system where the control enters in a linear fashion a linear combination of u_i

$$\dot{\mathbf{x}} = \mathbf{f}[\mathbf{x}(t), t] + \mathbf{G}[\mathbf{x}(t), t]\mathbf{u}(t), \qquad \mathbf{x}(t_o) = \mathbf{x}_o \tag{5.3-1}$$

$$a_i \le u_i \le b_i, \qquad \forall i \tag{5.3-2}$$

and assume a performance index which, likewise, contains only linear terms in the control variable, such that the Hamiltonian will also be linear in $\mathbf{u}(t)$.

$$J = \theta[\mathbf{x}(t_f), t_f] + \int_{t_o}^{t_f} \{\phi[\mathbf{x}(t), t] + \mathbf{h}^T[\mathbf{x}(t), t]\mathbf{u}(t)\}\, dt \tag{5.3-3}$$

$$\begin{aligned} H[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t] = \phi[\mathbf{x}(t), t] &+ \mathbf{h}^T[\mathbf{x}(t), t]\mathbf{u}(t) \\ &+ \boldsymbol{\lambda}^T(t)\{\mathbf{f}[\mathbf{x}(t), t] + \mathbf{G}[\mathbf{x}(t), t]\mathbf{u}(t)\} \end{aligned} \tag{5.3-4}$$

Since the Hamiltonian is linear in the control vector, $\mathbf{u}(t)$, minimization of the Hamiltonian with respect to $\mathbf{u}(t)$ requires that

$$u_i = \begin{cases} a_i & \text{if } \{\mathbf{h}^T[\mathbf{x}(t), t] + \boldsymbol{\lambda}^T(t)\mathbf{G}[\mathbf{x}(t), t]\}_i > 0 \\ b_i & \text{if } \{\mathbf{h}^T[\mathbf{x}(t), t] + \boldsymbol{\lambda}^T(t)\mathbf{G}[\mathbf{x}(t), t]\}_i < 0 \end{cases} \tag{5.3-5}$$

Thus we see that when the control vector appears [illegible]arly in both the equation of motion of the differential system and the performance index, and if in addition each component of the control vector is bounded, the optimal control is bang bang. The only exception to this occurs in cases where

where ith component is 0.

$$\mathbf{h}^T[\mathbf{x}(t), t] + \boldsymbol{\lambda}^T(t)\mathbf{G}[\mathbf{x}(t), t] = \mathbf{0} \quad \text{or } [\ \ ]_i = 0 \tag{5.3-6}$$

for then the Hamiltonian is not a function of $\mathbf{u}(t)$ and cannot be minimized with respect to $\mathbf{u}(t)$. When Eq. (5.3-6) holds for more than isolated points in time, the optimization problem is said to possess a singular solution, a problem which we will discuss in detail in the next section. A singular solution is possible with respect to a particular control component, u_i, if the ith component of Eq. (5.3-6) is zero.

For this problem, the canonic equations are obtained as

$$\dot{\mathbf{x}} = \frac{\partial H}{\partial \boldsymbol{\lambda}} = \mathbf{f}[\mathbf{x}(t), t] + \mathbf{G}[\mathbf{x}(t), t]\mathbf{u}(t) \tag{5.3-7}$$

$$\begin{aligned} -\dot{\boldsymbol{\lambda}} = \frac{\partial H}{\partial \mathbf{x}} &= \frac{\partial \phi[\mathbf{x}(t), t]}{\partial \mathbf{x}} + \frac{\partial \mathbf{h}^T[\mathbf{x}(t), t]}{\partial \mathbf{x}}\mathbf{u}(t) \\ &+ \frac{\partial \mathbf{f}^T[\mathbf{x}(t), t]}{\partial \mathbf{x}}\boldsymbol{\lambda}(t) + \frac{\partial \{\mathbf{G}[\mathbf{x}(t), t]\mathbf{u}(t)\}^T}{\partial \mathbf{x}(t)}\boldsymbol{\lambda}(t) \end{aligned} \tag{5.3-8}$$

where $\mathbf{u}(t)$ is determined via Eq. (5.3-5). Since we have not specifically stated the end conditions, we have carried the general problem about as far as is possible. When we specify information concerning the desired states at the terminal time and the initial condition vector, we have, as before, a two-point boundary value problem with half of the conditions specified at the initial time and half at the terminal time. A possible method of solution of the canonic equations for this formulation consists of reversing time in the canonic equations. Starting at the determined or specified terminal vector, which often is the origin of the state vector, we integrate back from this point with a constant control until a switching point is obtained from Eq. (5.3-5). Since no terminal conditions are present for half of the state variables, the method is, of necessity, cut and try. Chapters 13, 14, and 15 provide more systematic methods for solving this type of two-point boundary value problem.

We shall now illustrate various solutions to a particular case which results in bang bang control—the minimum time problem for constant linear systems with a scalar input. In this problem, we desire to transfer an n vector constant differential system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t), \qquad \mathbf{x}(t_o) = \mathbf{x}_o \tag{5.3-9}$$

to the origin, $\mathbf{x}(t_f) = \mathbf{0}$, in minimum time, such that we have for the cost function

$$J = \int_{t_o}^{t_f} (1)\, dt = t_f - t_o \tag{5.3-10}$$

al becomes zero the discrete Kalman
0) through (10.3-40) as well as the
on algorithms of Eqns. (10.3-51) through
– Kalman filter algorithms of Fig (9.3-2)

) $t_2 \ge t_1$
riance becomes
$|t_1)\} = \varphi(t_2,t_1) P(t_1) \varphi^T(t_2,t_1)$
$(k) Q(k) \Gamma^T(k) \varphi^T(t_2,k)$
ession for the error variance in

int boundary value problem Eqs. (9.4-19)
two pt. boundary value problem of
yields fixed interval smoothing
k_f). Show that the Kalman filter
oundary value problems result if
$\lambda(t|t_f) + \hat{x}(t)$
k) $\lambda(k|k_f) + \hat{x}(k)$
. the smoothing and filtering solutions.

11

CONTROLLABILITY AND OBSERVABILITY —THE SEPARATION THEOREM

In our previous work with the regulator and servomechanism problems, we noted that there were certain requirements, in addition to the definiteness of certain matrices, which must exist in order for the problem to have a meaningful solution. In this chapter we wish to examine these requirements, which we have postponed until now so that we might explore them using optimum control and filtering theory.

First we will examine an intrinsic characterization of the manner in which the output of a system is constrained with respect to the ability to observe system states. Then we will examine the dual requirement and find the characterization of the manner in which a system is constrained with respect to control of the system states or system outputs. We will consider these requirements for both continuous and discrete systems and will thus prove the observability and controllability requirements for linear systems. Original efforts in this area are due to Kalman Ho and Narendra [1, 2, 3, 4], Kreindler and Sarachik [5], Lee [6], and Gilbert [7].

We shall then turn our attention to systems that are partially observable in that the output vector contains all information necessary for the unique recovery of each component of the state vector. We discuss two methods for the construction of observers, the first due to Kalman [8], and the second to Luenberger [9].

Finally, we pose the problem of combined estimation and control in which we not only have the requirement for state estimation but also the requirement to use the estimated state in such way as to generate an optimal control law. This problem has been treated by Kalman [10], Joseph and Tou [11], Gunckel and Franklin [12], and others [13, 14]. It lays the foundation for the optimal adaptive problem which we shall consider in later chapters.

11.1 Observability in linear dynamic systems

In Chapters 8, 9, and 10 we developed various concepts concerning state estimation in linear continuous and linear discrete systems. To accomplish state estimation, it is necessary that certain requirements with respect to observability be met.

For a system to be observable, it must be possible to determine the state of an unforced system from the knowledge of the output of the system over some time interval. Specifically, in an unobservable system, it is impossible to determine an initial state vector $\mathbf{x}(t_o)$ from a knowledge of the output, $\mathbf{z}(t)$. Of course, we must be able to do this if we are concerned with control of system state variables as we are in the regulator problem. We shall first discuss the observability requirement for linear discrete systems and then proceed to a discussion of linear continuous systems.

11.1-1 Observability in time-varying discrete systems

Let us suppose that we have a system whose state is described by the unforced vector difference equation

$$\mathbf{x}(k+1) = \mathbf{A}(k)\mathbf{x}(k) \tag{11.1-1}$$

and suppose that we observe a vector $\mathbf{z}(k)$ which is a linear combination of the system states plus an additive noise term

$$\mathbf{z}(k) = \mathbf{C}(k)\mathbf{x}(k) + \mathbf{v}(k) \tag{11.1-2}$$

We desire to find the best least-squares estimate, $\hat{\mathbf{x}}(k)$, of $\mathbf{x}(k)$ by minimizing

$$J = \tfrac{1}{2}\sum_{k=k_o}^{k_f} \|\mathbf{z}(k) - \mathbf{C}(k)\hat{\mathbf{x}}(k)\|^2_{\mathbf{R}^{-1}(k)} \tag{11.1-3}$$

subject to the constraint of Eq. (11.1-1) with $\mathbf{x}(k)$ replaced by $\hat{\mathbf{x}}(k)$. This is a multistage decision process, and since Eq. (11-1.1) holds, we can write

$$\mathbf{x}(k_o+1) = \mathbf{A}(k_o)\mathbf{x}(k_o),$$

$$\mathbf{x}(k_o+2) = \mathbf{A}(k_o+1)\mathbf{x}(k_o+1) = \mathbf{A}(k_o+1)\mathbf{A}(k_o)\mathbf{x}(k_o)$$

Thus it is clear that

$$\mathbf{x}(k_o+k) = \boldsymbol{\varphi}(k_o+k, k_o)\mathbf{x}(k_o) \tag{11.1-4}$$

where

$$\boldsymbol{\varphi}(k_o+k, k_o) = \mathbf{A}(k_o+k-1)\ldots\mathbf{A}(k_o+1)\mathbf{A}(k_o) = \prod_{k=k_o}^{k_o+k-1}\mathbf{A}(k) \tag{11.1-5}$$

$$\boldsymbol{\varphi}(k_o, k_o) = \mathbf{I} \tag{11.1-6}$$

Since matrix multiplication is not commutative, we realize that we must form the product in Eq. (11.1-5) in the proper order. Now we can write

$$\mathbf{x}(k) = \boldsymbol{\varphi}(k, k_o)\mathbf{x}(k_o) \tag{11.1-7}$$

By using Eq. (11.1-7), we can write the cost function as

$$J = \tfrac{1}{2}\sum_{k=k_o}^{k_f} \|\mathbf{z}(k) - \mathbf{C}(k)\boldsymbol{\varphi}(k, k_o)\hat{\mathbf{x}}(k_o)\|^2_{\mathbf{R}^{-1}(k)} \tag{11.1-8}$$

which includes the constraint Eq. (11.1-1), since it has been used to formulate the equation.

We wish to minimize Eq. (11.1-8). To do this we will solve $\partial J/\partial\hat{\mathbf{x}}(k_o) = \mathbf{0}$, which is the usual necessary condition for a minimum. In doing this we obtain from Eq. (11.1-8)

$$\sum_{k=k_o}^{k_f} \boldsymbol{\varphi}^T(k, k_o)\mathbf{C}^T(k)\mathbf{R}^{-1}(k)[\mathbf{z}(k) - \mathbf{C}(k)\boldsymbol{\varphi}(k, k_o)\hat{\mathbf{x}}(k_o)] = \mathbf{0} \tag{11.1-9}$$

We note that $\hat{\mathbf{x}}(k_o)$ may be removed from the summation sign. By doing this and solving the resulting equation, we obtain

$$\hat{\mathbf{x}}(k_o) = \mathbf{M}^{-1}(k_o, k_f)\sum_{k=k_o}^{k_f} \boldsymbol{\varphi}^T(k, k_o)\mathbf{C}^T(k)\mathbf{R}^{-1}(k)\mathbf{z}(k) \tag{11.1-10}$$

as the best initial condition, where we have defined

$$\mathbf{M}(k_o, k_f) = \sum_{k=k_o}^{k_f} \boldsymbol{\varphi}^T(k, k_o)\mathbf{C}^T(k)\mathbf{R}^{-1}(k)\mathbf{C}(k)\boldsymbol{\varphi}(k, k_o) \tag{11.1-11}$$

Clearly, $\mathbf{M}(k_f, k_o)$ must have an inverse and, therefore, must be nonsingular. Kalman's condition for observability goes even further, in that it requires $\mathbf{M}(k_f, k_o)$ to be positive-definite. We recall that a positive-definite matrix $\mathbf{F}$ is defined as one such that $\mathbf{x}^T\mathbf{F}\mathbf{x} > 0$ for any nonzero $\mathbf{x}$. Also real symmetric matrix $\mathbf{F}$ is positive-definite if and only if there exists a nonsingular matrix $\mathbf{D}$ such that $\mathbf{F} = \mathbf{D}^T\mathbf{D}$. We note that $\mathbf{D}$, being nonsingular, implies that $\mathbf{F}$ is nonsingular also, since $\det(\mathbf{F}) = [\det(\mathbf{D})]^2$. Since $\mathbf{M}$ is of the form $\mathbf{D}^T\mathbf{D}$, the positive-definite requirement really only requires that $\mathbf{M}$ be nonsingular. For observability, we are not at all concerned with the specific nature of the positive-definite weighting matrix $\mathbf{R}$, and thus we set $\mathbf{R} = \mathbf{I}$ in Eq. (11.1-11).

Example 11.1-1

Suppose we have two integrators in cascade as in Fig. 11.1-1a. We ask: Can we estimate $\mathbf{x}^T = [x_1, x_2]$ by observing $\mathbf{z}$? Obviously not, because we do not know the initial condition on the second integrator. In this case we would find $\mathbf{M}$ to be singular and thus not positive definite.

Now suppose that we add a switch to the system as shown in Fig. 11.1-1b. We begin by observing $\mathbf{z}^T = [z_1, z_2]$ at some time $t_o < t_1$. Can we estimate $\mathbf{x}$? We would find that $\mathbf{M}$ is singular for $t < t_1$ and nonsingular thereafter, indicating that the system is observable for $t > t_1$, and nonobservable for $t < t_1$. This is

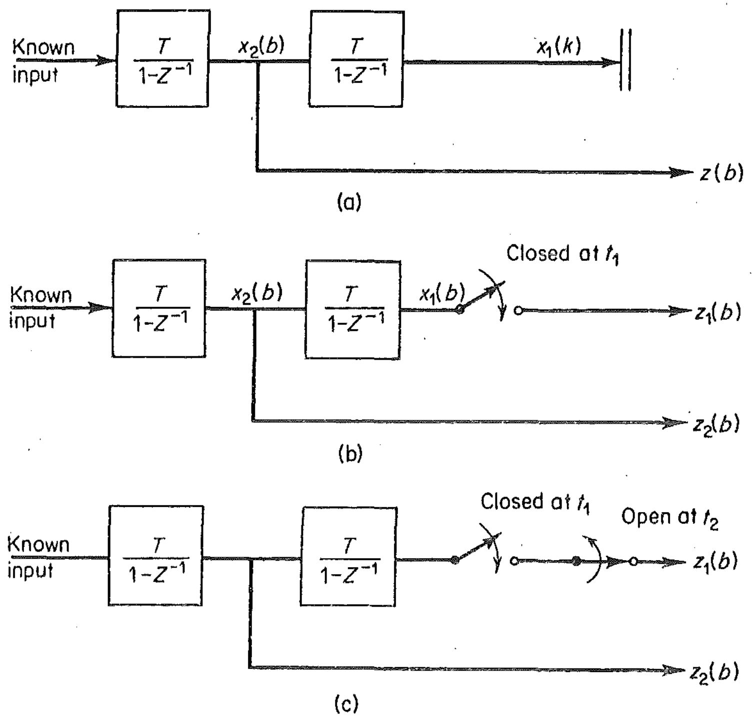


Fig. 11.1-1 A simple system which is a) unobservable b), c) observable for $t > t_1$.

what we could expect intuitively. Lastly, we add another switch, which we open at time t_2 as shown in Fig. 11.1-1c. In this case, the system would be nonobservable for $t < t_1$, but observable thereafter, even for $t > t_2$. This is because of the fact that, once we know the value of x_1 for some time t_o, we know x_1 for all time, provided x_2 is known, and we are always observing x_2. Thus, M will be singular for $t < t_1$ and nonsingular thereafter. There is a general theorem we could have applied to the third part of this example [2] which states that the rank of $\mathbf{M}(k_f, k_o)$ is nondecreasing with increasing time or, here, increasing k_f.

It is not necessary that we interpret the observability condition through the use of a least-squares curve fitting procedure. From Eqs. (11.1-2) and (11.1-7) we can set up a vector Z composed of

$$\boldsymbol{Z} = \begin{bmatrix} \mathbf{z}(k_o) \\ \mathbf{z}(k_o+1) \\ \mathbf{z}(k_o+2) \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{z}(k_f) \end{bmatrix} = \begin{bmatrix} \mathbf{C}(k_o) \\ \mathbf{C}(k_o+1)\boldsymbol{\varphi}(k_o+1, k_o) \\ \mathbf{C}(k_o+2)\boldsymbol{\varphi}(k_o+2, k_o) \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{C}(k_f)\boldsymbol{\varphi}(k_f, k_o) \end{bmatrix} \hat{\mathbf{x}}(k_o) = \boldsymbol{\Delta}^T(k_o, k_f)\hat{\mathbf{x}}(k_o) \tag{11.1-12}$$

such that

$$\boldsymbol{\Delta}(k_o, k_f) = [\mathbf{C}^T(k_o) \mid \boldsymbol{\varphi}^T(k_o+1, k_o)\mathbf{C}^T(k_o+1) \mid \cdots \mid \boldsymbol{\varphi}^T(k_f, k_o)\mathbf{C}^T(k_f)] \tag{11.1-13}$$

To solve for $\hat{\mathbf{x}}(t_o)$, it is necessary that $\boldsymbol{\Delta}(k_o, k_f)$ be of rank n ($\mathbf{x}$ is an n vector). This provides us with an alternative test for observability. If we premultiply Eq. (11.1-12) by $\boldsymbol{\Delta}(k_o, k_f)$, we have

$$\sum_{k=k_o}^{k_f} \boldsymbol{\varphi}^T(k, k_o)\mathbf{C}^T(k)\mathbf{z}(k) = [\sum_{k=k_o}^{k=k_f} \boldsymbol{\varphi}^T(k, k_o)\mathbf{C}^T(k)\mathbf{C}(k)\boldsymbol{\varphi}(k, k_o)]\mathbf{x}(k_o) \tag{11.1-14}$$

Thus we again have

$$\hat{\mathbf{x}}(k_o) = \mathbf{M}^{-1}(k_o, k_f) \sum_{k=k_o}^{k_f} \boldsymbol{\varphi}^T(k, k_o)\mathbf{C}^T(k)\mathbf{z}(k) \tag{11.1-15}$$

where $\mathbf{M}(k_o, k_f)$ has been previously defined by Eq. (11.1-11). The matrix $\mathbf{M}(k_o, k_f)$ is sometimes called the Gramian matrix and is nonsingular if and only if the matrix $\boldsymbol{\Delta}(k_o, k_f)$ is of rank n. Thus there certainly must be at least n columns in $\boldsymbol{\Delta}(k_o, k_f)$, which requires that the minimum sequence length, $k_f - k_o$ is $(n/m - 1)$, where $\mathbf{x}$ is an n vector and $\mathbf{z}$ is an m vector.

For constant discrete systems where A and C are stage invariant, these results simplify somewhat since $\boldsymbol{\varphi}(k, k_o) = \mathbf{A}^{(k-k_o)}$, $\mathbf{C}(k) = \mathbf{C}$, and the observability requirement becomes that the matrix

$$\boldsymbol{\Delta}(k) = [\mathbf{C}^T \mid \mathbf{A}^T\mathbf{C}^T \mid \mathbf{A}^{T^2}\mathbf{C}^T \mid \mathbf{A}^{T^3}\mathbf{C}^T \mid \cdots \mid \mathbf{A}^{T^{k-1}}\mathbf{C}^T] \tag{11.1-16}$$

be of rank n. If a constant system is not observable on a sequence of length $k = n$, it is, of course, not observable on any sequence. This is not the case for stage-varying or nonconstant systems as indicated in Example 11.1-1. In many cases, it will be computationally more convenient to determine whether or not the $n \times n$ matrix $\boldsymbol{\Delta}\boldsymbol{\Delta}^T$ is of rank n rather than the $n \times nm$ matrix $\boldsymbol{\Delta}$ of Eq. (11.1-16). This statement will apply to the many matrices of the form of Eq. (11.1-16) which we will encounter in this section and the next.

11.1-2 Observability in continuous systems

We have previously derived the observability condition for discrete static and dynamic systems. Now consider a continuous dynamic system represented by the n vector equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) \tag{11.1-17}$$

where we observe (measure) an m vector output

$$\mathbf{z}(t) = \mathbf{C}(t)\mathbf{x}(t) + \mathbf{v}(t) \tag{11.1-18}$$

where $\mathbf{v}(t)$ is additive measurement noise. We wish to find the best least-square estimator, $\hat{\mathbf{x}}(t)$, of $\mathbf{x}(t)$ such that the cost function

$$J = \tfrac{1}{2} \int_{t_o}^{t_f} \| \mathbf{z}(t) - \mathbf{C}(t)\hat{\mathbf{x}}(t) \|^2_{\mathbf{R}^{-1}(t)} \, dt \tag{11.1-19}$$

is minimized, subject to the constraint

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{A}(t)\hat{\mathbf{x}}(t) \tag{11.1-20}$$

We could obviously apply the maximum principle, but instead, we will use another, simpler approach as follows. The solution to Eq. (11.1-20) is

$$\hat{\mathbf{x}}(t) = \boldsymbol{\varphi}(t, \tau)\mathbf{x}(\tau) \tag{11.1-21}$$

where

$$\frac{\partial \boldsymbol{\varphi}(t, \tau)}{\partial t} = \mathbf{A}(t)\boldsymbol{\varphi}(t, \tau), \boldsymbol{\varphi}(t, t) = \mathbf{I} \tag{11.1-22}$$

Therefore, given $\hat{\mathbf{x}}(t_1)$ at some time $t_1 = \tau$, we can find $\hat{\mathbf{x}}(t_1)$ at any other time $t_1 = t$ by choosing the proper transition matrix $\boldsymbol{\varphi}(t, \tau)$.

We can use Eq. (11.1-21) to replace $\hat{\mathbf{x}}(t)$ in the cost function, Eq. (11.1-19). In so doing, we are free to choose any value of t we desire. It seems that a reasonable choice is $t = t_f$, since we will then obtain a solution for that value of $\hat{\mathbf{x}}(t_f)$ (i.e., the final state) which gives least-square error. In addition, we have previously given the solution for $\mathbf{x}(k_o)$ for the discrete case. Thus, the cost function becomes

$$J = \tfrac{1}{2}\int_{t_o}^{t_f} \|\mathbf{z}(t) - \mathbf{C}(t)\boldsymbol{\varphi}(t, t_f)\hat{\mathbf{x}}(t_f)\|^2_{\mathbf{R}^{-1}(t)}\, dt \tag{11.1-23}$$

To determine the particular $\hat{\mathbf{x}}(t_f)$ that minimizes Eq. (11.1-23), we must solve

$$\frac{\partial J}{\partial \hat{\mathbf{x}}(t_f)} = 0 = \int_{t_o}^{t_f} \boldsymbol{\varphi}^T(t, t_f)\mathbf{C}^T(t)\mathbf{R}^{-1}(t)[\mathbf{z}(t) - \mathbf{C}(t)\boldsymbol{\varphi}(t, t_f)\hat{\mathbf{x}}(t_f)]\, dt \tag{11.1-24}$$

which gives

$$\left[\int_{t_o}^{t_f} \boldsymbol{\varphi}^T(t, t_f)\mathbf{C}^T(t)\mathbf{R}^{-1}(t)\mathbf{C}(t)\boldsymbol{\varphi}(t, t_f)\, dt\right]\hat{\mathbf{x}}(t_f) = \int_{t_o}^{t_f} \boldsymbol{\varphi}^T(t, t_f)\mathbf{C}^T(t)\mathbf{R}^{-1}(t)\mathbf{z}(t)\, dt \tag{11.1-25}$$

We now define

$$\mathbf{N}(t_o, t_f) = \int_{t_o}^{t_f} \boldsymbol{\varphi}^T(t, t_f)\mathbf{C}^T(t)\mathbf{R}^{-1}(t)\mathbf{C}(t)\boldsymbol{\varphi}(t, t_f)\, dt \tag{11.1-26}$$

so that

$$\hat{\mathbf{x}}(t_f) = \mathbf{N}^{-1}(t_o, t_f)\int_{t_o}^{t_f} \boldsymbol{\varphi}^T(t, t_f)\mathbf{C}(t)\mathbf{R}^{-1}(t)\mathbf{z}(t)\, dt \tag{11.1-27}$$

Clearly, the matrix of Eq. (11.1-26) must have an inverse or, in other words, must be nonsingular. Furthermore, by computing the second derivative $\partial^2 J/\partial \mathbf{x}^2$, we see that we require $\mathbf{N}(t_o, t_f)$ to be positive-definite in order to establish sufficient conditions for a minimum of the cost function. Thus, a system becomes observable at time t_f when the matrix $\mathbf{N}(t_o, t_f)$ is positive-definite for $t_f, t_f > t_o$. Again, it can be shown that the rank of the matrix $\mathbf{N}(t_o, t_f)$ is nondecreasing with time. In other words, once a system becomes observable at $t = t_1$, it remains observable for all $t > t_1$. For observability, the matrix $\mathbf{R}$ is again set equal to the identity matrix $\mathbf{I}$.

We will again offer an alternate derivation of the observability requirement. The output of the system $\mathbf{z}(t)$ is from Eqs. (11.1-18) and (11.1-21).

$z(t) = C(t)\,\hat{x}(t).$

$$\mathbf{z}(t) = \mathbf{C}(t)\boldsymbol{\varphi}(t, t_f)\hat{\mathbf{x}}(t_f) \tag{11.1-28}$$

By premultiplying this equation by $\boldsymbol{\varphi}^T(t, t_f)\mathbf{C}^T(t)$ and integrating, we obtain

$$\int_{t_o}^{t_f} \boldsymbol{\varphi}^T(t, t_f)\mathbf{C}^T(t)\mathbf{z}(t)\, dt = \left[\int_{t_o}^{t_f} \boldsymbol{\varphi}^T(t, t_f)\mathbf{C}^T(t)\mathbf{C}(t)\boldsymbol{\varphi}(t, t_f)\, dt\right]\hat{\mathbf{x}}(t_f) \tag{11.1-29}$$

Thus

$$\hat{\mathbf{x}}(t_f) = \mathbf{N}^{-1}(t_o, t_f)\int_{t_o}^{t_f} \boldsymbol{\varphi}^T(t, t_f)\mathbf{C}^T(t)\mathbf{z}(t)\, dt \tag{11.1-30}$$

where $\mathbf{N}(t_o, t_f)$ is as defined before:

$$\mathbf{N}(t_o, t_f) = \int_{t_o}^{t_f} \boldsymbol{\varphi}^T(t, t_f)\mathbf{C}^T(t)\mathbf{C}(t)\boldsymbol{\varphi}(t, t_f)\, dt \tag{11.1-31}$$

We can clearly solve for $\hat{\mathbf{x}}(t_o)$ also by

$$\hat{\mathbf{x}}(t_o) = \mathbf{M}^{-1}(t_o, t_f)\int_{t_o}^{t_f} \boldsymbol{\varphi}^T(t, t_o)\mathbf{C}^T(t)\mathbf{z}(t)\, dt \tag{11.1-32}$$

where

$$\mathbf{M}(t_o, t_f) = \int_{t_o}^{t_f} \boldsymbol{\varphi}^T(t, t_o)\mathbf{C}^T(t)\mathbf{C}(t)\boldsymbol{\varphi}(t, t_o)\, dt \tag{11.1-33}$$

and we can easily show that

$$\mathbf{M}(t_o, t_f) = \boldsymbol{\varphi}^T(t_f, t_o)\mathbf{N}(t_o, t_f)\boldsymbol{\varphi}(t_f, t_o) \tag{11.1-34}$$

From Eq. (11.1-28) we see that a necessary condition for the system to be observable (on the interval $[t_o, t_f]$) is that the columns of $\mathbf{C}(t)\boldsymbol{\varphi}(t, t_f)$ be linearly independent. Mathematically, we may write this condition of linear independence in terms of an m vector $\boldsymbol{\eta}$ as [15, 16]

$C(t)\,\phi(t,$

$$\boldsymbol{\eta}^T\mathbf{C}(t)\boldsymbol{\varphi}(t, t_f) \neq \mathbf{0}^T, \quad \forall t \in [t_o, t_f], \qquad \boldsymbol{\eta} \neq \mathbf{0} \tag{11.1-35}$$

This condition may be developed into a test for observability as follows. If we assume that the conditions of Eq. (11.1-35) are not fulfilled, and differentiate Eq. (11.1-35) repeatedly, noting that $\partial\boldsymbol{\varphi}(t, t_f)/\partial t = \mathbf{A}(t)\boldsymbol{\varphi}(t, t_f)$, we obtain the set of equations

$$\boldsymbol{\eta}^T\boldsymbol{\Gamma}_j^T(t)\boldsymbol{\varphi}(t, t_f) = \mathbf{0}^T, \qquad j = 1, 2, \ldots, n \tag{11.1-36}$$

where

$$\begin{aligned} \boldsymbol{\Gamma}_1 &= \mathbf{C}^T(t) \\ \boldsymbol{\Gamma}_k &= \frac{\partial \boldsymbol{\Gamma}_{k-1}}{\partial t} + \mathbf{A}^T(t)\boldsymbol{\Gamma}_{k-1} \end{aligned} \tag{11.1-37}$$

Now if we define

$$\boldsymbol{\Gamma} = [\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, \ldots, \boldsymbol{\Gamma}_n] \tag{11.1-38}$$

$\mu = n\eta =$

we see that for $n\eta$ vectors which we call ν we have

$$\nu^T\Gamma^T(t)\varphi(t, t_f) = 0^T \qquad (11.1\text{-}39)$$

which, since φ is always nonsingular, implies that Γ is singular. But Eq. (11.1-35) does not express an equality, so none of these relations, Eq. (11.1-36), could hold, and Γ cannot be singular if the system is observable. Thus, if the Γ matrix of Eq. (11.1-38) is of rank n, where Γ_j is defined in Eq. (11.1-37), the system is observable.

The matrices $M(t_o, t_f)$ and $N(t_o, t_f)$ are known as Gramian matrices and must be positive-definite for an observable system. This is an alternate and equivalent criterion to requiring the Γ matrix to be of rank n. For a constant system, it is considerably simpler to determine the rank of the Γ matrix than to evaluate either of the Gramian matrices. Thus for a constant system, the easiest criterion for observability is to use the requirement that the $n \times nm$ matrix

$$\Gamma = [C^T \,\vdots\, A^TC^T \,\vdots\, A^{T2}C^T \,\vdots\, \cdots \,\vdots\, A^{Tn-1}C^T] \qquad (11.1\text{-}40)$$

be of rank n. This may be accomplished if we determine whether the $n \times n$ matrix $\Gamma\Gamma^T$ is of rank n.

We may now distinguish between several types of observability. A system is said to be observable on the interval $[t_o, t_f]$ if, for a specified t_o and specified t_f, every state $x(t_o)$ may be determined from knowledge of $z(t)\ \forall\ t \in [t_o, t_f]$. In other words, the M matrix is positive-definite or the rank test is satisfied for the fixed t_o and fixed t_f. If this is true for all t_o and some $t_f > t_o$, we say that the system is completely observable. If this is true for every t_o and every $t_f > t_o$, the system is said to be totally observable. The only modification to this statement needed to treat discrete systems is that there are a finite number of states, as discussed in Section 11.1-1, before a discrete system will become observable. Finally, we remark that application of the state estimation techniques of the previous two chapters to unobservable systems often leads to impossible computational problems in determining the solution to the error variance equation. A remedy is to attempt to estimate only those components of the state vector which are observable in the output vector.

11.2 Controllability in linear systems

In Chapters 9 and 10, we saw that the linear state estimation and the regulator problem were duals of one another. Thus it is reasonable to expect a dual of the observability criterion, and we shall call it the controllablility criterion. We will say that a system is state controllable if any initial state vector $x(t_o)$ can be transferred to any final state vector $x(t_f)$, where t_o and

t_f are fixed by means of some control $u(t)$.† More precise definitions of co trollability, as well as a discussion of the implications of duality, will be giv at the end of this section. We shall first consider state controllability a output controllability for continuous systems. The close similarity of t results will then be noted. As suits the dual to observability, we shall initia our approach by considering the transfer of the system from the initial sta to a final state which, since linear systems are being considered, can be co sidered to be the origin without loss of generality.

Suppose we wish to determine whether the system described by

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \qquad (11.2\text{-}$$

$$z(t) = C(t)x(t) \qquad (11.2\text{-}$$

is controllable. In other words, we wish to find whether there is a contr $u(t)$, such that $x(t_o) = x_o$ and $x(t_f) = 0$. We will find the control which accor plishes this (if it exists) and which minimizes the cost function

$$J = \tfrac{1}{2}\int_{t_o}^{t_f} \|u(t)\|^2_{R(t)}\, dt \qquad (11.2\text{-}$$

We will use this cost function to "get a handle" on the problem, i. to determine if there is a $u(t)$ such that we can bring the system from $x(t_o) =$ to $x(t_f) = 0$. Another "sensible" cost function would work equally we To do this, we shall use the maximum principle. Thus, we form the Ham tonian

$$H[x(t), u(t), \lambda(t), t] = \tfrac{1}{2}\|u(t)\|^2_{R(t)} + \lambda^T(t)[A(t)x(t) + B(t)u(t)] \qquad (11.2\text{-}$$

and obtain in the usual way

$$\frac{\partial H}{\partial \lambda} = \dot{x} = A(t)x(t) + B(t)u(t), \qquad x(t_o) = x_o \qquad (11.2\text{-}$$

$$\frac{\partial H}{\partial x} = -\dot{\lambda} = A^T\lambda(t), \qquad x(t_f) = 0 \qquad (11.2\text{-}$$

To obtain the minimum H, we set

$$\frac{\partial H}{\partial u} = 0 \qquad (11.2\text{-}$$

which gives

$$u(t) = -R^{-1}(t)B^T(t)\lambda(t) \qquad (11.2\text{-}$$

By combining these last four equations, we obtain

$$\dot{x} = A(t)x(t) - B(t)R^{-1}(t)B^T(t)\lambda(t), \qquad x(t_o) = x_o \qquad (11.2\text{-}$$

$$\dot{\lambda} = -A^T(t)\lambda(t), \qquad x(t_f) = 0 \qquad (11.2\text{-}1$$

---

†In a similar way, a system will be called output controllable if there exists an in $u(t)$ which transfers an initial output vector $z(t_o)$ to any final output vector $z(t_f)$.

In a fashion similar to that which we have used many times before, we obtain the solution to these two equations as

$$\mathbf{x}(t_f) = \varphi(t_f, t_o)\mathbf{x}(t_o) - \int_{t_o}^{t_f} \varphi(t_f, \tau)\mathbf{B}(\tau)\mathbf{R}^{-1}(\tau)\mathbf{B}^T(\tau)\boldsymbol{\lambda}(\tau)\, d\tau \qquad (11.2\text{-}11)$$

$$\boldsymbol{\lambda}(t) = \varphi^T(t_f, t)\boldsymbol{\lambda}(t_f) \qquad (11.2\text{-}12)$$

By combining Eqs. (11.2-11) and (11.2-12), we obtain

$$\mathbf{x}(t_f) = \varphi(t_f, t_o)\mathbf{x}(t_o) - \int_{t_o}^{t_f} \varphi(t_f, \tau)\mathbf{B}(\tau)\mathbf{R}^{-1}(\tau)\mathbf{B}^T(\tau)\varphi^T(t_f, \tau)\boldsymbol{\lambda}(t_f)\, d\tau \quad (11.2\text{-}13)$$

which must be zero. An alternate approach is to write

$$\mathbf{x}(t) = \varphi(t, t_f)\mathbf{x}(t_f) + \int_{t_f}^{t} \varphi(t, \tau)\mathbf{B}(\tau)\mathbf{R}^{-1}(\tau)\mathbf{B}^T(\tau)\boldsymbol{\lambda}(\tau)\, d\tau \qquad (11.2\text{-}14)$$

which, since $\mathbf{x}(t_f) = \mathbf{0}$ becomes just

$$\mathbf{x}(t) = -\int_{t_f}^{t_o} \varphi(t, \tau)\mathbf{B}(\tau)\mathbf{R}^{-1}(\tau)\mathbf{B}^T(\tau)\boldsymbol{\lambda}(\tau)\, d\tau \qquad (11.2\text{-}15)$$

But, since

$$\boldsymbol{\lambda}(t) = \varphi^T(t_o, t)\boldsymbol{\lambda}(t_o) \qquad (11.2\text{-}16)$$

Eq. (11.2-15) can be written, if we choose $t = t_o$, as

$$\mathbf{x}(t_o) = \int_{t_f}^{t_o} \varphi(t_o, \tau)\mathbf{B}(\tau)\mathbf{R}^{-1}(\tau)\mathbf{B}^T(\tau)\varphi^T(t_o, \tau)\boldsymbol{\lambda}(t_o)\, d\tau \qquad (11.2\text{-}17)$$

Now we can solve either Eq. (11.2-13) for $\boldsymbol{\lambda}(t_f)$ or Eq. (11.2-17) for $\boldsymbol{\lambda}(t_o)$. Suppose we choose the latter. Then

$$\boldsymbol{\lambda}(t_o) = -\mathbf{W}^{-1}(t_o, t_f)\mathbf{x}(t_o) \qquad (11.2\text{-}18)$$

where

$$\mathbf{W}(t_o, t_f) = \int_{t_o}^{t_f} \varphi(t_o, \tau)\mathbf{B}(\tau)\mathbf{R}^{-1}(\tau)\mathbf{B}^T(\tau)\varphi^T(t_o, \tau)\, d\tau \qquad (11.2\text{-}19)$$

If a system is state controllable, $\mathbf{W}(t_o, t_f)$ must have an inverse and also be positive-definite as the second variation would show. Again $\mathbf{R}$ may be set equal to the identity matrix. In Section 9.2, we had a relation very similar to Eq. (11.2-19), which we converted to a differential equation. We found that it was very much easier to solve the differential equation than to evaluate the integral. Let us now try the same approach here. Differentiation of Eq. (11.2-19) gives

$$\begin{aligned}\frac{\partial \mathbf{W}(t_o, t_f)}{\partial t_o} = &-\varphi(t_o, t_o)\mathbf{B}(t_o)\mathbf{R}^{-1}(t_o)\mathbf{B}^T(t_o)\varphi^T(t_o, t_o)\\ &+ \int_{t_o}^{t_f} \frac{\partial \varphi(t_o, \tau)}{\partial t_o}\mathbf{B}(\tau)\mathbf{R}^{-1}(\tau)\mathbf{B}^T(\tau)\varphi^T(t_o, \tau)\, d\tau \qquad (11.2\text{-}20)\\ &+ \int_{t_o}^{t_f} \varphi(t_o, \tau)\mathbf{B}(\tau)\mathbf{R}^{-1}(\tau)\mathbf{B}^T(\tau)\frac{\partial \varphi^T(t_o, \tau)}{\partial t_o}\, d\tau\end{aligned}$$

---

which becomes, since $\partial\varphi(t, t_o)/\partial t = \mathbf{A}(t)\varphi(t, t_o)$ and $\varphi(t, t) = \mathbf{I}$,

$$\begin{aligned}\frac{\partial \mathbf{W}(t_o, t_f)}{\partial t_o} = &-\mathbf{B}(t_o)\mathbf{R}^{-1}(t_o)\mathbf{B}^T(t_o)\\ &+ \mathbf{A}(t_o)\int_{t_o}^{t_f} \varphi(t_o, \tau)\mathbf{B}(\tau)\mathbf{R}^{-1}(\tau)\mathbf{B}^T(\tau)\varphi^T(t_o, \tau)\, d\tau \qquad (11.2\text{-}2\\ &+ \int_{t_o}^{t_f} \varphi(t_o, \tau)\mathbf{B}(\tau)\mathbf{R}^{-1}(\tau)\mathbf{B}^T(\tau)\varphi^T(t_o, \tau)\mathbf{A}^T(t_o)\, d\tau\end{aligned}$$

But, by Eq. (11.2-19), the two integrals are just $\mathbf{W}(t_o, t_f)$. Therefore,

$$\begin{aligned}\frac{\partial \mathbf{W}(t_o, t_f)}{\partial t_o} = &-\mathbf{B}(t_o)\mathbf{R}^{-1}(t_o)\mathbf{B}^T(t_o) + \mathbf{A}(t_o)\mathbf{W}(t_o, t_f) \qquad (11.2\text{-}2\\ &+ \mathbf{W}(t_o, t_f)\mathbf{A}^T(t_o), \qquad \mathbf{W}(t_f, t_f) = \mathbf{0}\end{aligned}$$

We have, therefore, succeeded in obtaining a differential equation for $\mathbf{W}(t_o, t$ which should be easier to solve than the defining relation for $\mathbf{W}(t_o, t_f)$ Eq. (11.2-19).

It is interesting now to evaluate the cost function of Eq. (11.2-3) whic by Eq. (11.2-8), becomes

$$J = \tfrac{1}{2}\int_{t_o}^{t_f} \boldsymbol{\lambda}^T(t)\mathbf{B}(t)\mathbf{R}^{-T}(t)\mathbf{R}(t)\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\boldsymbol{\lambda}(t)\, dt \qquad (11.2\text{-}2$$

But $\mathbf{R}(t)$ is symmetric, so that

$$J = \tfrac{1}{2}\int_{t_o}^{t_f} \boldsymbol{\lambda}^T(t)\mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\boldsymbol{\lambda}(t)\, dt \qquad (11.2\text{-}2$$

From Eqs. (11.2-16) and (11.2-18), we see that

$$\boldsymbol{\lambda}(t) = \varphi^T(t_o, t)\boldsymbol{\lambda}(t_o) = -\varphi^T(t_o, t)\mathbf{W}^{-1}(t_o, t_f)\mathbf{x}(t_o) \qquad (11.2\text{-}2$$

From the defining relation for $\mathbf{W}(t_o, t_f)$, we know that it is symmetric; hen Eq. (11.2-24) becomes

$$J = \tfrac{1}{2}\int_{t_o}^{t_f} \mathbf{x}^T(t_o)\mathbf{W}^{-1}(t_o, t_f)\varphi(t_o, t)\mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\varphi^T(t_o, t)\mathbf{W}^{-1}(t_o, t_f)\mathbf{x}(t_o)\, dt \qquad (11.2\text{-}2$$

By excluding those terms from the integral which do not involve t, we s that

$$J = \tfrac{1}{2}\left\{\mathbf{x}^T(t_o)\mathbf{W}^{-1}(t_o, t_f)\left[\int_{t_o}^{t_f} \varphi(t_o, t)\mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\varphi^T(t_o, t)\, dt\right]\mathbf{W}^{-1}(t_o, t_f)\mathbf{x}(t\right. \qquad (11.2\text{-}2$$

Or, since the integral in the brackets is just the definition of $\mathbf{W}(t_o, t_f)$, we ha finally

$$J = \tfrac{1}{2}\mathbf{x}^T(t_o)\mathbf{W}^{-1}(t_o, t_f)\mathbf{x}(t_o) = \tfrac{1}{2}\|\mathbf{x}(t_o)\|^2_{\mathbf{W}^{-1}(t_o, t_f)} \qquad (11.2\text{-}2$$

Equation (11.2-28) allows an interesting interpretation of controllabili Suppose that we are given some definite value for the cost J. Then, if can determine $\mathbf{W}^{-1}(t_o, t_f)$, we can find all initial conditions such that I

(11.2-28) is satisfied. We can thus plot a surface in n-space representing those initial conditions from which we can take the system to the origin with a cost of J. This problem is known as the reachable zone problem, which is considered in Problem 4 of this chapter.

We can offer an alternative approach to this problem. We shall do this now for the output controllability problem which reduces to the state controllability problem when $\mathbf{C}(t) = \mathbf{I}$. The solution to Eqs. (11.2-1) and (11.2-2) is the m vector output due to the r vector control

$$\mathbf{z}(t) - \mathbf{C}(t)\boldsymbol{\varphi}(t, t_o)\mathbf{x}(t_o) = \mathbf{C}(t) \int_{t_o}^{t} \boldsymbol{\varphi}(t, \tau)\mathbf{B}(\tau)\mathbf{u}(\tau)\, d\tau \tag{11.2-29}$$

At time t_f, the left-hand side of this equation is simply equal to some specified value $\mathbf{z}_d(t_f)$ such that we may write

$$\mathbf{z}_d(t_f) = \mathbf{z}(t_f) - \mathbf{C}(t_f)\boldsymbol{\varphi}(t_f, t_o)\mathbf{x}(t_o) = \int_{t_o}^{t_f} \mathbf{C}(t_f)\boldsymbol{\varphi}(t_f, \tau)\mathbf{B}(\tau)\mathbf{u}(\tau)\, d\tau \tag{11.2-30}$$

A sufficient condition for output controllability on $[t_o, t_f]$ is that the columns of $\mathbf{C}(t_f)\boldsymbol{\varphi}(t_f, \tau)\mathbf{B}(\tau)$ be linearly independent, which means that, for arbitrary m vector $\boldsymbol{\eta}$, we have the r vector equation [15, 16]

$$\boldsymbol{\eta}^T\mathbf{C}(t_f)\boldsymbol{\varphi}(t_f, \tau)\mathbf{B}(\tau) \neq \mathbf{0}^T, \qquad t_o \leq \tau \leq t_f \tag{11.2-31}$$

We may develop another output controllability condition from this condition. This proof will proceed by the method of contradiction. Suppose that there exists at least one nonzero vector $\boldsymbol{\eta}$, such that Eq. (11.2-31) is, in fact, true. Repeated differentiation of Eq. (11.2-31) with respect to τ yields

$$\boldsymbol{\eta}^T\mathbf{C}(t_f)\boldsymbol{\varphi}(t_f, \tau)\boldsymbol{\Gamma}_j(\tau) = \mathbf{0}^T, \qquad j = 1, 2, \ldots, n \tag{11.2-32}$$

where, since $\partial\boldsymbol{\varphi}(t_f, \tau)/\partial\tau = -\boldsymbol{\varphi}(t_f, \tau)\mathbf{A}(\tau)$,

$$\begin{aligned} \boldsymbol{\Gamma}_1(\tau) &= \mathbf{B}(\tau) \\ \boldsymbol{\Gamma}_k(\tau) &= \frac{\partial\boldsymbol{\Gamma}_{k-1}(\tau)}{\partial\tau} - \mathbf{A}(\tau)\boldsymbol{\Gamma}_{k-1}(\tau) \end{aligned} \tag{11.2-33}$$

Then, if we define the n by nm matrix $\boldsymbol{\Gamma}$

$$\boldsymbol{\Gamma} = [\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, \ldots, \boldsymbol{\Gamma}_n] \tag{11.2-34}$$

the condition of Eq. (11.2-32) becomes, for the $n\eta$ vectors $\mathcal{N}$,

$$\mathcal{N}^T\mathbf{C}(t_f)\boldsymbol{\varphi}(t_f, \tau)\boldsymbol{\Gamma} = \mathbf{0}^T \tag{11.2-35}$$

which would tell us that $\boldsymbol{\Gamma}$ could not be of rank n since $\boldsymbol{\varphi}$ is nonsingular (excluding for the moment the possibility of $\mathbf{C}$ being singular). But Eq. (11.2-35) cannot be zero by Eq. (11.2-31), and so $\boldsymbol{\Gamma}$ must then be of rank n, and Eq. (11.2-35) will not, in fact, be zero. Although this requirement holds for time-varying systems, it is particularly easy to apply in the case of constant systems, for then, as is easily verified, for $\boldsymbol{\Gamma}' = [\boldsymbol{\Gamma}_1, -\boldsymbol{\Gamma}_2, \ldots (-1)^{n+1}\boldsymbol{\Gamma}_n]$,

$$\boldsymbol{\Gamma}' = [\mathbf{B} \mid \mathbf{AB} \mid \mathbf{A}^2\mathbf{B} \mid \ldots \mid \mathbf{A}^{n-1}\mathbf{B}] \qquad (11.$$

and this must be of rank n. This is only the requirement for state con lability since, if a constant system is controllable at all, it is control at $t_f = t_o$ (impulse control required). Therefore, from Eq. (11.2-35), the put controllability requirement is that

$$[\mathbf{CB} \mid \mathbf{CAB} \mid \mathbf{CA}^2\mathbf{B} \mid \ldots \mid \mathbf{CA}^{n-1}\mathbf{B}] \qquad (11.$$

be of rank m. For the general time-varying case, the $\mathbf{C}(t_f)\boldsymbol{\Gamma}$ term o (11.2-35) must be of rank m since we know that $\boldsymbol{\varphi}$ must be nonsingul

If, in Eq. (11.2-30), we let

$$\mathbf{u}(t) = \mathbf{B}^T(t)\boldsymbol{\varphi}^T(t_f, t)\mathbf{C}^T(t_f)\boldsymbol{\lambda}(t_f) \qquad (11.$$

we have

$$\boldsymbol{\lambda}(t_f) = -\mathbf{V}^{-1}(t_o, t_f)\mathbf{z}_d(t_f) \qquad (11.$$

where

$$\mathbf{V}(t_o, t_f) = \int_{t_o}^{t_f} \mathbf{C}(t_f)\boldsymbol{\varphi}(t_f, \tau)\mathbf{B}(\tau)\mathbf{B}^T(\tau)\boldsymbol{\varphi}^T(t_f, t)\mathbf{C}^T(t_f)\, d\tau \qquad (11.$$

and must be positive-definite for a controllable system. For state co lability, we may treat $\mathbf{C} = \mathbf{I}$; then we can easily show that

$$\mathbf{V}(t_o, t_f) = \boldsymbol{\varphi}(t_f, t_o)\mathbf{W}(t_o, t_f)\boldsymbol{\varphi}^T(t_f, t_o) \qquad (11.$$

where $\mathbf{W}(t_o, t_f)$ is defined by Eq. (11.2-19).

It is quite easy for us to show that all of these results carry over ex to the discrete system described by

$$\mathbf{x}(k + 1) = \mathbf{A}(k)\mathbf{x}(k) + \mathbf{B}(k)\mathbf{u}(k) \qquad (11.$$

$$\mathbf{z}(k) = \mathbf{C}(k)\mathbf{x}(k) \qquad (11.$$

except that discrete transition matrices and summations are used r than continuous transition matrices and integrations. The time interval [is then replaced by the sequence $k_o, k_o + 1, \ldots, k_f$. Thus, for instanc discrete equivalent of Eq. (11.2-19) is

$$\mathbf{W}(k_o, k_f) = \sum_{k=k_o}^{k=k_f} \boldsymbol{\varphi}(k_o, k)\mathbf{B}(k)\mathbf{R}^{-1}(k)\mathbf{B}^T(k)\boldsymbol{\varphi}^T(k_o, k) \qquad (11.$$

Analogous to the discrete observability requirement, a controllable di system can be transferred to the origin in at most n stages, where $\mathbf{x}$ is vector.

Just as in the case of observability, there are several different ty controllability. We will give these definitions for the case of state co lability. Output controllability definitions follow merely by replac of $\mathbf{x}(t_f)$ by $\mathbf{z}(t_f)$ in the definitions.

We will say that a system is state controllable for a given t_o and t_f i initial state $\mathbf{x}(t_o)$ can be transferred to any final state $\mathbf{x}(t_f)$ using any c $\mathbf{u}(t)$ over the interval $[t_o, t_f]$. A system will be said to be completely stat trollable if, for any t_o, each initial state $\mathbf{x}(t_o)$ can be transferred to any final state and given final time $\mathbf{x}(t_f)$ where, of course, $t_f \geq t_o$. To obtain total state controllability, the system must be completely state controllable for every t_o and every t_f.

Example 11.2-1

Let us consider the linear system described by

$$\dot{x}_1 = x_2(t) + u(t), \qquad z_1(t) = x_1(t)$$
$$\dot{x}_2 = -x_1(t) - 2x_2(t) - u(t), \quad z_2(t) = x_1(t) + x_2(t)$$

The system dynamics can also be written as

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t), \qquad \mathbf{z}(t) = \mathbf{C}\mathbf{x}(t)$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

We wish to determine the observability and controllability of the system. From the preceding section we know that the system is observable if the $n \times nm$ matrix

$$[\mathbf{C}^T \,|\, \mathbf{A}^T\mathbf{C}^T \,|\, \cdots \,|\, \mathbf{A}^{T^{n-1}}\mathbf{C}^T] = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 1 & -1 \end{bmatrix}$$

is of rank 2. This is the case, and so the system is observable. To discern state controllability, we must examine the matrix

$$[\mathbf{B} \,|\, \mathbf{AB} \,|\, \mathbf{A}^2\mathbf{B} \,|\, \cdots \,|\, \mathbf{A}^{n-1}\mathbf{B}] = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

to see if it is of rank 2. Clearly it is not, and so this system is not state controllable. Neither is the system output controllable, because the matrix

$$[\mathbf{CB} \,|\, \mathbf{CAB} \,|\, \mathbf{CA}^2\mathbf{B} \,|\, \cdots \,|\, \mathbf{CA}^{n-1}\mathbf{B}] = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$$

is not of rank 2.

Let us now examine the reasons for this uncontrollability. Figure 11.2-1 illustrates a possible block diagram for this system. Appropriate transfer functions for the system are

$$\frac{x_1(s)}{u(s)} = \frac{1}{s+1}, \qquad \frac{x_2(s)}{u(s)} = \frac{-1}{s+1}$$

and we observe that the physical reason the system is not state controllable is that the state vector $\mathbf{x}(t)$ can be controlled only along or parallel to a straight line $x_1(t) + x_2(t) = 0$. This is certainly not in two dimensions; therefore the system is not state controllable. Appropriate transfer functions for the output state are

$$\frac{z_1(s)}{u(s)} = \frac{x_1(s)}{u(s)} = \frac{1}{s+1}, \qquad \frac{z_2(s)}{u(s)} = \frac{x_1(s) + x_2(s)}{u(s)} = 0$$

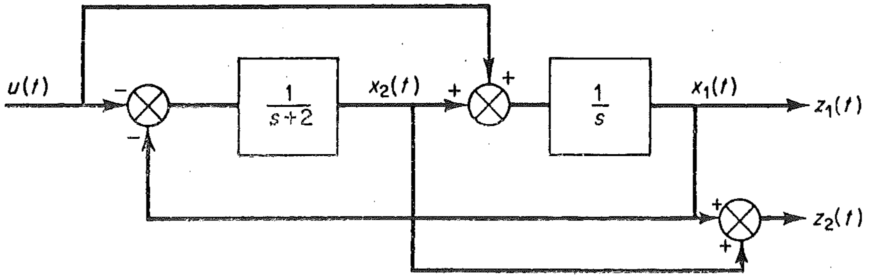


Fig. 11.2-1 Block diagram of uncontrollable system, Example (11.2-1).

Since the output $z_2(t)$ cannot be controlled by the input, the entire system is not output controllable. If the output were just $z_1(t)$, a scalar, then the system is not state controllable but is output controllable. This means that we could determine an input which could drive $z_1(t)$ to any given value but could not drive $x_1(t)$ and $x_2(t)$ to any value which lies off the line $x_1(t) + x_2(t) = 0$. We note that we were given a second order system but found first order transfer function from control input to state and output state variables. This implies that that the given system is "reducible" in order. Choate and Sage [16] have shown that systems which are not totally controllable must be reducible.

Earlier we remarked that the dual of an unobservable system is an uncontrollable system. This can easily be seen if we observe the observability criteria where the adjoint system ($\mathbf{A}^* = -\mathbf{A}^T$, $\mathbf{B}^* = \mathbf{C}^T$, $\mathbf{C}^* = \mathbf{B}^T$) is used and if we note that the observability criteria becomes the controllability criteria. Thus we may say that a system is controllable if the adjoint system is observable. Since the dual system is defined by $\mathbf{A}^*(t^*) = \mathbf{A}^T(t)$, $\mathbf{B}^*(t^*) = \mathbf{C}^T(t)$, $\mathbf{C}^*(t^*) = \mathbf{B}^T(t)$, $t^* = -t$, we see that the similar statement for dual systems, a system is uncontrollable (unobservable) if its dual is unobservable (uncontrollable), applies.

For successful control, it is normally necessary that systems be both controllable and observable. For example, if a subsystem which is unobservable is part of a closed-loop system, instabilities in the unobservable part of the system cannot be detected or stabilized by the closed loop. If a system is not state controllable, it is not possible to control a portion of the system, and thus persistent transients may exist. If the system is not output controllable, then it appears that all is lost unless it is possible to change input and/or output state variables.

Even though a system may be observable, not all components of the state variable, $\mathbf{x}(t)$, may be recoverable immediately from the observation $\mathbf{z}(t)$. We recall that $\mathbf{z}(t)$ may well be a scalar, $\mathbf{x}(t)$ may well be a 100 vector, and the system may certainly be observable. In the next section we shall discuss methods of state-variable recovery from observable output vectors.

THE VARIATIONAL APPROACH TO OPTIMAL CONTROL PROBLEMS

I. NECESSARY CONDITIONS FOR OPTIMAL CONTROL

GENERAL PROBLEM IS TO MINIMIZE

$$J(u) = h(x(t_f), t_f) + \int_{t_0}^{t_f} g[x(t), u(t), t]\,dt$$

SUBJECT TO $\dot{x}(t) = a[x(t), u(t), t]$

$x \rightarrow n \times 1$ VECTOR $\qquad$ $u \rightarrow m \times 1$ VECTOR

NOW

$$h[x(t_f, t)] = \int_{t_0}^{t_f} \frac{d}{dt} h[x(t_f, t)]\,dt + h[x(t_0), t_0]$$

$$\Rightarrow J(u) = \int_{t_0}^{t_f} \left[g[x, u, t] + \frac{d}{dt} h(x, t)\right] dt + h[x(t_0), t_0]$$

SINCE $h[x(t_0) t_0]$ IS CONSTANT, WE CAN DROP IT:

$$J(u) = \int_{t_0}^{t_f} \left[g[x, u, t] + \frac{d}{dt} h(x, t)\right] dt$$

BUT

$$\frac{d}{dt} h(x, t) = \left[\frac{\delta h}{\delta x}\right]^T \dot{x} + \frac{\delta h}{\delta t}$$

THUS

$$J(u) = \int_{t_0}^{t_f} \left[g + \frac{\delta h^T}{\delta x} \dot{x} + \frac{\delta h}{\delta t}\right] dt$$

TO INCLUDE RESTRAINTS, WE WRITE THE AUGMENTED J AS

$$J_a(u) = \int_{t_0}^{t_f} \left[g + \frac{\delta h}{\delta x} \dot{x} + \frac{\delta h}{\delta t} + p^T(t)\left[a(x, u, t) - \dot{x}\right]\right] dt$$

WHERE $p(t)$ ARE THE (n) LAGRANGE MULTIPLIERS.

DEFINE AUGMENTED g:

$$g_a[x,\dot{x},u,p,t] = g[x,u,t] + p^T(t)\,[a(x,u,t) - \dot{x}] + \frac{\delta h^T}{\delta x}\dot{x} + \frac{\delta h}{\delta t}$$

OUR PROBLEM NOW BECOMES ONE OF MINIMIZING

$$J_a(u) = \int_{t_0}^{t_f} g_a \, dt$$

RECOGNIZING THAT g_a IS INDEPENDENT OF $\dot{u}$ AND $\dot{p}$, AND PERFORMING THE STANDARD PARTS INTEGRATION TRICK GIVES THE VARIATION

$$\delta J_a = 0 = \frac{\delta g_a^T}{\delta \dot{x}}\delta x\Big|_{t_f} + \left[g_a - \frac{\delta g_a^T}{\delta \dot{x}}\dot{x}\right]\Big|_{t_f}\delta t_f + \int_{t_0}^{t_f}\left\{\left[\frac{\delta g_a^T}{\delta x} - \frac{d}{dt}\frac{\delta g_a^T}{\delta \dot{x}}\right]\delta x + \frac{\delta g_a^T}{\delta u}\delta u + \frac{\delta g_a^T}{\delta p}\delta p\right\}dt$$

NOTE THAT WE HAVE ALLOWED NON-SPECIFICATION OF t_f AND $x(t_f)$. THE TERMS IN THIS EXPRESSION WHICH ARE DEPENDENT ON h INSIDE INTEGRAL ARE

$$\frac{\delta}{\delta x}\left[\frac{\delta h^T}{\delta x}\dot{x} + \frac{\delta h}{\delta t}\right] - \frac{d}{dt}\frac{\delta}{\delta \dot{x}}\frac{\delta h^T}{\delta x}\dot{x}$$

OR

$$\frac{\delta^2 h}{\delta x^2}\dot{X} + \frac{\delta^2 h}{\delta x \delta t} - \frac{d}{dt}\frac{\delta h}{\delta x}$$

APPLYING CHAIN RULE TO LAST TERM:

$$\frac{\delta^2 h}{\delta x^2}\dot{x} + \cancel{\frac{\delta^2 h}{\delta x \delta t}} - \frac{\delta^2 h}{\delta x^2}\dot{x} - \cancel{\frac{\delta^2 h}{\delta x \delta t}}$$

$$= \frac{\delta^2 h}{\delta x^2}\dot{x} - \frac{\delta^2 h}{\delta x^2}\dot{x} = 0$$

THUS

$$\delta J_a = \int_{t_0}^{t_f}\left\{\left[\frac{\delta g^T}{\delta x} + p^T\frac{\delta a}{\delta x} + \frac{dp^T}{dt}\right]\delta x + \left[\frac{\delta g^T}{\delta u} + p^T\frac{\delta a}{\delta u}\right]\delta u + [a-\dot{x}]^T\delta p\right\}dt = 0$$

WE ARE READY TO MAKE OUR NECESSARY CONDITIONS WITH REGARD TO THIS EQUATION.

FIRST OFF

① $$a = \dot{x}$$

$\therefore$ THE COEFFIECENT OF δp ARE ZERO.

CHOOSE p SUCH THAT THE COEFFICIENT OF δx IS ZERO:

② $$\dot{p}(t) = -\frac{\delta a^T}{\delta x} p - \frac{\delta g}{\delta x}$$

THIS IS THE <u>COSTATE EQUATION</u>.

THE COEFFICIENT OF δu MUST LIKEWISE BE ZERO. THUS

③ $$\frac{\delta g}{\delta u} + \frac{\delta a^T}{\delta u} p = 0$$

THESE THREE EQUATIONS ARE IMPORTANT.

WE HAVE YET TO DEAL WITH THE BOUNDRY CONDITION:

④ $$\frac{\delta h^T}{\delta x} \delta x \Big|_{t_f} + \left[g + \frac{\delta h}{\delta t} + p^T a \right] \delta t_f = 0$$

MORE ON THIS LATER

WE MAY SIMPLIFY THE ABOVE EQUATIONS BY INTRODUCING THE HAMILTONIAN

$$\mathcal{H}[x,u,p,t] \triangleq g[x,u,t] + p^T(t)\, a(x,u,t)$$

OUR EQUATIONS BECOME, $\forall\ t \in [t_0, t_f]$

① $\dot{x} = \frac{\delta \mathcal{H}}{\delta p}$

② $\dot{p} = -\delta H / \delta x$

③ $0 = \delta \mathcal{H} / \delta u$

④ $\left[\frac{\delta h}{\delta x} - p\right]^T \Big|_{t_f} + \left[\mathcal{H}\Big|_{t_f} + \frac{\delta h}{\delta t}\Big|_{t_f}\right] \delta t_f = 0$

A. BOUNDRY CONDITIONS

1. FIXED t_f

a. $x(t_f)$ SPECIFIED

$$\delta x_f = 0, \quad \delta t_f = 0$$

$$\Longrightarrow x(t_f) = x_f$$

b. $x(t_f)$ FREE

$$\delta t_f = 0$$

$$\frac{\delta h}{\delta x} - p \Big|_{t_f} = 0$$

~~c. FINAL STATE ON SURFACE $m[x(t)] = 0$~~

~~EXAMPLE $m(x) = (x_1 - 3)^2 + (x_2 - 4) - 4 = 0$~~

~~(THIS IS A CIRCLE)~~

~~THE ADMISSIBLE VALUES OF $x(t_f)$ ARE (TO FIRST ORDER) NORMAL~~

~~$\frac{\delta m}{\delta x}\Big|_{t_f} = \begin{bmatrix} 2[x_1 - 3] \\ 2[x_2 - 4] \end{bmatrix}$~~

~~$\delta x(t_f)$ IS NORMAL TO GRADIANT:~~

~~$\frac{\delta m^T}{\delta x} \delta x \Big|_{t_f} = 2[x_1(t_f) - 3]\, \delta x_1(t_f) + 2[x_2(t_f) - 4]\, \delta x_2(t_f) = 0$~~

~~$\Rightarrow \delta x_2(t_f) = - \frac{[x_1 - 3]}{[x_2 - 4]} \delta x_1 \Big|_{t_f}$~~

~~SUBSTITUTING THIS INTO ④:~~

~~$\left[\frac{\delta h}{\delta x} - p \right]^T \begin{bmatrix} 1 \\ -\frac{(x_1 - 3)}{(x_2 - 4)} \end{bmatrix} \Bigg|_{t_f} = 0$~~

~~WE MUST SATISFY THIS AND~~

~~$m(x)\big|_{t_f} = 0$~~

~~NOW, LETS GENERALIZE $\Longrightarrow$~~

- CONTROLLABILITY

  FOR LINEAR TIME-INVARIANT SYSTEMS

  $[B \vdots AB \vdots A^2B \vdots \ldots \vdots A^{n-1}B]$ HAS RANK n

  $\dot{x} = Ax + BU$

- OBSERVABLE

  FOR LINEAR TIME-INVARIANT SYSTEMS

  $G = [C^T \vdots A^T C^T \vdots A^{T2} C^T \vdots \ldots \vdots A^{Tn-1} C^T]$

  HAS RANK n $Y = CX$

- PERFORMANCE MEASURES

  - MINIMUM TIME: $J = t_f - t_0$
  - TRACKING : $J = \int_{t_0}^{t_f} \| x(t) - R(t) \|_Q^2 \, dt$
  - MINIMUM CONTROL EFFORT: $J = \int_{t_0}^{t_f} \| u(t) \|_R^2 \, dt$

VARIATIONAL CALCULUS

$$J = \int_{t_0}^{t_f} \phi[x(t), \dot{x}(t), t]\,dt$$

THEN

$$\Rightarrow \quad \delta J = \frac{\delta \phi}{\delta \dot{x}} \delta x \Big|_{t_0}^{t_f} + \left[\phi - \frac{\delta \phi}{\delta \dot{x}} \dot{x}\right] \delta t \Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \left[\frac{\delta \phi}{\delta x} - \frac{d}{dt} \frac{\delta \phi}{\delta \dot{x}}\right] \delta x\, dt$$

EULER'S EQUATION:

$$\Rightarrow \quad \frac{\delta \phi}{\delta x} - \frac{d}{dt} \frac{\delta \phi}{\delta \dot{x}} = 0$$

TRANSVERSALITY CONDITION

$$\Rightarrow \quad \frac{\delta \phi}{\delta \dot{x}} \delta x \Big|_{t_0}^{t_f} + \left[\phi - \frac{\delta \phi}{\delta \dot{x}} \dot{x}\right] \delta t \Big|_{t_0}^{t_f} = 0$$

FOR TERMINAL TIME SPECIFIED: $\delta t = 0$

FOR TERMINAL STATE SPECIFIED: $\delta x = 0$

FOR $x(t_f) = \Theta(t_f)$ THEN $\frac{\delta \Theta}{\delta t} = \frac{\delta x_f}{\delta t_f}$

$$\Rightarrow \quad \frac{\delta \phi}{\delta \dot{x}}[\dot{\Theta} - \dot{x}] + \phi = 0 \quad ; t = t_f$$

EXAMPLE $\quad J = \int_{t_0}^{t_f} [1 + \dot{x}^2]^{1/2} dt$

$$\Theta(t) = -5t + 15$$

$$\frac{\delta \phi}{\delta x} - \frac{d}{dt} \frac{\delta \phi}{\delta \dot{x}} = 0 - \frac{d}{dt} \frac{\dot{x}}{\sqrt{1 + \dot{x}^2}} = 0$$

$$\Rightarrow \ddot{x} = 0 \Rightarrow x = at + b$$

$$\frac{\delta \phi}{\delta \dot{x}}[\dot{\Theta} - \dot{x}] + \phi = 0 = \frac{\dot{x}}{\sqrt{1 + \dot{x}^2}}[-5 - \dot{x}] + \sqrt{1 + \dot{x}^2} = 0$$

$$+\dot{x}[5 + \dot{x}] = 1 + \dot{x}^2$$

$$-5\dot{x} + 1 = 0 \quad ; t = t_f$$

$$\Rightarrow \dot{x} = \frac{1}{5} \Rightarrow x = \frac{1}{5} t_f + C = -5t_f + 15$$

THUS $\quad t_f = 75/26$

NOW $\quad \dot{x} = a$

$$\dot{x}(t_f = 75/26) = \frac{1}{5} = a$$

$$\Rightarrow x = \frac{1}{5} t + b$$

$$x(t_f) = \Theta(t_f)$$

$$\Rightarrow x = \frac{1}{5} t$$

- WEIERSTRASS - ERDMAN CORNER CONDITIONS

t_1 AND $x(t_1)$ UNSPECIFIED

$$\begin{cases} \delta\phi/\delta\dot{x}\big|_{t_1^-} = \delta\phi/\delta\dot{x}\big|_{t_1^+} \\ \phi - \dot{x}\,\delta\phi/\delta\dot{x}\big|_{t_1^-} = \phi - \dot{x}\,\delta\phi/\delta\dot{x}\big|_{t_1^+} \end{cases}$$

MUST ALSO SATISFY EULER LAGRANGE & TRANS. CONDITIONS

EXAMPLE $J = \int_0^t \dot{x}^2(t)\,[1-\dot{x}(t)]^2\,dt$

SOLUTION IS $x = at + b$

$$\delta\phi/\delta\dot{x} = 2\dot{x}[1-2\dot{x}][1-\dot{x}]$$

$$\phi - \delta\phi/\delta\dot{x}\,\dot{x} = \dot{x}(1-\dot{x})(3\dot{x}-1)$$

$$\dot{x} = 0, \tfrac{1}{2}, 1$$

$$\Rightarrow \dot{x} = 0, 1, \tfrac{1}{3}$$

THUS, WE REQUIRE $\dot{x} = 0, 1$

$\Rightarrow \dot{x}(t_1^-) = 0$ AND $x(t_1^+) = 1$

OR $\dot{x}(t_1^-) = 1$ AND $x(t_1^+) = 0$

EXAMPLE $J(x) = \int_0^{\pi/2} [\dot{x}^2 - x^2]\,dt$

$x(0) = 0, \quad x(\pi/2) = 1$

TURNS OUT $x = c_3 \cos t + c_4 \sin t$

$$\frac{\delta\phi}{\delta\dot{x}} = 2\dot{x}$$

$$\Rightarrow \dot{x}(t_1^+) = \dot{x}(t_1^-)$$

$$\phi - \dot{x}\,\frac{\delta\phi}{\delta\dot{x}} = \dot{x}^2 - x^2 - 2\dot{x}^2 = -\dot{x}^2 - x^2 = 0$$

WE CAN HAVE NO CORNERS SINCE DERIVATIVES ARE CONTINUOUS.

CONSTRAINED EXTREMA (POINT CONSTRAINTS)

MINIMIZE $J(w) = \int_{t_0}^{t_f} \phi[w, \dot{w}, t]dt$

SUBJECT TO $f_i[w, t] = 0, \; i = 1, \ldots, n$

$W = \left.\begin{matrix} x \\ u \end{matrix}\right] \Leftarrow n+m$ VECTOR

USE LAGRANGE MULTIPLIERS AND WRITE

$\phi_a(w, \dot{w}, \lambda, t) = \phi(w, \dot{w}, t) + \lambda^T f(w, t)$

$\Rightarrow J_a(w, p) = \int_{t_0}^{t_f} \phi_a(w, \dot{w}, \lambda, t)dt$

EULER'S EQUATION IS: $\frac{\delta\phi_a}{\delta W} - \frac{d}{dt}\frac{\delta\phi_a}{\delta\dot{x}} = 0$

<u>EXAMPLE</u> $J = \int_{t_0}^{t_f} [1 + \dot{w}_1^2 + \dot{w}_2^2]^{1/2} dt$, w_0, t_0, w_f, t_f SPECIFIED

$w_1^2 + w_2^2 + t^2 = R^2$ ①

$\phi_a = 1 + \dot{w}_1^2 + \dot{w}_2^2 + \lambda[w_1^2 + w_2^2 + t^2 - R^2]$

$\frac{\delta\phi_a}{\delta w_1} - \frac{d}{dt}\frac{\delta\phi_a}{\delta\dot{w}_1} = 0$ ② $\frac{\delta\phi_a}{\delta w_2} - \frac{d}{dt}\frac{\delta\phi}{\delta\dot{w}_2} = 0$ ③

①, ②, & ③ WILL SOLVE IT

(DIFFERENTIAL CONSTRAINTS)

$f(w, \dot{w}, t] = 0$ (USE SAME EQUATIONS)

EXAMPLE $J = \frac{1}{2}\int_0^2 (\ddot{\theta})^2 dt$

$\theta(0) = 1, \; \dot{\theta}(0) = 1, \; \theta(2) = 0, \; \dot{\theta}(2) = 0$

LET $x_1 = \theta, \; x_2 = \dot{\theta}$

$\Rightarrow \dot{x}_1 = x_2, \; \dot{x}_2 = u$

$\Rightarrow \phi_a = \frac{1}{2}u^2 + \lambda_1[\dot{x}_1 - x_2] + \lambda_2(\dot{x}_2 - u)$

LET $w = \begin{bmatrix} x \\ u \end{bmatrix}$

$\frac{\delta\phi_a}{\delta x_1} - \frac{d}{dt}\frac{\delta\phi_a}{\delta\dot{x}_1} = 0 \Rightarrow \dot{\lambda}_1 = 0 \Rightarrow \lambda_1 = C$

$\frac{\delta\phi_a}{\delta x_2} - \frac{d}{dt}\frac{\delta\phi_a}{\delta\dot{x}_2} = 0 \Rightarrow \dot{\lambda}_2 = -\lambda_1 = -C$

ETC

EXAMPLE

$$J = \frac{1}{2}\int_{t_0}^{t}\left[\|U(t)\|^2_{R(t)} + \|x(t) - r(t)\|^2_{Q(t)}\right]dt$$

$$\dot{x} = Ax + BU \qquad ; \; x(t_0) = x_0$$

$$\phi_a = \frac{1}{2}\|U(t)\|^2_{R(t)} + \frac{1}{2}\|x - r\|^2_{Q(t)} + \lambda^T(Ax + BU - \dot{x})$$

EULER'S EQUATION

$$0 = \frac{\delta\phi_a}{\delta x} - \frac{d}{dt}\frac{\delta\phi_a}{\delta\dot{x}} = Q(x - r) + A^T\lambda + \frac{\delta\lambda}{\delta t} = 0$$

$$0 = \frac{\delta\phi_a}{\delta U} - \frac{d}{dt}\frac{\delta\phi_a}{\delta\dot{U}} = RU + B^T\lambda \Rightarrow U = -R^{-1}B^T\lambda$$

TRANSVERSALITY CONDITIONS ARE

$$\delta w^T \frac{\delta\phi_a}{\delta\dot{w}}\Big|_{t_0}^{t_f} = 0$$

$$\underbrace{\delta x \quad \delta u} \left[\begin{array}{l} \delta\phi_a/\delta\dot{x} \quad = 0 \\ 0 \end{array}\right.$$

$$\Rightarrow \frac{\delta\phi_a}{\delta x}\Big|_{t_f} = -\lambda\Big|_{t_f} = -\lambda(t_f) = 0$$

- INEQUALITY CONSTRAINTS

EXTREMIZE J SUBJECT TO

$$f(w,\dot{w},t)=0 \qquad \Gamma_{MIN} \leq \Gamma(w,\dot{w},t) \leq \Gamma_{MAX}$$

$$\Rightarrow (\Gamma_{MAX}-\Gamma)(\Gamma-\Gamma_{min}) = \alpha^2$$

$$\phi_a = \phi + \lambda^T f + p^T\left[\alpha^2 - (\Gamma_{MAX}-\Gamma)(\Gamma-\Gamma_{MIN})\right]$$

- ISOPERIMETRIC CONSTRAINTS

MINIMIZE $J = \int_{t_0}^{t_f} \phi \, dt$

SUBUECT TO $\int_{t_0}^{t_f} e(w,\dot{w},t)\,dt = C$

$$z(t) = \int_{t_0}^{t} e(w,\dot{w},t)\,dt \Rightarrow \dot{z} = e$$

$$z(t_0)=0 \qquad z(t_f)=C$$

$$\Rightarrow \phi_a = \phi + \lambda^T(\dot{z}-e)$$

NOTE $\frac{\delta\phi_a}{\delta z} - \frac{d}{dt}\frac{\delta\phi_a}{\delta \dot{z}} = 0 = \lambda = CONST$

<u>EXAMPLE:</u>

$$J = \int_{t_0}^{t_f} \tfrac{1}{2}(w_1^2 + w_2^2 + 2\dot{w}_1\dot{w}_2)\,dt$$

SUBJECT TO $\int_{t_0}^{t_f} w_2^2\,dt = C$

$$\Rightarrow \dot{z} = w_2^2 \quad ; \quad z(t_0)=0, \quad z(t_f)=C$$

$$\phi_a = \tfrac{1}{2}(w_1^2 + w_2^2 + 2\dot{w}_1\dot{w}_2) + \lambda^T(\dot{z} - w_2^2)$$

$$\delta\phi_o/\delta w_1 = w_1 \qquad \delta\phi_a/\delta\dot{w}_1 = 2\dot{w}_2$$

$$\Rightarrow w_1 - 2\ddot{w}_2 = 0$$

$$\delta\phi_o/\delta w_2 = 2w_2 + 2\lambda w_2 \qquad \delta\phi_o/\delta\dot{w}_2 = 2\dot{w}_1$$

$$\Rightarrow w_2 + 2\lambda w_2 - 2\ddot{w}_1 = 0$$

$\lambda = CONST$

THE LINEAR REGULATOR

$$\begin{cases} \dot{x} = Ax + BU \qquad x(t_0) = x_0 \\ J = \frac{1}{2} x^T(t_f) S x(t_f) + \frac{1}{2} \int_{t_0}^{t_f} [x^T Q x + U^T R U] d\tau \end{cases}$$

$$H[x, u, \lambda, t] = \frac{1}{2} x^T Q x + U^T R U + \lambda^T (Ax + BU)$$

APPLY MINIMUM PRINCIPLE:

$$\frac{\delta H}{\delta U} = 0 = RU + B^T \lambda$$

$$\frac{\delta H}{\delta x} = -\dot{\lambda} = Qx + A^T \lambda$$

ALSO, WE HAVE TERMINAL CONDITION:

$$\lambda(t_f) = \left. \frac{\delta \theta}{\delta x} \right|_{t_f} = S x(t_f)$$

$$U = -R^{-1} B^T \lambda$$

ASSUME THAT $\lambda(t) = P(t) x(t)$

THEN $\dot{x} = Ax - BR^{-1} B^T P x$

AND $\dot{\lambda} = \dot{P} x + P \dot{x} = -Qx - A^T P x$

COMBINING THESE

$$[\dot{P} + PA + A^T P - PBR^{-1} B^T P + Q] x(t) = 0$$

THE COEFFICIENT MUST BE ZERO,

THIS GIVES US THE RICCATI EQ:

$$\dot{P} = -PA - A^T P + PBR^{-1} B^T P - Q$$

WHERE $P(t_f) = S$ (THUS, WE MUST INTEGRATE BACKWARDS)

ASSUME A FEEDBACK RELATION:

$$u(t) = K(t)\,x(t)$$

K(t) IS THE KALMAN GAIN. NOW

$$P P^{-1} = I$$

$\Rightarrow$ DIFFERENTIATING:

$$\dot{P} P^{-1} + P \dot{P}^{-1} = 0 \Rightarrow -\dot{P}^{-1} = P^{-1} \dot{P} P^{-1}$$

USING THIS ON OUR ORIGINAL EQUATION:

$$\dot{P}^{-1} = A P^{-1} + P^{-1} A - B R^{-1} B^{T} + P^{-1} Q P^{-1}$$

WITH $P^{-1}(t_f) = S^{-1}$

EXAMPLE

$$\dot{x} = -\frac{1}{2} x + u$$

$$J = \frac{1}{2} S x^2(t_f) + \frac{1}{2} \int_{t_0}^{t_f} [2x^2 + u^2]\, dt$$

$$\dot{P} = -PA - A^T P - P B R^{-1} B^T P - Q$$

$$= +\frac{1}{2} P + \frac{1}{2} P - P(1)(1)^{-1}(1) P - 2 = P + P^2 - 2$$

THE BOLZA PROBLEM

A. FIXED BEGINNING AND TERMINAL TIMES

MINIMIZE $J = \Theta[x(t), t]_{t_0}^{t_f} + \int_{t_0}^{t_f} \phi(x, u, t)\, dt$

SUBJECT TO $\dot{x} = f(x, u, t) \Leftarrow n$ VECTOR

$M(t_0)\, x(t_0) = m_0 \Leftarrow r$ VECTOR

$N(t_f)\, x(t_f) = n_f \Leftarrow q$ VECTOR

$$J_a = \Theta[x, t]\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \left\{ \phi(x, u, t) + \lambda^T [f(x, u, t) - \dot{x}] \right\} dt$$

DEFINE THE HAMILTONIAN:

$$H(x, u, \lambda, t) = \phi(x, u, t) + \lambda^T f(x, u, t)$$

$$\Rightarrow J_a = \Theta[x, t]_{t_0}^{t_f} + \int_{t_0}^{t_f} [H(x, u, \lambda, t) - \lambda^T \dot{x}]\, dt$$

INTEGRATING BY PARTS:

$$J_a = \Theta[x, t]\Big|_{t_0}^{t_f} - \lambda^T x \Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} [H(x, u, \lambda, t) + \dot{\lambda}^T x]\, dt$$

PERFORM VARIATION IN J_a:

$$\delta J_a = \left\{ \delta x^T \left[\frac{\delta \Theta}{\delta x} - \lambda \right] \right\}\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \left\{ \delta x^T \left[\frac{\delta H}{\delta x} + \dot{\lambda} \right] + \delta u^T \frac{\delta H}{\delta u} \right\} dt$$

$\delta J_a = 0$ GIVES

$$\begin{cases} \delta x^T \left[\frac{\delta \Theta}{\delta x} - \lambda \right] = 0 & ;\ t = t_0, t_f \\ \dot{\lambda} = -\frac{\delta H}{\delta x} & ;\ \dot{x} = f(x, u, t) = \frac{\delta H}{\delta \lambda} \\ \frac{\delta H}{\delta u} = 0 & \end{cases}$$

EXAMPLE

$$\dot{x}_1 = x_2 \qquad x_1(0)=0$$
$$\dot{x}_2 = x_3 \qquad x_2(0)=0$$
$$\dot{x}_3 = U \qquad x_3(0)=0$$

TERMINAL MANIFOLD: $x_1^2(1) + x_2^2(1) = 1$

$$J = \frac{1}{2}\int_0^1 U^2 dt$$
$$H = \phi + \lambda^T f$$
$$= \frac{1}{2}U^2 + \lambda_1 x_2 + \lambda_2 x_3 + \lambda_3 U$$
$$\frac{\delta H}{\delta U} = 0 = U + \lambda_3 -$$
$$\dot{\lambda} = -\frac{\delta H}{\delta x} \Rightarrow \dot{\lambda}_1 = 0 , \quad \dot{\lambda}_2 = \lambda_1 \quad \dot{\lambda}_3 = \lambda_2$$
$$N(x(t_f), t_f) = 0 = x_1^2(1) + x_2^2(1) - 1$$
$$\frac{\delta \Theta}{\delta x} + \frac{\delta N^T}{\delta x} \nu = \lambda(1) \Rightarrow \lambda_1(1) = 2x_1(1)\nu$$
$$\lambda_2(1) = 2x_2(1)\nu$$
$$\lambda_3(1) = 0$$

OUR SYSTEM OF EQUATIONS ARE THUS

$$\dot{x}_1 = x_2 \quad ; \; x_1(0) = 0$$
$$\dot{x}_2 = x_3 \quad ; \; x_2(0) = 0$$
$$\dot{x}_3 = U \quad ; \; x_3(0) = 0$$
$$\lambda_3 = -U$$
$$\left.\begin{array}{ll} \dot{\lambda}_1 = 0 & \lambda_1(1) = 2x_1(1)\nu \\ \dot{\lambda}_2 = \lambda_1 & \lambda_2(1) = 2x_2(1)\nu \\ \dot{\lambda}_3 = \lambda_2 & \lambda_3(1) = 0 \end{array}\right\} x_1^2(1) + x_2^2(1) = 1$$

SAME PROBLEM WITH INITIAL & TERMINAL RESTRAINTS;

$$M[x(t_0), t_0]=0, \quad N[x(t_f), t_f]=0$$

AUGMENT J AS

$$J_a = \Theta[x,t]_{t_0}^{t_f} - \xi^T M[x(t_0), t_0] + \nu^T N[x(t_f), t_f]$$
$$+ \int_{t_0}^{t_f} [H - \lambda^T \dot{x}] dt$$

PROCEEDING AS BEFORE GIVES:

$$\delta J_a = \delta x^T \left[\frac{\delta \Theta}{\delta x} - \lambda \right]_{t_0}^{t_f} - \delta x^T \frac{\delta M^T}{\delta x} \xi \Big|_{t_0}$$
$$+ \delta x^T \frac{\delta N}{\delta x}\Big|_{t_f} + \int_{t_0}^{t_f} \left[\delta x^T \frac{\delta H}{\delta x} + \delta \dot{x}^T \dot{\lambda} + \delta u^T \frac{\delta H}{\delta u} \right] dt$$

THUS, IN ADDITION TO OUR PREVIOUS CONSTRAINTS WE HAVE

$$\lambda(t_0) = \frac{\delta \Theta}{\delta x} + \frac{\delta M^T}{\delta x} \xi \quad ; \; t = t_0$$
$$\lambda(t_f) = \frac{\delta \Theta}{\delta x} + \frac{\delta N^T}{\delta x} \nu \quad ; \; t = t_f$$

B. FIXED BEGINNING, t_f IS FREE, $x(t_f)$ IS FREE

$$\begin{cases} J = \Theta[x(t_f), t_f] + \int_{t_0}^{t_f} \phi(x,u,t)\,dt \\ \dot{x} = f(x,u,t) \qquad ; \; x(0) = x_0 \\ N[x(t_f)\; t_f] = 0 \end{cases}$$

AUGMENT J:

$$J_a = \Theta\Big|_{t_f} + \nu^T N\Big|_{t_f} + \int_{t_0}^{t_f}[\phi + \lambda^T(f - \dot{x})]dt$$

$$= \Theta\Big|_{t_f} + \nu^T N\Big|_{t_f} + \int_{t_0}^{t_f}[H - \lambda^T\dot{x}]dt$$

PEFORM AN INTEGRATION BY PARTS:

$$J_a = \Theta\Big|_{t_f} + \nu^T N\Big|_{t_f} + \int_{t_0}^{t_f} H dt$$

$$- \left(\lambda^T x\Big|_{t_0}^{t_f} - \int_{t_0}^{t_f} \dot{\lambda}^T x dt\right)$$

LET $\odot = \Theta + \nu^T N$

$$J_a = \odot\Big|_{t_f} - \lambda^T x\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f}[H + \dot{\lambda}^T x]dt$$

PERFORM VARIATION:

$$\delta J_a = \delta x^T \frac{\delta\odot}{\delta x}\Big|_{t_f} + \delta t_f \frac{\delta\odot}{\delta t_f}$$

$$+ \delta t_f \left(\frac{\delta\odot}{\delta x}\right)\frac{dx}{dt}\Big|_{t_f}$$

$$- \Big[\delta x^T \frac{\delta \lambda^T x}{\delta x} + \delta t \frac{\delta(\lambda^T x)^T}{\delta \lambda}\frac{d\lambda}{dt}$$

$$+ \delta t \left(\frac{\delta \lambda^T x}{\delta x}\right)^T \frac{dx}{dt}\Big]\Big|_{t_0}^{t_f}$$

$$+ \int_{t_0}^{t_f} \delta x^T \left[\frac{\delta(H + \dot{\lambda}^T x)}{\delta x} + \delta u^T \frac{\delta H}{\delta u}\right]dt$$

$$+ (H + \dot{\lambda}^T x)\Big|_{t_f} \delta t_f$$

REARRANGING AND RECOGNIZING THAT

$$\dot{\lambda} = - \delta\odot/\delta x$$

GIVES

$$\delta J = \delta t_f \left[H + \frac{\delta\odot}{\delta t_f}\right]_{t_f} + \delta x^T\left[\frac{\delta\odot}{\delta x} - \lambda\right]_{t_f}$$

$$+ \int_{t_0}^{t}\left[\delta x^T\left(\frac{\delta H}{\delta x} + \dot{\lambda}\right) + \delta u^T\left(\frac{\delta H}{\delta u}\right)\right]dt$$

SETTING $\delta J = 0$ GIVES

$$\begin{cases} H = \phi + \lambda^T f \\ \frac{\delta H}{\delta \lambda} = \dot{x} = f \\ \frac{\delta H}{\delta x} = -\dot{\lambda} = \frac{\delta f^T}{\delta x} \lambda + \frac{\delta \phi}{\delta x} \\ \frac{\delta H}{\delta U} = 0 = \frac{\delta \phi}{\delta U} + \frac{\delta f^T}{\delta U} \lambda \end{cases}$$

TRANSVERSALITY CONDITIONS ARE:

$$\begin{cases} \lambda(t_f) = \delta \Theta / \delta x(t_f) \\ N = 0 \\ H + \frac{\delta \Theta}{\delta t_f} + \frac{\delta N^T}{\delta t_f} \gamma = 0 \end{cases}$$

BOLZA PROBLEM WITH INEQUALITY CONSTRAINTS

THE PONTRYAGIN MAXIMUM PRINCIPLE

$$\begin{cases} \dot{x} = f(x, u, t) \\ N[x(t_f), t_f] = 0 \qquad x(t_0) = x_0 \\ g[x, u, t] \geq 0 \\ J = \theta[x(t_f), t_f] + \int_{t_0}^{t_f} \phi(x, u, t)\,dt \end{cases}$$

LET $\dot{z}^2 = g > 0$

$$J = \theta[x(t_f), t_f] + \xi^T x(t_0) + \nu^T N[x(t_f), t_f]$$
$$+ \int_{t_0}^{t_f} [H(x, \dot{w}, \lambda, t) - \lambda^T \dot{x} - \Gamma^T [g - \dot{z}^2]]\,dt$$

$H = \phi + \lambda^T f$

$\dot{w} = u \qquad w(t_0) = 0$

$\Phi \triangleq H - \lambda^T \dot{x} - \Gamma^T (g - \dot{z}^2) \Leftarrow$ LAGRANGIAN

EULER LAGRANGE BECOMES

$$\frac{d}{dt} \frac{\delta \Phi}{\delta \dot{x}} - \frac{\delta \Phi}{\delta x} = 0$$

$$\frac{d}{dt} \frac{\delta \Phi}{\delta \dot{w}} = 0$$

$$\frac{d}{dt} \frac{\delta \Phi}{\delta z} = 0$$

LET $\ominus = \theta + \nu^T N$

TRANSVERSALITY CONDITIONS BECOME

$$\frac{\delta \ominus}{\delta t_f} + \frac{\delta \ominus}{\delta x}^T \dot{x} + \Phi \Big|_{t_f} = 0$$

EXAMPLE:

$J = t_f$

$\dot{x} = Ax + Bu$

$x(t_0) = x_0$

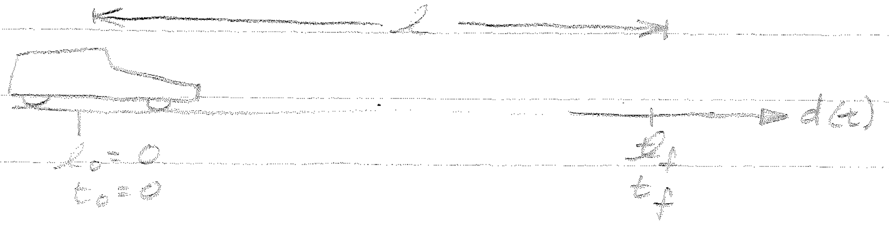
$H = \lambda^T [Ax + Bu]$

$\|U\| \leq 1$

OPTIMUM WHEN $\lambda^T BU \geq \lambda^T B\hat{U}$

OCCURS WHEN $u_i = -\operatorname{sgn} q_i = -\operatorname{sgn} B^T \lambda$

EXAMPLE CONTROL PROBLEM



$$\frac{d^2}{dt^2} d(t) = \underset{\underset{\text{ACCELERATION}}{\uparrow}}{\alpha(t)} + \underset{\underset{\text{BREAKING}}{\nwarrow}}{\beta(t)}$$

IN NORMAL FORM $x_1 = d$ $x_2 = \frac{d}{dt} d$ $u_1 = \alpha$ $u_2 = \beta$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

WHEN A MATRIX IS OF THE FORM

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & & & \\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots \end{bmatrix}$$

IT IS IN "PHASE VARIABLE" FORM

INPUT CONSTRAINTS: $0 \le \alpha(t) \le M$

$-N \le \beta(t) \le 0$

STATE VARIABLE CONSTRAINTS

$$\left.\begin{array}{ll} x_1(t_0) = \ell_0 = 0 & x_2(t_0) = 0 \\ x_1(t_0) = \ell & x_2(t_0) = 0 \end{array}\right\} \text{CAR'S SPEED AT THESE POINTS IS 0}$$

IF THE CAR ONLY GOES FORWARD

$0 \le x(t) \le \ell$ $\quad 0 \le x_2(t)$

SOLUTIONS FOLLOWING THESE CONSTRAINTS ARE "ADMISSABLE TRAJECTORIES"

STATE VARIABLE SYSTEM REPRESENTATION

NONLINEAR: $\dot{x}(t) = f[x, u, t]$

LINEAR: $\dot{x}(t) = A(t)\,x(t) + B(t)\,u(t)$; $x(t_0) = x_0$

HOMOGENEOUS SOLUTION ($u=0$) IS

$$x(t) = \Phi(t, t_0)\, x_0$$

$\Phi(t,\tau)$, THE "STATE TRANSITION" MATRIX, HAS THE FOLLOWING PROPERTIES:

(1) $\Phi(t_0, t_0) = I$

PROOF: $x(t_0) = \Phi(t_0, t_0)\, x(t_0)$

(2) $\Phi(t_1, t_2)\,\Phi(t_2 t_3) = \Phi(t_1, t_3)$

PROOF: $x(t_3) = \Phi(t_3 t_2)\, x(t_2)$

$x(t_2) = \Phi(t_2 t_1)\, x(t_1)$

$\Rightarrow x(t_3) = \Phi(t_3 t_2)\,\Phi(t_2 t_1)\, x(t_1) = \Phi(t_3 t_1)\, x(t_1)$

(3) $\Phi(t_1, t_2) = \Phi^{-1}(t_2 t_1)$

PROOF: $x(t_2) = \Phi(t_2 t_1)\, x(t_1)$

$\Rightarrow x(t_1) = \Phi^{-1}(t_2 t_1)\, x(t_2) = \Phi(t_1 t_2)\, x(t_2)$

THE GENERAL SOLUTION TO $\dot{x} = Ax + Bu$ IS

$$x(t) = \Phi(t\, t_0)\, x_0 + \int_{t_0}^{t} \Phi(t,\tau)\, B(\tau)\, u(\tau)\, d\tau$$

PROOF: $\dot{x}(t) = \frac{d}{dt}\Phi(t\, t_0)x_0 + \int_{t_0}^{t} \frac{d}{dt}\Phi(t,\tau)\, B(\tau)\, u(\tau)\, d\tau$
$+ \Phi(t, t)\, B(t)\, u(t)$

BUT $\dot{\Phi}(t,\tau) = A(t)\Phi(t,\tau)$

AND $\Phi(t, t) = I$

$\Rightarrow \dot{x}(t) = A(t)\,\Phi(t, t_0)\, x_0 + \int_{t_0}^{t} A(t)\,\Phi(t,\tau)\, B(\tau)\, u(\tau)\, d\tau$
$+ B(t)\, u(t)$

$= A(t)\left[\Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t,\tau)\, B(\tau)\, u(\tau)\, d\tau\right]$
$+ B(t)\, u(t)$

$= A(t)\, x(t) + B(t)\, u(t)$

SOLUTION FOR TIME INVARIANT CASE

$$\mathring{x}(t) = A\,x(t) \Rightarrow x(t) = \Phi(t - t_0)\,x(t_0)$$

(1) $\mathring{x}(t) = A\,x(t) \Rightarrow x(t) = e^{At}\,x(t_0)$

RECALL $x(t) = \Phi(t - t_0)\,x(t_0)$

$\Rightarrow \Phi(t\ ) = e^{At}$

(2) $\mathring{x}(t) = A\,x(t) + B\,U(t)$

$s\,\underline{X}(s) - x(0) = A\,\underline{X}(s) + B\,\underline{U}(s)$

$[sI - A]\,\underline{X}(s) = x(0) + B\,\underline{U}(s)$

$\underline{X}(s) = [sI - A]^{-1}x(0) + [sI - A]^{-1}B\,\underline{U}(s)$

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

OR $sI - A = \mathcal{L}\left[\Phi(t)\right]$

(3) EIGEN VALUE APPROACH

- DIAGONALIZE "A" MATRIX $\Rightarrow$

$AP = PJ$

$$\begin{bmatrix} & A & \end{bmatrix}\begin{bmatrix} P \end{bmatrix} = \begin{bmatrix} P \end{bmatrix}\begin{bmatrix} & J & \end{bmatrix}$$

THEN $\Phi(t) = e^{At}$

$= e^{PJP^Tt}$

$= P\,e^{Jt}\,P^T$

$$\Rightarrow \Phi(t) = P\begin{bmatrix} e^{\lambda_1 t} & & & 0 \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ 0 & & & e^{\lambda_n t} \end{bmatrix}P^T$$

SOLUTION IS CLOSED FORM

CONTROLLABILITY

$$\dot{x} = f[x, u, t] \quad , \; x(t_0) = x_0$$

A SYSTEM IS COMPLETELY STATE CONTROLLABLE IF $\forall \; x_0 = x(t_0)$ AND $x(t_f) \; \exists \; u(t) \ni \|u(t)\| < \infty$ $\ni$ $x(t_0)$ CAN BE TRANSFERRED TO $x(t_f)$ IN FINITE TIME $t_f - t_0 < \infty$.

THEOREM: THE SYSTEM

$$\dot{x} = Ax + Bu \qquad y = Cx$$

IS COMPLETELY STATE CONTROLLABLE ON $[t_0, t_f]$ IFF

$$M_0(t_0, t_f) \triangleq \int_{t_0}^{t_f} H_x(t_f, \tau) H_x^T(t_f, \tau) d\tau \qquad ①$$

IS NON-SINGULAR. THESE QUANTITIES ARE DEFINED AS FOLLOWS

$$\dot{x}(t) = A(t)x(t) + B(t)u(t)$$

$$x(t_f) = \Phi(t_f, t_0)x(t_0) + \int_{t_0}^{t_f} \Phi(t_f, \tau)B(\tau)u(\tau)d\tau$$

LET $x^d(t_f) = \int_{t_0}^{t_f} \Phi(t_f, \tau)B(\tau)u(\tau)d\tau$

$$\mathcal{H}_x(t_f, \tau) = \Phi(t_f, \tau)B(\tau)$$

$$x^o(t_f) = \Phi(t_f, t_0)x(t_0)$$

$$\Rightarrow x(t_f) - x^o(t_f) = x^d(t_f) = \int_{t_0}^{t_f} \mathcal{H}_x(t_f, \tau)u(\tau)d\tau$$

① IS A NECESSARY AND SUFFICIENT CONDITION

PROOF: (SUFFICIENCY)

LET $u(\tau) = \mathcal{H}_x^T(t_f, \tau)\lambda \neq 0$

THEN $x^d(t_f) = \int_{t_0}^{t_f} \mathcal{H}(t_f, \tau)H^T(t_f, \tau)d\tau \, \lambda$

NOW, IF $M_0(t_0, t_f)$ IS NON-SINGULAR THEN $\exists \lambda \ni \lambda = M_0^{-1}(t_0, t_f)\, x^d(t_f)$

(NECESSITY)

PROOF BY COUNTERDICTION

ASSUME $M_0(t_0, t_f)$ IS SINGULAR & THE SYSTEM IS STATE CONTROLLABLE. THEN $\exists$ $C \ni$

$$M_0 C = 0 \quad \ni C \neq 0$$

$$C^T M_0^T = 0$$

NOTE THAT $M_0 = M_0^T \Rightarrow C^T M_0 = 0$, OR, $\exists$ $C \neq 0 \ni$

$$C^T \int_{t_0}^{t_f} H_x(t_f, \tau) H_x^T(t_f, \tau) d\tau = 0$$

$$\Rightarrow C^T M_0 C = \int_{t_0}^{t_f} C^T \mathcal{H}_x(t_f, \tau) H_x^T(t_f, \tau) C \, d\tau$$

$$= \int_{t_0}^{t_f} [\mathcal{H}^T(t_f, \tau) C]^T [H^T(t_f, \tau) C] d\tau$$

$H^T(t_f, \tau) C$ IS A $m \times 1$ COLUMN VECTOR

LET $\eta = H^T(t_f, \tau) C$

$$\Rightarrow C^T M_0 C = \int_{t_0}^{t_f} \eta^T \eta \, d\tau$$

$$= \int_{t_0}^{t_f} (\eta_1^2 + \eta_2^2 + \ldots + \eta_m^2) d\tau$$

($\Rightarrow C^T M_0 C$ IS A POS. DEF. MATRIX WHICH COUNTERDICTS THE ASSUMPTION $C^T M_0 = 0$)

WE HAVE

$$x(t_f) - x^0(t_f) = \int_{t_0}^{t_f} H_x(t_f, \tau) u(\tau) d\tau$$

FOR CONTROLLABILITY, THE MUST BE TRUE $\forall$ $x(t_f)$

LET $x(t_f) = C$ (AND, WLOG, $x(t_0) = 0$)

$$\Rightarrow C = \int_{t_0}^{t_f} H_x(t_f, \tau) u(\tau) d\tau$$

$$C^T C = \int_{t_0}^{t_f} \underbrace{C^T H_x(t_f, \tau)}_{=0} u(\tau) d\tau$$

$$= 0$$

$$\Rightarrow C^T C = 0 \Rightarrow C_1 = C_2 = C_3 = \ldots = C_m = 0$$

THIS COUNTERDICTS THE INITIAL ASSUMPTION THAT $C \neq 0$

OUTPUT CONTROLLABILITY

$$Y = CX$$

$$= C\left[\Phi(t,t_0)X_0 + \int_{t_0}^{t_f} \Phi(t,\tau)B(\tau)u(\tau)d\tau \right]$$

LET $\quad Y_0 = C\Phi(t,t_0)X(t_0)$

$$H_Y(t,\tau) = C\Phi(t,\tau)B(\tau)$$

$$\Rightarrow Y = Y_0 + \int_{t_0}^{t_f} H_Y(t,\tau)u(\tau)d\tau$$

THIS IS THE SAME TYPE OF FORMULATION WE HAD BEFORE.

$\Rightarrow$ THE SYSTEM IS OUTPUT CONTROLLABLE WHEN THE MATRIX $N_0(t_f,t_0) = \int_{t_0}^{t_f} H_Y(t_f,\tau) H_Y^T(t_f,\tau)d\tau$ IS NONSINGULAR, THEN THE SYSTEM IS COMPLETELY OUTPUT CONTROLLABLE. ON THE TIME INTERVAL $[t_0, t_f]$

IF N_0 (OR M_0) IS NON-SINGULAR, THEN THE COLUMN VECTORS OF $H_x^T(t,\tau)$ ARE LINEARLY INDEPENDENT.

~~IF $H_x^T(t,\tau)$ IS NOT SQUARE, THEN AT LEAST n(= # OF STATES) COLUMNS MUST BE LINEARLY INDEPENDENT.~~

SIMPLE TEST FOR CONTROLLABILITY

$$\dot{x} = AX + BU \qquad Y = CX$$

$$M_c(t_0,t_f) = \int_{t_0}^{t_f} H_x(t_f,\tau) H_x^T(t_f,\tau)\,d\tau$$

$$H_x(t_f,\tau) = \Phi(t_f,\tau)B(\tau)$$

WE HAVE ESTABLISHED THAT THE SYSTEM IS CONTROLLABLE IF M_c IS NON-SINGULAR, OR, EQUIVALENTLY, IF ALL COLUMNS OF $\mathcal{H}_x^T(t_f,\tau)$.

THUS, $\forall\ \eta \neq 0$

$$\mathcal{H}_x^T(t_f,\tau)\,\eta \neq 0$$

OR $\eta^T \mathcal{H}_x(t_f,\tau) \neq 0$

OR $\eta^T \Phi(t_f,\tau)B(\tau) \neq 0$

CONSIDER NON-CONTROLLABITY SITUATION

$$\eta^T \Phi(t_f,\tau)B(\tau) = 0 \qquad ①$$

NOW, WE KNOW $\Phi(t_f,\tau)$ IS NON-SINGULAR AND

$$\frac{d}{d\tau}\Phi(t_f,\tau) = -\Phi(t_f,\tau)A(\tau)$$

$$\frac{d}{dt}\Phi(t,t_0) = A(t)\Phi(t,\tau)$$

DIFFERENTIATING ① W.R.T. τ:

$$\eta^T[\dot{\Phi}(t_f,\tau)B(\tau) + \Phi(t_f,\tau)\dot{B}(\tau)] = 0$$

$$\eta^T[-\Phi(t_f,\tau)A(\tau)B(\tau) + \Phi(t_f,\tau)\dot{B}(\tau)] = 0$$

$$\eta^T\Phi(t_f,\tau)[\dot{B}(\tau) - A(\tau)B(\tau)] = 0$$

LET $\Gamma_1(\tau) = B(\tau)$

$$\Rightarrow \eta\,\Phi(t_f,\tau)\Gamma_1(\tau) = 0 \qquad ②$$

AND $\eta^T\Phi(t_f,\tau)[\dot{\Gamma}_1(\tau) - A(\tau)\Gamma_1(\tau)] = 0$

LET $\Gamma_2(\tau) = \dot{\Gamma}_1(\tau) - A(\tau)\Gamma_1(\tau)$

$$\Rightarrow \eta\,\Phi(t_f,\tau)\Gamma_2(\tau) = 0 \qquad ③$$

FROM ② AND ③, WE HAVE SET A PATTERN. THE $n-1$ 'ST DERIVATIVE IS

$$\eta^T \Phi(t_f, \tau) \Gamma_n(\tau) = 0$$

WHERE $\Gamma_n(\tau) = \dot{\Gamma}_{n-1}(\tau) - A(\tau)\Gamma_{n-1}(\tau)$

IN GENERAL, WE HAVE THE RECURRSION:

$$\Gamma_{K+1}(\tau) = \Gamma_K(\tau) - A(\tau)\Gamma_K(\tau) \quad ; \; K = 1, 2, .., n-1$$

COMBINING ALL OF THESE RELATIONS INTO A SINGLE EQUATION GIVES

$$\eta^T \Phi(t_f, \tau) \left[\Gamma_1 \,\vdots\, \Gamma_2 \,\vdots\, \cdots \,\vdots\, \Gamma_n \right] \neq 0$$

WHERE WE HAVE ONCE AGAIN INTRODUCED THE INEQUALITY NECESSARY FOR (STATE) CONTROLLABILITY.

CALCULUS OF VARIATIONS

A. FUNDEMENTAL CONCEPTS

- FUNCTION: f IS A MAPPING ASSIGNING TO EACH ELEMENT q IN A CERTAIN SET $\mathscr{D}$ A UNIQUE ELEMENT IN A SET $\mathscr{R}$. $\mathscr{D}$ IS THE DOMAIN, $\mathscr{R}$ THE RANGE

- FUNCTIONAL: J IS A MAPPING ASSIGNING EACH FUNCTION X IN A CLASS Ω A REAL NUMBER. Ω IS THE DOMAIN OF THE FUNCTIONAL AND THE ASSOCIATED NUMBERS ARE THE "RANGE"

- LINEAR: $f(\cdot)$ IS LINEAR IFF $f(\alpha q_1 + \beta q_2)$
  $= \alpha f(q_1) + \beta f(q_2)$

- NORM: IN AN n DIMENSIONAL EUCLIDEAN SPACE, THE NORM IS A RULE OF CORRESPONDENCE ASSINING EACH POINT q OF THE SPACE A REAL NUMBER. THE NORM, $\|q\|$, MUST SATISFY
  (1) $\|q\| \geq 0$ WITH $\|q\| = 0$ IFF $q = 0$
  (2) $\|\alpha q\| = |\alpha| \, \|q\|$ $\forall$ REAL α
  (3) $\|q_1 + q_2\| \leq \|q_1\| + \|q_2\|$

- CLOSE: TWO FUNCTIONS OR FUNCTIONALS ARE "CLOSE" IF $\|q_1 - q_2\|$ IS "SMALL

-INCREMENT OF A FUNCTIONAL

$$\Delta f \triangleq f(q+\Delta q) - f(q)$$

OR $\Delta J \triangleq J(x+\delta x) - J(x)$

- VARIATION OF A FUNCTIONAL

$$df = \frac{\delta f}{\delta q_1}\Delta q_1 + \frac{\delta f}{\delta q_2}\Delta q_2 + \ldots + \frac{\delta f}{\delta q_n}\Delta q_n$$

OR

$$\Delta J(x, \delta x) = \delta J(x, \delta x) + g(x, \delta x)\|\delta x\|$$

$\delta J(x, dx)$, THE VARIATION OF J, IS A LINEARIZATION ON ΔJ. THAT IS, BY EXPANDING ΔJ IN A TAYLOR SERIES, AND RETAINING ONLY FIRST ORDER TERMS, IN $\delta x(t)$, WE OBTAIN $\delta J(x, \delta x)$

- FUNCTION MINIMA ∓ MAXIMA AT q^*

MAXIMUM IF

$$\Delta f = f(q) - f(q^*) \leq 0$$

MINIMUM IF

$$\Delta f \geq 0$$

$\ni \|q - q^*\| < \epsilon$ IN A LIMIT SENSE

- GLOBAL EXTREMUM

IF ABOVE CONDITIONS HOLD FOR ALL ϵ.

- FUNCTIONAL MINIMA & MAXIMA

J IS EXTREMUM @ x^* IF $\exists \epsilon > 0$

$\ni \forall x \in \Omega, \ \|x - x^*\| < \epsilon$

$$\Delta J = J(x) - J(x^*) \geq 0 \Rightarrow \text{MIN}$$

$$\Delta J = J(x) - J(x^*) \leq \ \Rightarrow \text{MAX}$$

IF CONDITION HOLDS $\forall \epsilon$, THEN $J(x^*)$ IS GLOBAL EXTREMUM.

- FUNDEMENTAL THEOREM OF THE CALCULUS OF VARIATIONS

LET $x(t) \in \Omega$ AND $J(x)$ BE A DIFFERENTIABLE FUNCTION OF x.

IF Ω IS <u>NOT</u> CONSTRAINED BY ANY BOUNDRIES, THEN, IF x^* EXTREMIZES J, THEN

$$\delta J(x^*, \delta x) = 0$$

FOR ALL ADMISSIBLE δx

B. FUNCTIONALS OF A SINGLE FUNCTION

1. THE SIMPLEST PROBLEM (t_0, t_f, X_0 ‡ X_f SPECIFIED)

$$J(x) = \int_{t_0}^{t_f} g[x(t), \dot{x}(t), t]\,dt$$

WE WISH TO FIND X^* TO EXTREMIZE THIS FUNCTION GIVEN

(a) $g(\;)$ IS TWICE DIFFERENTIABLE

(b) t_0 AND t_f ARE FIXED

$$x_0 = X(t_0) \qquad x_f = x(t_f)$$

NOW

$$\Delta J(x, \delta x) = J(x + \delta x) - J(x)$$
$$= \int_{t_0}^{t_f} g[x + \delta x, \dot{x} + \delta \dot{x}, t]\,dt - \int_{t_0}^{t_f} g(x, \dot{x}, t)\,dt$$

EXPANDING IN A TAYLOR SERIES:

$$\Delta J = \int_{t_0}^{t_f} \Big[\Big[\cancel{g(x, \dot{x}, t)} + \Big\{ \frac{\delta g(x,\dot{x},t)}{\delta x}\delta x + \frac{\delta g(x,\dot{x},t)}{\delta \dot{x}}\delta \dot{x} \Big\}$$
$$+ \frac{1}{2}\Big\{ \frac{\delta^2 g(x,\dot{x},t)}{\delta x^2}\delta^2 x + 2\frac{\delta^2 g(x,\dot{x},t)}{\delta x \delta \dot{x}}\delta x \delta \dot{x} + \frac{\delta^2 g(x,\dot{x},t)}{\delta \dot{x}^2} \Big\}$$
$$+ 0\{\delta^2 x, \delta^2 \dot{x}\Big] - \cancel{g(x, \dot{x}, t)}\Big]\,dt$$

AND

$$\delta J = \int_{t_0}^{t_f} \Big\{ \frac{\delta g}{\delta x}\delta x + \frac{\delta g}{\delta \dot{x}}\delta \dot{x} \Big\}\,dt$$

NEXT, INTEGRATE BY PARTS. NOTE THAT

$$\delta x(t) = \delta x(t_0) + \int_{t_0}^{t} \delta \dot{x}(t)$$

$$U = \frac{\delta g}{\delta \dot{x}} \qquad dv = \delta \dot{x}\, dt$$

$$dU = \frac{d}{dt}\frac{\delta g}{\delta \dot{x}}\, dt \qquad V = \delta x$$

$$\Rightarrow \delta J = \frac{\delta g}{\delta \dot{x}}\,\delta x \Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \left[\frac{\delta g}{\delta x} - \frac{d}{dt}\frac{\delta g}{\delta \dot{x}} \right] \delta x\, dt$$

SINCE t_0 AND t_f ARE FIXED, $\delta x(t_0) = \delta x(t_f) = 0$ AND THE FIRST TERM VANISHES. TO FIND EXTREMA, WE APPLY THE FUND. THEOREM OF VARIATIONAL CALCULUS & WRITE

$$\delta J = 0 = \int_{t_0}^{t_f} \left[\frac{\delta g}{\delta x} - \frac{d}{dt}\frac{\delta g}{\delta \dot{x}} \right] \delta x\, dt$$

- FUNDEMENTAL LEMMA OF THE CALCULUS OF VARIATIONS:

IF $\int_{t_0}^{t_f} h(t)\, \delta x(t)\, dt = 0$

$\forall\ \delta x(t)$ CONTINUOUS ON $[t_0, t_f]$,

THEN $h(t) = 0$ ON $[t_0, t_f]$

THUS, OUR EXTREMA PROBLEM REDUCES TO "EULER'S EQUATION":

$$\frac{\delta g}{\delta x} - \frac{d}{dt}\frac{\delta g}{\delta \dot{x}} = 0$$

THE VALUE OF $X = X^*$ WHICH SATISEIES THIS, IS OUR OPTIMAL SOLUTION FOR THE CASE WHERE THE END POINTS ARE FIXED. (NOTE THE "SPLIT" BOUNDRY CONDITIONS)

2. t_f SPECIFIED, $X(t_f)$ FREE

(t_0 AND $X(t_0)$ SPECIFIED)

TO SOLVE THIS CASE, WE RECALL THE "INTEGRATION BY PARTS" SOLUTION OF THE PREVIOUS SECTION.

$$\delta J(x, \delta x) = \frac{\delta g}{\delta \dot{x}} \delta x \Big|_{t_0}^{t_f}$$

$$+ \int_{t_0}^{t_f} \left[\frac{\delta g}{\delta x} - \frac{d}{dt} \frac{\delta g}{\delta \dot{x}} \right] dt \; (= 0)$$

STILL, $X(t_0) = 0$. THUS, IN ADDITION TO SATISFYING EULER'S EQUATION, WE MUST SATISFY

$$\frac{\delta g[x(t_f), \dot{x}(t_f), t_f]}{\delta \dot{x}} \quad \delta x(t_f) = 0$$

SINCE $\delta x(t_f)$ IS ARBITRARY (SINCE $X(t_f)$ IS FIXED), THIS BECOMES

$$\frac{\delta g}{\delta \dot{x}} \Big|_{t = t_f} = 0$$

THIS IS A "NATURAL BOUNDRY CONDITION".

2. t_f, t_0, x_0 SPECIFIED, $x(t_f)$ FREE

IN OUR LAST SECTION, THE RESULT GIVEN AFTER PARTS INTEGRATION WAS:

$$\delta J(x, \delta x) = \frac{\delta g}{\delta \dot{x}} \delta x \Big|_{t_0}^{t_f}$$

$$+ \int_{t_0}^{t_f} \left[\frac{\delta g}{\delta x} - \frac{d}{dt} \frac{\delta g}{\delta \dot{x}} \right] \delta x dt$$

SO, AS BEFORE, WE MUST SATISFY EULER'S EQUATION:

$$\frac{\delta g}{\delta x} - \frac{d}{dt} \frac{\delta g}{\delta \dot{x}} = 0$$

<u>PLUS</u> THE CONDITION

$$\frac{\delta g}{\delta \dot{x}} \delta x \Big|_{t_0}^{t_f} = 0$$

NOW, SINCE $x_0 = x(t_0)$ IS SPECIFIED, THIS REDUCES TO

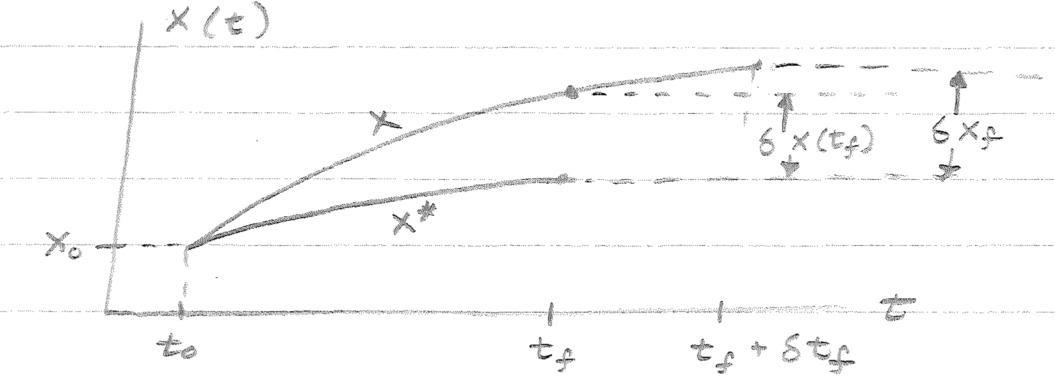
$$\frac{\delta g}{\delta \dot{x}} \delta x \Big|_{t_f} = 0$$

3. BOTH t_f AND $x(t_f)$ FREE

WE STILL WISH TO MINIMIZE

$$J = \int_{t_0}^{t_f} g\,[x(t), \dot{x}(t), t\,]\,dt$$

WITH t_f AND $x(t_f)$ FREE:



NOW

$$\Delta J = \int_{t_0}^{t_f+\delta t_f} g\,[x, \dot{x}, t\,]\,dt - \int_{t_0}^{t_f} g\,(x^*, \dot{x}^*, t)\,dt$$

$$= \int_{t_0}^{t_f} \{ g\,[x, \dot{x}, t\,] - g\,[x^*, \dot{x}^*, t\,] \}\,dt + \int_{t_f}^{t_f+\delta t_f} g\,[x, \dot{x}, t\,]\,dt$$

FIRST APPROXIMATE

$$\int_{t_f}^{t_f+\delta t_f} g[x, \dot{x}, t\,]\,dt = g\,[x(t_f), \dot{x}(t_f), t_f\,]\,\delta t_f$$

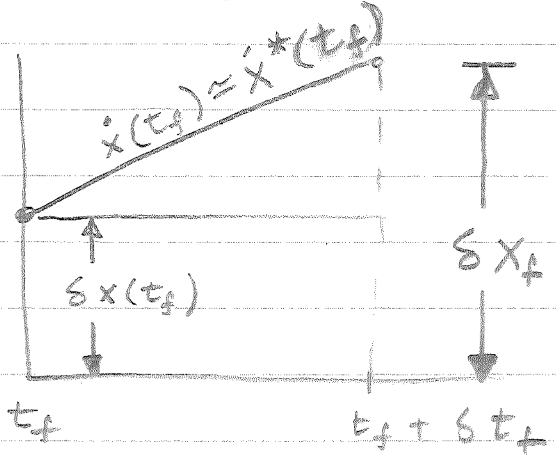
SECONDLY, INTERATE BY PARTS (AS BEFORE) RECOGNIZING THAT $\delta x(t_0) = 0$:

$$\int_{t_0}^{t_f} [\,g\,[x, \dot{x}, t\,] - g\,[x^*, \dot{x}^*, t\,]\,]\,dt = \frac{\delta g}{\delta x}\,\delta x \Big|^{t_f} + \int_{t_0}^{t_f} \left[\frac{\delta g}{\delta x} - \frac{d}{dt}\frac{\delta g}{\delta x} \right] dt\,\delta x$$

THUS, WE MAY APPROXIMATE:

$$\Delta J \cong \frac{\delta g\,[x(t_f), \dot{x}(t_f), t_f\,]}{\delta \dot{x}}\,\delta x(t_f) + g\,[x(t_f), \dot{x}(t_f), t_f\,]\,\delta t_f + \int_{t_0}^{t_f} \left[\frac{\delta g}{\delta x} - \frac{d}{dt}\,\delta g / \delta \dot{x} \right] \delta x\,dt$$

WE MUST NOW RELATE $\delta x(t_f)$ TO δt_f AND δx_f



$$\delta x_f \approx \delta x(t_f) + \dot{x}^*(t_f)\,\delta t_f$$

THUS, $\delta x(t_f) \approx \delta x_f - \dot{x}^*(t_f)\,\delta t_f$

SUBSTITUTING THIS & COLLECTING TERM;

$$\delta J = 0 = \frac{\delta g[x^*(t_f)\,\dot{x}^*(t_f)\,t_f]}{\delta \dot{x}}\,\delta x_f + \left[g(x^*(t_f), \dot{x}^*(t_f), t_f) - \frac{\delta g[x^*(t_f)\,\dot{x}^*(t_f), t_f]}{\delta \dot{x}}\,\dot{x}^*(t_f)\right]\delta t_f + \int_{t_0}^{t_f}\left[\frac{\delta g}{\delta x} - \frac{d}{dt}\frac{\delta g}{\delta \dot{x}}\right]\delta x\,dt$$

INTERRETATIONS

a. t_f SPECIFIED

HERE, $\delta t_f = 0$ AND

$$\delta J = 0 = \frac{\delta g^*}{\delta \dot{x}}\,\delta x\Big|_{t_f} + \int_{t_0}^{t_f}\left[\frac{\delta g}{\delta x} - \frac{d}{dt}\frac{\delta g}{\delta \dot{x}}\right]\delta x\,dt$$

(THIS IS THE SAME CASE CONSIDERED TWO PAGES AGO)

b. $x(t_f)$ SPECIFIED

HERE, $\delta x_f = 0$ AND

$$\delta J = \left[g^* - \frac{\delta g^*}{\delta \dot{x}} \dot{x} \right]\Big|_{t_f} \delta t_f + \int_{t_0}^{t_f} \left[\frac{\delta g}{\delta x} - \frac{d}{dt} \frac{\delta g}{\delta \dot{x}} \right] \delta x \, dt$$

THUS, WE MUST SATISFY

$$g^* - \frac{\delta g^*}{\delta \dot{x}} \dot{x} \Big|_{t_f} = 0$$

AND $\delta g / \delta x - \frac{d}{dt} \delta g / \delta \dot{x} = 0$

c. t_f & $x(t_f)$ UNSPECIFIED AND INDEPENDENT

HERE, WE MUST SATISFY ALL OF THE RELATIONS IN (a) & (b)

SPECIFICALLY:

$$\delta g / \delta \dot{x} \Big|_{t_f} = 0$$

$$g - \delta g / \delta \dot{x} \times \dot{x} \Big|_{t_f} = 0$$

AND EULER'S EQ.

d. $x(t_f)$ & t_f RELATED:

$$x(t_f) = \Theta(t_f)$$

THEN $\frac{\delta x_f}{\delta t_f} \approx \frac{\delta \Theta(t)}{\delta t} \Rightarrow \delta x_f = \frac{\delta \Theta(t_f)}{\delta t} \delta t_f$

SUBSTITUTING & SIMPLIFYING LEAVES:

$$\frac{\delta g}{\delta \dot{x}}\Big|_{t_f} \left[\frac{d\Theta(t_f)}{dt} - \dot{x}(t_f) \right]$$

$$+ g \Big|_{t_f} = 0$$

C. FUNCTIONS INVOLVING SEVERAL INDEPENDENT FUNCTIONS

HERE, JUST GENERALIZE PREVIOUS NOTIONS TO n SPACE. EULER'S EQUATION BECOMES, FOR EXAMPLE,

$$\frac{\delta g(\bar{x})}{\delta \bar{x}} - \frac{d}{dt}\frac{\delta g(\bar{x})}{\delta \dot{\bar{x}}} = 0$$

WHERE g IS A SCALER AND $\bar{X}$ IS AN n VECTOR. FOR X_i, THE EQUATION IS

$$\delta g(\bar{x})/\delta x_i - \frac{d}{dt}\,\delta g(\bar{x})/\delta \dot{x}_i = 0$$

WE HAVE n EQUATIONS & n UNKNOWNS.

FOR t_f AND $X(t_f)$ FREE, OUR EQUATIONS BECOME

$$\frac{\delta g[\bar{x}^*, \dot{\bar{x}}^*, t]}{\delta \bar{x}} - \frac{d}{dt}\frac{\delta g[\bar{x}^*, \dot{\bar{x}}^*, t]}{\delta \dot{\bar{x}}} = 0$$

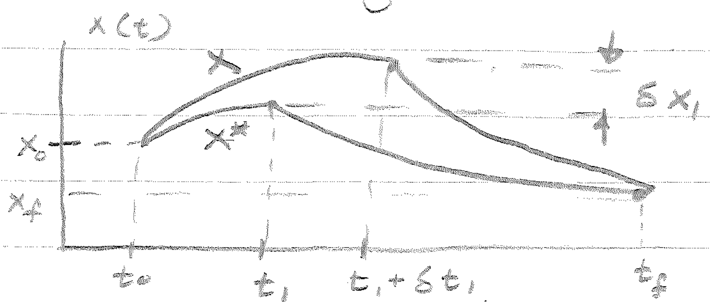
AND

$$\left[\frac{\delta g[\bar{x}^*(t_f), \dot{\bar{x}}^*(t_f)\ t_f]}{\delta \bar{x}^*}\right]^T \delta \bar{x}^*(t_f)$$

$$+ \left[g\{\bar{x}^*(t_f), \dot{x}^*(t_f), t_f\} - \frac{\delta g\{\bar{x}^*(t_f)\ \dot{\bar{x}}^*(t_f), t_f\}}{\delta \dot{\bar{x}}^*}\ \dot{\bar{x}}^*(t_f)\right] \delta t_f$$

D. PIECEWISE SMOOTH EXTREMALS

$$J(x) = \int_{t_0}^{t_f} g[x, \dot{x}, t]\,dt$$



ASSUME t_0 t_f x_0 & x_f ARE FIXED.

LET $J = J_1(x) + J_2(x)$

$$= \int_{t_0}^{t_1} g\,dt + \int_{t_1}^{t_f} g\,dt$$

NOW $\delta J = \delta J_1 + \delta J_2$

$$= \left[\frac{\delta g^*}{\delta \dot{x}}\Big|_{t_1-} \delta x_1 + \left\{ g^*\Big|_{t_1-} - \frac{\delta g^*}{\delta \dot{x}} \dot{x}^*\Big|_{t_1-} \right\} \delta t_1 + \int_{t_0}^{t_1} \left[\delta g^*/\delta x - \frac{d}{dt} \frac{\delta g^*}{\delta \dot{x}} \right] \delta x\,dt \right]$$

$$+ \left[- \frac{\delta g^*}{\delta \dot{x}}\Big|_{t_1+} \delta x_1 - \left\{ g^*\Big|_{t_1+} - \frac{\delta g^*}{\delta \dot{x}} \dot{x}^*\Big|_{t_1+} \right\} \delta t_1 + \int_{t_1}^{t_f} \left[\delta g^*/\delta x - \frac{d}{dt} \frac{\delta g^*}{\delta \dot{x}} \right] \delta x\,dt \right]$$

EULER'S EQUATION MUST BE SATISFIED ON BOTH INTERVALS. THE TRANSVERSALITY CONDITIONS ARE

$$\left\{ \frac{\delta g}{\delta \dot{x}}\Big|_{t_1-} - \frac{\delta g}{\delta \dot{x}}\Big|_{t_1+} \right\} \delta x_1$$

$$+ \left[\left\{ g^*\Big|_{t_1-} - \frac{\delta g^*}{\delta \dot{x}} \dot{x}^*\Big|_{t_1-} \right\} - \left\{ g^*\Big|_{t_1+} - \frac{\delta g^*}{\delta \dot{x}} \dot{x}^*\Big|_{t_1+} \right\} \right] dt = 0$$

(a) WHEN t_f ∮ $x(t_f)$ ARE INDEPENDENT, WE HAVE THE "WEIERSTRASS-ERDMAN CORNER CONDITIONS":

$$\frac{\delta g}{\delta \dot{x}}\Big|_{t_1^-} = \frac{\delta g}{\delta \dot{x}}\Big|_{t_1^+}$$

$$\left[g - \frac{\delta g}{\delta \dot{x}}\dot{x}\right]\Big|_{t_1^-} = \left[g - \frac{\delta g}{\delta \dot{x}}\dot{x}\right]\Big|_{t_1^+}$$

(b) WHEN $x(t_1) = \Theta(t_1)$:

$$\delta x_1 = \frac{d\Theta(t_1)}{dt}\delta t_1$$

THIS GIVES

$$\delta g/\delta \dot{x}\Big|_{t_1^-}\left[\frac{d\theta}{dt} - \dot{x}\right]\Big|_{t_1^-} + g\Big|_{t_1^-}$$

$$= \delta g/\delta \dot{x}\Big|_{t_1^+}\left[\frac{d\theta}{dt} - \dot{x}\right]\Big|_{t_1^+} + g\Big|_{t_1^+}$$

(c) FOR MANY FUNCTIONS, THE WEIERSTRASS-ERDMAN CORNER CONDITIONS ARE:

$$\frac{\delta g}{\delta \dot{\mathbf{x}}}\Big|_{t_1^-} = \frac{\delta g}{\delta \dot{\mathbf{x}}}\Big|_{t_1^+}$$

$$\left[g - \left[\frac{\delta g}{\delta \dot{\mathbf{x}}}\right]^T \dot{\mathbf{x}}\right]_{t_1^-} = \left[g - \left[\frac{\delta g}{\delta \dot{\mathbf{x}}}\right]^T \dot{\mathbf{x}}\right]_{t_1^+}$$

E. CONSTRAINED EXTREMA

1. EXAMPLE: SCALER USE OF LAGRANGE MULTIPLIERS

MINIMIZE $f(Y_1 Y_2) = Y_1^2 + Y_2^2$

SUBJECT TO $Y_1 + Y_2 = 5$

OUR AUGMENTED f IS

$$f_a = Y_1^2 + Y_2^2 + p(Y_1 + Y_2 - 5)$$

$p \equiv$ LAGRANGE MULTIPLIER

NOW

$$d f_a(Y_1 Y_2 p) = 0 = \frac{\delta f_a}{\delta Y_1} \delta P_1 + \frac{\delta f_a}{\delta Y_2} \delta Y_2 + \frac{\delta f_a}{\delta p} \delta p$$

$$= [2Y_1 + p] \delta Y_1 + [2Y_2 + p] \delta Y_2 + [Y_1 + Y_2 - 5] \delta p$$

$Y_1 + Y_5 - 5$ IS ALREADY ZERO

WE MUST ALSO REQUIRE THAT

$2Y_2 + p = 0$ AND $2Y_1 + p = 0$

WE HAVE 3 EQS. & 3 UNKNOWNS.

SOLUTION IS

$$Y_1 = 2.5, \quad Y_2 = 2.5, \quad p = -5$$

2. CONSTRAINED PROBLEM USING LAGRANGE MULTIPLIERS

MINIMIZE $f(Y_1, Y_2, \ldots, Y_{n+m})$

SUBJECT TO $a_i(Y_1, \ldots Y_{n+m}) = 0 \; ; \; i = 1, \ldots, n$

FORM, FIRST, THE AUGMENTED FUNCTION:

$$f_a(Y_1, Y_2, \ldots Y_{n+m}, p_1 \ldots p_n) = f(Y_1 Y_2 \ldots Y_{n+m}) + \Sigma_{i=1}^{n} p_i a_i(Y_1, Y_2, \ldots, Y_{n+m})$$

THEN

$$df_a = \Sigma_{K=1}^{n+m} \frac{\delta f_a}{\delta Y_K} dY_K + \Sigma_{i=1}^{n} \frac{\delta f_a}{\delta p_i} dp_i$$

BUT $\delta f_a / \delta p_i = a_i$

$$\Rightarrow df_a = \Sigma_{K=1}^{n+m} \frac{\delta f_a}{\delta Y_K} dY_K + \Sigma_{i=1}^{n} a_i dp_i$$

SETTING $df_a = 0$, THE EQUATIONS ARE

$$\left.\begin{array}{ll} a_i(Y_1, \ldots, Y_{n+m}) = 0 \; ; & i = 1, 2, \ldots, n \\ \frac{\delta f}{\delta Y_K} = 0 & K = 1, 2, \ldots n+m \end{array}\right\} \begin{array}{l} 2n+m \\ \text{EQUATIONS} \end{array}$$

● EXAMPLE: MINIMIZE: $f(Y_1 Y_2 Y_3) = Y_1^2 + Y_2^2 + Y_3^2$

SUBJECT TO $Y_3 = Y_1 Y_2 + 5$, $Y_1 + Y_2 + Y_3 = \underline{1}$

NOW: $f_a(Y_1 Y_2 Y_3 p_1 p_2) = (Y_1^2 + Y_2^2 + Y_3^2) + p_1[Y_1 Y_2 + 5 - Y_3] + p_2[Y_1 + Y_2 + Y_3 - 1]$

$a_i(Y_1, \ldots Y_{n+m}) = a_i(Y_1 Y_2 Y_3) = 0$

$$\Rightarrow Y_1 Y_2 + 5 - Y_3 = 0$$

$$Y_1 + Y_2 + Y_3 - 1 = 0$$

$\frac{\delta f_a}{\delta Y_K} = 0$

$$\Rightarrow \frac{\delta f_a}{\delta Y_1} = 0 = 2Y_1 + p_1 Y_2 + p_2$$

$$\frac{\delta f_a}{\delta Y_2} = 0 = 2Y_2 + p_1 Y_1 + P_2$$

$$\frac{\delta f_a}{\delta Y_3} = 0 = 2Y_3 - p_1 + P_2 = 0$$

SOLVING GIVES

$$(Y_1 Y_2 Y_3) = \begin{cases} 2, -2, 1 \\ \text{OR } -2, 2, 1 \end{cases}$$

$$\Rightarrow f_{a_{MIN}} = 9$$

3. CONSTRAINED MINIMIZATION OF FUNCTIONALS

– POINT CONSTRAINTS

MINIMIZE $J(w) = \int_{t_0}^{t_f} g[w, \dot{w}, t]\,dt$

$\ni$ w IS A $n+m$ VECTOR, SUBJECT TO

$f_i[w, t] = 0$, $i = 1, 2, \ldots, n$

(THESE ARE <u>POINT</u> CONSTRAINTS)

m OF THE $n+m$ COMPONENTS OF w ARE THUS INDEPENDENT.

FORM THE AUGMENTED FUNCTIONAL

$$J_a(w,p) = \int_{t_0}^{t_f}\Big[g(w,\dot{w},t) + \textstyle\sum_{i=1}^{n} p_i(t) f_i(w,t)\Big]dt$$
$$= \int_{t_0}^{t_f}\big[g(w,\dot{w},t) + p^T(t) f(w,t)\big]dt$$

HERE, $p(t)$ IS AN n VECTOR OF LAGRANGE MULTIPLIERS.

THE VARIATION OF J_a IS

$$\delta J_a(w, \delta w, p, \delta p)$$
$$= \int_{t_0}^{t_f}\Big(\Big[\frac{\delta g^T}{\delta w} + p^T(t)\frac{\delta f}{\delta w}\Big]\delta w$$
$$+ \Big[\delta g^T/\delta\dot{w} + p^T(t)\,\delta f/\delta\dot{w}\Big]\delta\dot{w}$$
$$+ f^T\delta p\Big)dt$$

NOTE THAT

$$\frac{\delta f}{\delta \dot{w}} = \begin{bmatrix} \delta f_1/\delta\dot{w}_1 & \cdots & \delta f_1/\delta\dot{w}_{n+m} \\ \vdots & & \vdots \\ \delta f_n/\delta\dot{w}_1 & \cdots & \delta f_n/\delta\dot{w}_{n+m} \end{bmatrix}$$

WE NOW INTEGRATE BY PARTS THE TERMS CONTAINING $\delta \dot{w}$. THE RESULT IS:

$$\delta J_a(w, \delta w, p, \delta p) = \int_{t_0}^{t_f} \left\{ \left[\frac{\delta g}{\delta w} + p^T \frac{\delta f}{\delta w} - \frac{dg^T}{dt} + p^T \frac{\delta f}{\delta \dot{w}} \right] \delta w + f^T \delta p \right\} dt$$

LOOKS LIKE WE'VE SPECIFIED ALL BOUNDRY CONDITIONS.

FIRST OFF, $f(w, \dot{w}, t) = 0$. WE CAN CHOOSE OUR LAGRANGE MULTIPLIERS ARBITRARILY. OUR OTHER EQUATION IS:

$$\frac{\delta g}{\delta w} + \frac{\delta f^T}{\delta w} p - \frac{d}{dt}\left[\frac{\delta g}{\delta \dot{w}} + \frac{\delta f^T}{\delta \dot{w}} p \right] = 0$$

NOTE THAT IF WE SPECIFY OUR AUGMENTED g AS

$$g_a = g + p^T f$$

THEN THIS CONSTRAINT BECOMES

$$\frac{\delta g_a}{\delta w} - \frac{d}{dt} \frac{\delta g_a}{\delta \dot{w}} = 0$$

WHICH IS AKIN TO EULER'S EQ. THAT IS, WE COULD HAVE SIMPLIFIED THIS WHOLE MESS BY STARTING OUT WITH

$$J_a = \int_{t_0}^{t_f} g_a [w, \dot{w}, p, t] dt$$

EXAMPLE;

MINIMIZE $J(w) = \int_{t_0}^{t_f} \frac{1}{2}[w_1^2 + w_2^2]dt$

SUBJECT TO $\dot{w}_1 = w_2$

THEN

$$g_a = \frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 + p(t)[w_2 - \dot{w}_1]$$

$$\frac{\delta g_a}{\delta w} - \frac{d}{dt}\frac{\delta g_a}{\delta \dot{w}} = 0$$

THUS $\Rightarrow$ $w_1 + \dot{p} = 0$

$w_2 + p = 0$

ALSO $\dot{w}_1 = w_2$

3 EQUATIONS & THREE UNKNOWNS

EXAMPLE

MINIMIZE: $J(x, u) = \int_{t_0}^{t_f} \frac{1}{2}[x_1^2 + x_2^2 + u^2]dt$

SUBJECT TO: $\dot{x}_1 = x_2 - x_1$; $\dot{x}_2 = -2x_1 - 3x_2 + u$

LET $w_1 = x_1$, $w_2 = x_2$, AND $w_3 = u$

$$\Rightarrow J(w) = \int_{t_0}^{t_f} \frac{1}{2}[w_1^2 + w_2^2 + w_3^2]dt$$

$$\dot{w}_1 = w_2 - w_1 \quad ; \quad \dot{w}_2 = -2w_1 - 3w_2 + w_3$$

$$g_a = \frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 + \frac{1}{2}w_3^2 + p_1[w_2 - w_1 - \dot{w}_1] + p_2[-2w_1 - 3w_2 + w_3 - \dot{w}_2]$$

$$\frac{\delta}{\delta w} g_a - \frac{d}{dt}\frac{\delta}{\delta \dot{w}} g_a = 0$$

THUS

$$\begin{cases} w_1 - p_1 - 2p_2 - \dot{p}_1 = 0 \\ w_2 + p_1 - 3p_2 - \dot{p}_2 = 0 \\ w_3 + p_2 = 0 \end{cases}$$

UTILIZE THESE EQUATIONS WITH THE CONSTRAINTS:

$\dot{w}_1 = w_2 - w_1$

$\dot{w}_2 = -2w_1 - 2w_2 + w_3$

5 EQUATIONS & 5 UNKNOWNS

4. ISOPERIMETRIC CONSTAINTS

CONSTRAINTS ARE OF THE FORM:

$$\int_{t_0}^{t_f} e_i(w, \dot{w}, t)\,dt = C_i \qquad (i = 1, 2, \ldots, r)$$

THE C_i'S ARE SPECIFIED CONSTANTS

LET $z_i(t) = \int_{t_0}^{t} e_i(w, \dot{w}, t)\,dt$

THUS: $z_i(t_f) = C_i$

NOW $\dot{z}_i(t) = e_i(w, \dot{w}, t)$; $i = 1, 2, \ldots, r$

DEFINE

$$g_a(w, \dot{w}, p, \dot{z}, t) = g(w, \dot{w}, t) + p^T(t)\left[e_i - \dot{z}\right]$$

SOLVING AS BEFORE GIVES US r EQUATIONS

$$\frac{\delta g_a}{\delta z} - \frac{d}{dt}\frac{\delta}{\delta \dot{z}} g_a = 0$$

PLUS, WE STILL HAVE n+m EQUATIONS FROM

$$\frac{\delta g_a}{\delta w} - \frac{d}{dt}\frac{\delta}{\delta \dot{w}} g_a = 0$$

PLUS, THE r CONSTRAINTS:

$$\dot{z}(t) = e(w, \dot{w}, t)$$

THIS GIVES $n+m+2r$ EQUATIONS

NOTE THAT $\delta g_a / \delta z = 0$

AND $\frac{\delta}{\delta \dot{z}} g_a = -p(t)$

THIS EQUATION IS THUS ALWAYS

$$\dot{p}(t) = 0$$

∴ THE LAGRANGE MULTIPLIERS ARE ALWAYS CONSTANTS

EXAMPLE:

EXTREMIZE

$$J(w) = \int_{t_0}^{t_f} \frac{1}{2}\left[w_1^2 + w_2^2 + 2\dot{w}_1 \dot{w}_2\right] dt$$

SUBJECT TO $\int_{t_0}^{t_f} w_2^2(t)\,dt = c$

NOW, $\dot{z} = w_2^2(t)$

$$\Rightarrow g_a = g + p^T[e - \dot{z}]$$

$$= \frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 + 2\dot{w}_1\dot{w}_2 + p(t)\left[w_2^2 - \dot{z}\right]$$

$$\delta g_a/\delta w - \frac{d}{dt}\frac{\delta}{\delta \dot{w}} g_a = 0$$

$$\Rightarrow \delta g_a/\delta w_1 - \frac{d}{dt}\frac{\delta}{\delta \dot{w}_1} g_a$$

$$= 0 = w_1 - 2\ddot{w}_2 \qquad ①$$

$$\frac{\delta g_a}{\delta w_2} - \frac{d}{dt}\frac{\delta}{\delta \dot{w}_2} g_a$$

$$= 0 = w_2 + 2p w_2 - 2\ddot{w}_1 \qquad ②$$

$$\delta g_a/\delta z - \frac{d}{dt}\,\delta g_a/\delta \dot{z} = 0$$

$$\Rightarrow \dot{p}(t) = 0 \qquad ③$$

ALSO, WE HAVE CONSTRAINS

$$\dot{z} = w_2^2$$

AND THE BOUNDRY CONDITIONS

$$z(t_0) = 0 \qquad z(t_f) = c$$

EXAMPLE: EXTREMIZE

$$J(x,u) = \int_{t_0}^{t_f} \tfrac{1}{2}\left[x_1^2 + x_2^2\right] dt$$

SUBJECT TO: $\dot{x}_1 = -x_1 + x_2 + u$

$$\dot{x}_2 = -2x_1 - 3x_2 + u$$

$$\int_{t_0}^{t_f} u^2(t)\, dt = C$$

SET $w_1 = x_1,\ w_2 = x_2,\ u = w_3$

$$\Rightarrow J(w) = \int_{t_0}^{t_f} \tfrac{1}{2}\left[w_1^2 + w_2^2\right] dt$$

$$\dot{w}_1 = -w_1 + w_2 + w_3\ ,\ \dot{w}_2 = -2w_1 - 3w_2 + w_3$$

$$C = \int_{t_0}^{t_f} u^2(t)\, dt$$

SETTING $\dot{z} = w_3^2$ GIVES

$$g_a = \tfrac{1}{2}w_1^2 + \tfrac{1}{2}w_2^2 + p_1\left[-w_1 + w_2 + w_3 - \dot{w}_1\right]$$
$$+ p_2\left[-2w_1 - 3w_2 + w_3 - \dot{w}_2\right] + p_3\left[w_3^2 - \dot{z}\right]$$

$$\Rightarrow \frac{\delta g_a}{\delta w} - \frac{d}{dt}\frac{\delta g_a}{\delta \dot{w}} = 0$$

$$w_1 - p_1 - 2p_2 + \dot{p}_1 = 0$$

$$w_2 + p_1 - 3p_2 - \dot{p}_2 = 0$$

$$p_1 + p_2 + 2p_3 w_3 = 0$$

$$\Rightarrow \frac{\delta g_a}{\delta z} - \frac{d}{dt}\,\delta g_a/\delta\dot{z} = 0$$

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

$\Rightarrow$ CONSTRAINTS

$$\dot{w}_1 = -w_1 + w_2 + w_3$$

$$\dot{w}_2 = -2w_1 - 3w_2 + w_3$$

$$\dot{z}(t) = w_3^2$$

ALSO

$$z(t_0) = 0,\quad z(t_f) = C$$

BANG BANG CONTROL

DIFFER. EQ ASSUMED LINEAR:

$$\dot{x} = f(x,u,t) = f(x,t) + G(x,t)u(t)$$

$$J = \Theta[x(t_f), t_f] + \int_{t_0}^{t_f} [\phi(x,t) + h^T(x,t)u]\,dt$$

HAMILTONIAN IS

$$\mathcal{H} = \phi + h^T u + \lambda^T [f + Gu]$$

USE P'S MINUMUM PRINCIPLE

$$\mathcal{H}[x, \hat{u}, t] \le \mathcal{H}[x, u, t]$$

$$\Rightarrow (h^T + \lambda^T G)\hat{u} \le (h^T + \lambda^T G)u$$

PUT BOUNDS ON CONTROL:

$$a_i \le u_i \le b_i$$

$$u_i = \begin{cases} a_i & ; (h^T + \lambda^T G)_i > 0 \\ b_i & ; (h^T + \lambda^T G)_i < 0 \\ ? & ; (h^T + \lambda^T G)_i = 0 \quad \left(\begin{smallmatrix}\text{SINGULAR} \\ \text{PROB}\end{smallmatrix}\right) \end{cases}$$

IN ORDER TO SOLVE, WE MUST SOLVE

$$\dot{x} = f + Gu = \delta\mathcal{H}/\delta\lambda$$

$$\dot{\lambda} = -\delta\mathcal{H}/\delta x$$

MINIMUM TIME PROBLEM

$$\begin{cases} \dot{x} = Ax + bu \\ J = \int_{t_0=0}^{t_f} dt = t_f \\ -1 \le u \le 1 \end{cases}$$

$$\mathcal{H} = 1 + \lambda^T(Ax + bu)$$

P'S MINIMUM PRINCIPLE IS THUS

$$\lambda^T b \hat{u} \le \lambda^T b u$$

$$\Rightarrow \hat{u} = -\mathrm{sgn}(\lambda^T b)$$

THUS $\dot{x} = Ax + bu = Ax - b\,\mathrm{sgn}(\lambda^T b)$; $x(0) = 0$

SOLVE $\dot{\lambda} = -\frac{\delta \mathcal{H}}{\delta x} = -A^T \lambda$

$$\Rightarrow \lambda(t) = e^{-A^T(t - t_f)} \lambda(t_f)$$

LET $\tau = t_f - t$ AND $x(t) = x(-\tau + t_f) = \xi(\tau)$

THEN $\dot{x} = A\xi - b\,\mathrm{sgn}\,\lambda^T b$

$$= A\xi - b\,\mathrm{sgn}\{\lambda^T(t_f)\, e^{A^T t}\, b\}$$

SOLUTION IS

$$\xi(t) = \int_0^{\tau} e^{-A(\tau - s)} b\,\mathrm{sgn}\left[\lambda^T(t_f)\, e^{AS}\, b\right] ds$$

EXAMPLE

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \begin{bmatrix}0\\ 1\end{bmatrix} u(t) \qquad x(t_0) = x_0$$

$$x(t_f) = 0$$

$$J = \int_{t_0}^{t_f} dt \qquad -1 \leq u \leq 1$$

$$\hat{u} = -\operatorname{sgn} \lambda^T b = -\operatorname{sgn} \underline{\lambda_1\,\lambda_2}\ \begin{matrix}0\\ 1\end{matrix}\Big] = -\operatorname{sgn} \lambda_2$$

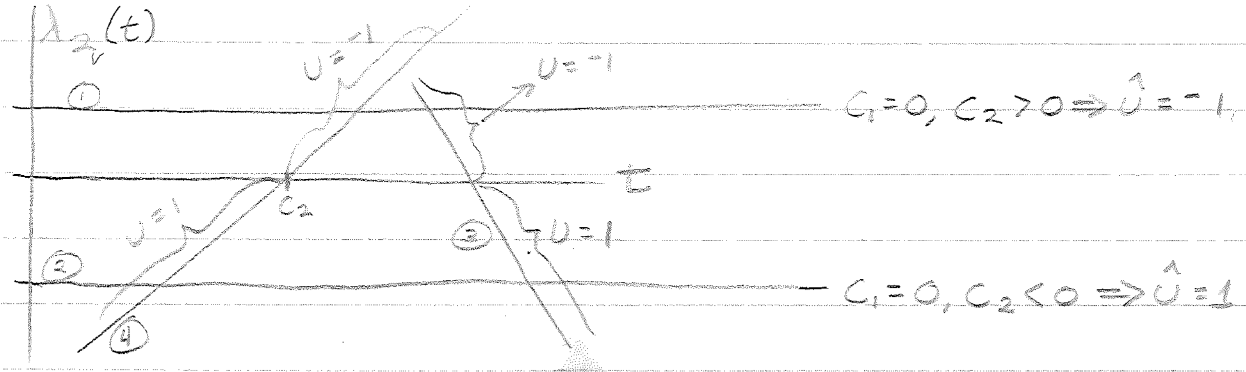
NOW $\dot{x}_1 = x_2$

$$\dot{x}_2 = u = \pm 1 \Rightarrow x_2(t) = \pm t + C_3 \Rightarrow x_1 = \pm\tfrac{1}{2}t^2 + C_3 t + C_4$$

COSTATE EQUATIONS:

$$\lambda' = -\delta \mathcal{H}/\delta x = -A^T\lambda \quad ; \quad \mathcal{H} = 1 + \lambda^T(Ax + bu)$$

$$\begin{bmatrix}\lambda_1'\\ \lambda_2'\end{bmatrix} = -\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}\begin{bmatrix}\lambda_1\\ \lambda_2\end{bmatrix} \Rightarrow \begin{cases}\dot{\lambda}_1 = 0 \Rightarrow \lambda_1 = C_1\\ \dot{\lambda}_2 = -\lambda_1 \Rightarrow \lambda_2 = -C_1 t + C_2\end{cases}$$



THE FOUR POSSIBLE OPTIMAL CONTOLS ARE:

① $\hat{u} = -1$, ② $\hat{u} = 1$, ③ $\hat{u}$ FROM -1 TO 1 ④ $\hat{u}$ FROM $+1$ TO -1

NOW $x_1 = \pm\frac{1}{2}t^2 + C_3 t + C_4$; $\frac{1}{2}x_2^2 = \frac{1}{2}t^2 \pm C_3 t + C_3^2$

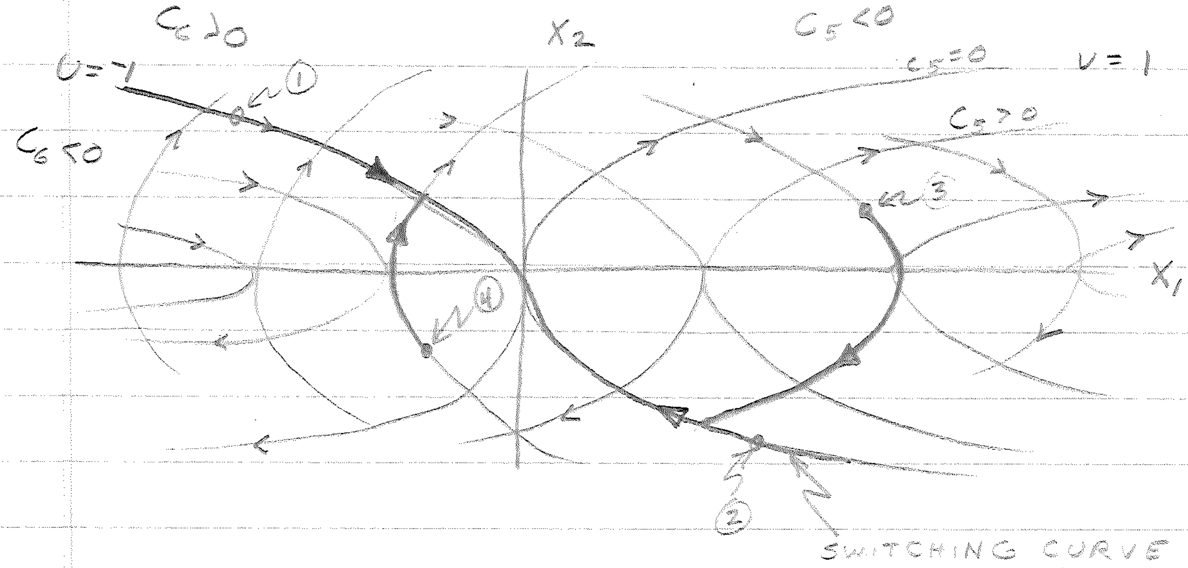
FOR $\hat{u} = 1$, USE "+" SIGN $\Rightarrow x_1^2 = \frac{1}{2}x_2^2 + C_5$

FOR $\hat{u} = -1$, USE "−" SIGN $\Rightarrow x_1 = -(\frac{1}{2}t^2 - C_3 t) + C_4$

$$= -(\tfrac{1}{2}x_2^2 - \tfrac{1}{2}C_3^2) + C_4$$

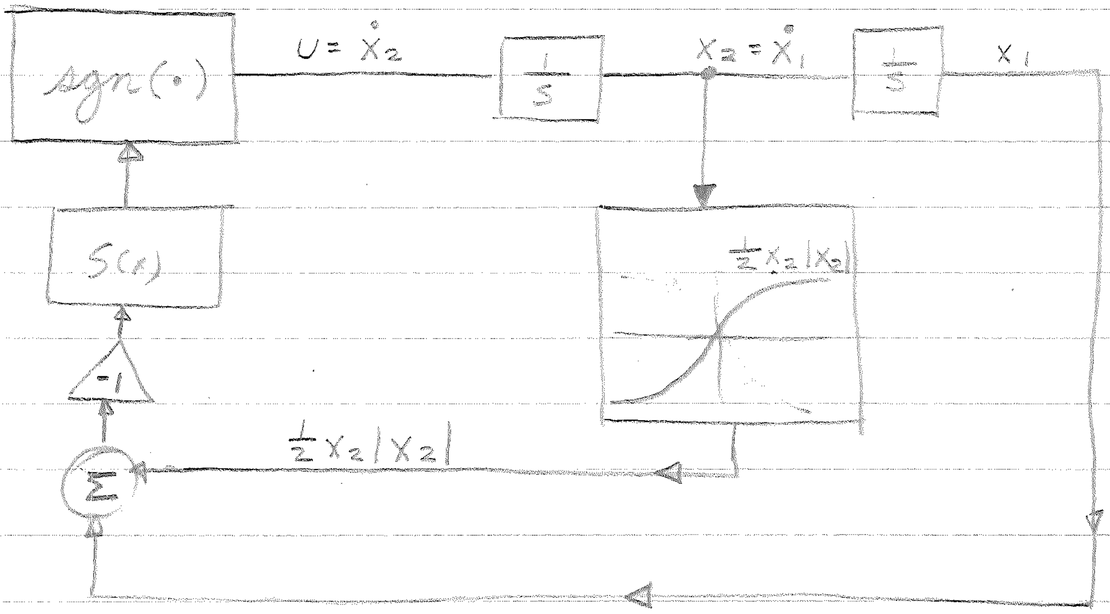
$$= -\tfrac{1}{2}x_2^2 + C_6$$

PLOT THESE BABIES $\Longrightarrow$



THE FOUR CASES ALLOWED BY THE λ'S CORRESPOND TO DIFFERENT CURVES AS SHOWN. THE SWITCHING CURVE IS GIVEN BY $x_1 = -\frac{1}{2}|x_2|x_2$

IMPLEMENTATION: $\dot{x}_1 = x_2$, $\dot{x}_2 = U$



MINIMUM FUEL PROBLEM

$$\begin{cases} \dot{x} = f(x,t) + G(x,t)\,u & ; x(t_0) = x_0 \quad x(t_f) = 0 \\ J = \int_{t_0}^{t_f} \left[k + \Sigma_{i=1}^{m} c_i |u_i| + \phi \right] dt \\ -1 \leq u_i \leq 1 \end{cases}$$

$$\mathcal{H} = k + \Sigma_{i=1}^{m} c_i |u_i| + \phi + \lambda^T (f + Gu)$$

APPLYING MINIMUM PRINCIPLE

$$\Sigma c_i |\hat{u}_i| + \lambda^T G \hat{u} \leq \Sigma c_i |u_i| + \lambda^T G u$$

NOW

$$\lambda^T G \hat{u} = \underline{\lambda_1 \lambda_2 \dots \lambda_n} \begin{bmatrix} g_{11} & \cdots & g_{1m} \\ \vdots & & \vdots \\ g_{n1} & \cdots & g_{nm} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}.$$

DENOTE COLUMNS OF G BY $g_1 \dots g_m$

$$\Rightarrow \lambda G \hat{u} = \underline{\lambda_1 \dots \lambda_n}\ \underline{g_1 \dots g_m} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}$$

$$= \lambda^T \Sigma_{i=1}^{m} g_i u_i$$

SO:

$$\Sigma_{i=1}^{m} c_i |\hat{u}_i| + \lambda^T \Sigma_{i=1}^{m} g_i \hat{u}_i \leq \Sigma_{i=1}^{m} c_i |u_i| + \lambda^T \Sigma_{i=1}^{m} g_i u_i$$

APPLY MIN PRINCIPLE TERMWISE:

$$c_i |\hat{u}_i| + \lambda^T g_i \hat{u}_i \geq c_i |u_i| + \lambda^T g_i u_i$$

$$c_i |\hat{u}_i| + \lambda^T g_i \hat{u}_1 = \begin{cases} (c_1 + \lambda^T g_1)\, \hat{u}_i & ; \hat{u}_1 > 0 \\ (-c_i + \lambda^T g_1)\, \hat{u} & ; \hat{u}_1 < 0 \end{cases}$$

ASSUMING $c_i = 1$,

$$|u_i| + \lambda^T g_i u_i = \begin{cases} (1 + \lambda^T g_i)\, \hat{u}_1 & ; \hat{u}_1 > 0 \\ (-1 + \lambda^T g_i)\, \hat{u}_i & ; \hat{u}_1 < 0 \end{cases}$$

WE WISH TO MINIMIZE THIS EXPRESSION.

TO DO SO, WE MUST CONSIDER SOME SPECIAL POLARITY CONDITIONS ON THE NUMBER $\lambda^T g_i$ $\Longrightarrow$

(i) LET $\lambda^T g_i > 1$

$$|u_i| + \lambda^T g_i u_i = \begin{cases} (1+\lambda^T g_i)\hat{u} \geq 0 & \text{FOR } \hat{u}_i \geq 0 \\ (-1+\lambda^T g_i)\hat{u} \leq 0 & \text{FOR } \hat{u}_i < 0 \end{cases}$$

TO MINIMIZE, WE WOULD CHOOSE $\hat{u}_i$ AS SMALL AS POSSIBLE $\Longrightarrow \hat{u}_i = -1$

(ii) $\lambda^T g_i = 1$

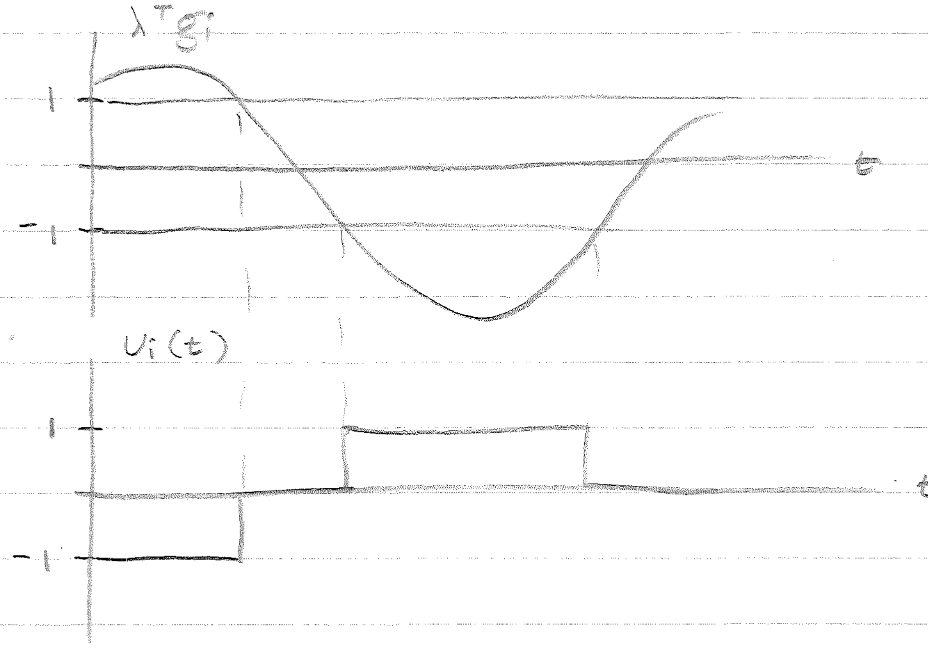
$$\left.\begin{aligned} &|u_i| + \lambda^T g_i u_i = \begin{cases} 2\hat{u}_i > 0 & \text{FOR } u_i > 0 \\ 0 & \text{FOR } u < 0 \end{cases} \\ &\text{CHOOSE } u_i \leq 0 \text{ FOR MINIMIZATION (?)} \end{aligned}\right\} \text{NO}$$

(iii) $0 < \lambda^T g < 1 \Rightarrow u_i = 0$

(iv) $-1 < \lambda^T g < 0 \Longrightarrow u_i = 0$

(v) $\lambda^T g_i < -1 \Longrightarrow u_i = 1$

(vi) $\lambda^T g_i = 1 \Longrightarrow u_i \geq 0$



THIS

"BANG-OFF-BANG" CONTROL;

$$\cancel{u = \begin{cases} 1 & ; \lambda_2 < -1 \\ 0 & ; |\lambda_2| < 1 \\ -1 & ; \lambda_2 > 0 \end{cases}}$$

EXAMPLE

$$J = \int_{t_0}^{t_f} \left[\tfrac{1}{2} x Q^T x + |u| \right] dt \qquad ; -1 < u < 1$$

$$\dot{x} = Ax + bu \qquad x(t_0) = x_0 \qquad x(t_f) = 0$$

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad b = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad Q = \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix}$$

TURNS OUT

$$\mathcal{H} = \tfrac{1}{2} q_1^2 x_1^2 + \tfrac{1}{2} q_2^2 x_2^2 + |u| + \lambda_1 x_1 + \lambda_2 u$$

HERE $\quad G = b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$$\lambda^T g_1 = \begin{bmatrix} \lambda_1 & \lambda_2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \lambda_2$$

THUS $\quad u = \begin{cases} 1 & ; \ \lambda_2 < -1 \\ 0 & ; \ -1 < \lambda_2 < 1 \\ -1 & ; \ \lambda_2 > 1 \end{cases}$

FOR $u = \pm 1$, WE HAVE THE SAME PROBLEM AS BEFORE

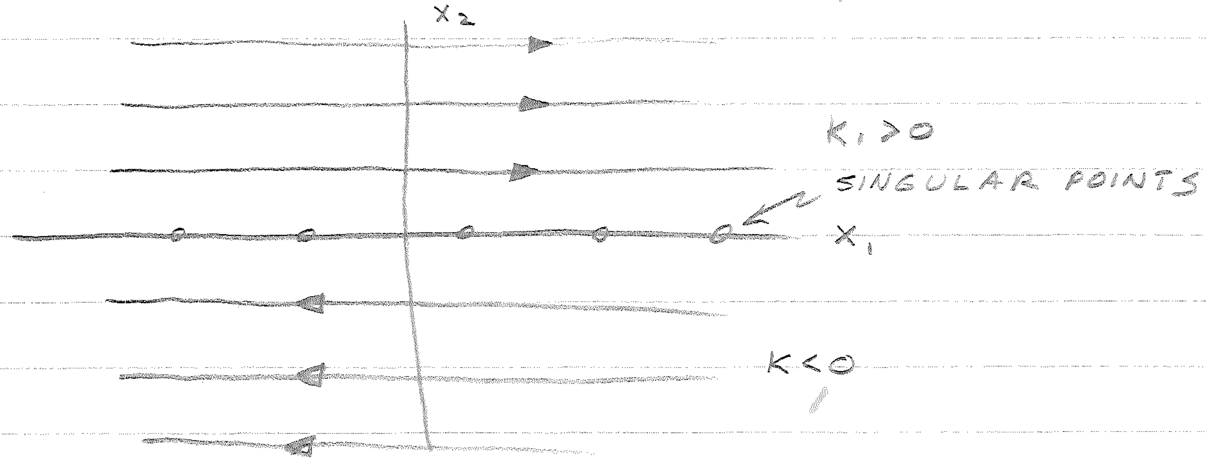
FOR $u = 0$, WE HAVE TO PROCEED. NOW

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ u \end{bmatrix} \qquad x_1(t_f) = x_2(t_f) = 0$$

$$-\frac{\delta \mathcal{H}}{\delta x} = \begin{bmatrix} \dot{\lambda}_1 \\ \dot{\lambda}_2 \end{bmatrix} = -\begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$$

NOW $u = 0 \Rightarrow \dot{x}_2 = u \Longrightarrow x_2 = K_1$

$$\dot{x}_1 = x_2 = K_1 \Longrightarrow x_1 = K_1 t + K_2$$



SINGULAR POINTS OCCUR WHEN $K_1 = 0$.

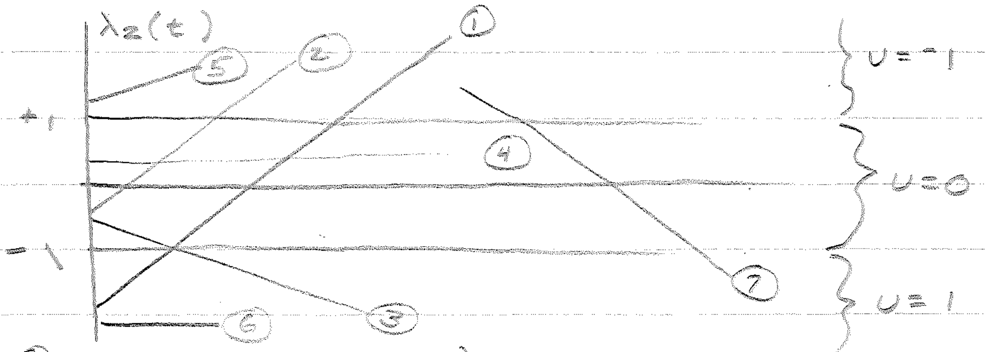
WE MUST NOW SOLVE (IN GENERAL) FOR λ'S.

LET $Q = 0$

$\Rightarrow \dot{\lambda}_1 = 0 \Rightarrow \lambda_1 = q_1$

$\dot{\lambda}_2 = \lambda_1 \Rightarrow \lambda_2 = q_1 t + q_2$

OUR CONTROL IS DETERMINED BY λ_2 WHICH IS A STRAIGHT LINE:



① ⇒ $U = (1, 0, -1)$

② ⇒ $U = (0, -1)$

③ ⇒ $U = (0, 1)$

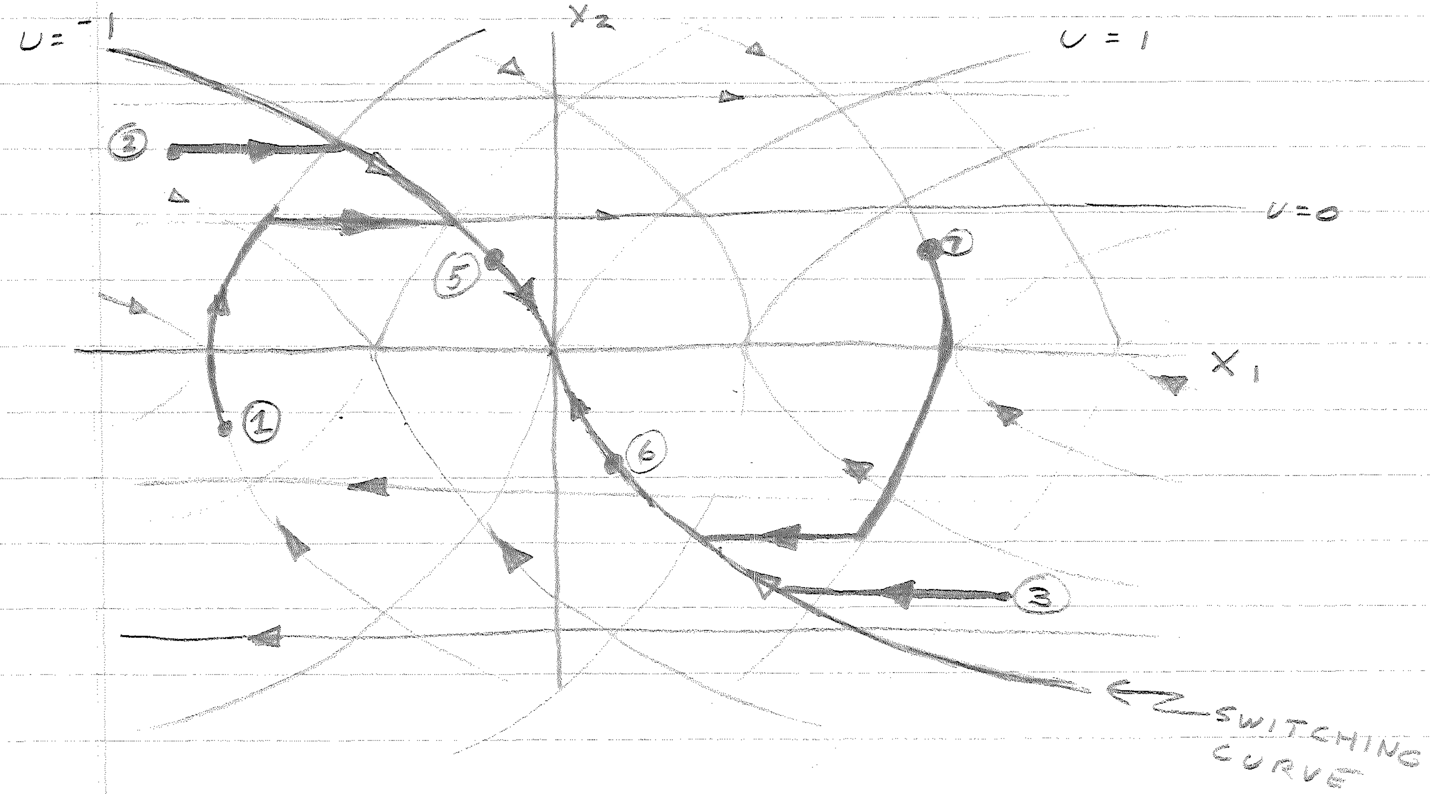
④ ⇒ $U = 0$ ← WON'T WORK DUE TO STATIONARY POINTS

⑤ $U = -1$

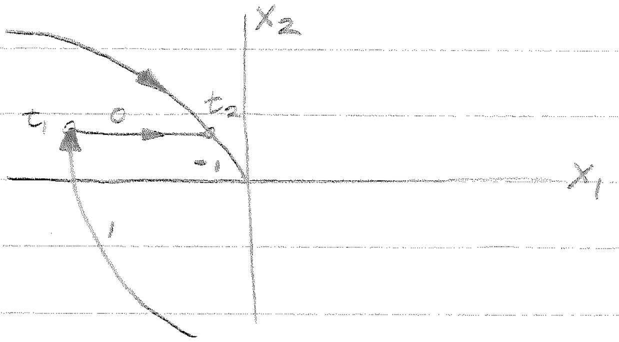
⑥ $U = +1$

⑦ $U = (-1, 0, 1)$

NOTE: <u>CAN'T</u> HAVE, FOR EX, $U = (1, 0)$



CONSIDER THE CONTROL $u = 1, 0, -1$



WE WISH TO SHOW THAT $t_2 - t_1 = \infty$

@ $t = t_2$ WE HAVE $x_1(t_2) = -\frac{1}{2} x_2^2(t_2)$

$$x_2(t_2) = x_2(t_1) = k_1$$

@ $t = t_1$ $\quad x_1(t_1) = k_1 t_1 + k_2 = x_2(t_1)\, t_1 + k_2$

$$\Rightarrow k_2 = x_1(t_1) - x_2(t_1)\, t_1$$

THUS $x_1(t) = k_1 t + k_2$

$$= x_2(t_1)\, t + x_1(t_1) - x_2(t_1)\, t_1$$

$$= x_1(t_1) + x_2(t_1)(t - t_1)$$

RECALL THAT $\begin{bmatrix} \dot\lambda_1 \\ \dot\lambda_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$ FOR $q = 0$

$$\Rightarrow \dot\lambda_1 = 0 \Rightarrow \lambda_1 = a$$

$$\dot\lambda_2 = -\lambda_1 \Rightarrow \lambda_2 = -at + b$$

ALSO $\quad u = \begin{cases} -1 & ; \lambda_2 > 1 \\ 0 & ; |\lambda_2| < 1 \\ 1 & ; \lambda_2 < -1 \end{cases}$

@ $t = t_1$, u GOES FROM 1 TO 0 $\Rightarrow \lambda_2(t_1) = -at_1 + b = -1$

@ $t = t_2$, u GOES FROM 0 TO -1 $\Rightarrow \lambda_2(t_2) = -at_2 + b = 1$

SOLVING FOR a:

$$-2a(t_2 - t_1) = 2 \Rightarrow t_2 - t_1 = -1/a$$

NOW $\mathcal{H} = |u| + \lambda^T (Ax + bu)$

FOR $u = 0$

$$\mathcal{H} = \lambda^T Ax = \begin{bmatrix} \lambda_1 & \lambda_2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda_1 x_2 = a x_2$$

FOR $u=0$, $\mathcal{H}$ IS EXPLICITLY TIME INDEPENDENT. THUS, IT'S OPTIMUM TRAJECTORY MUST MAKE $\mathcal{H} = 0$. THIS GIVES $a=0$. BUT

$$t_2 - t_1 = \frac{-2}{a} = \infty$$

THIS PROBLEM CAN BE ALLEVIATED BY USING $J = \int_{t_0}^{t_f} (|u) + k)\,dt$

THE SINGULAR PROBLEM

- BANG-BANG CONTROL

$$\dot{x} = Ax + bu$$

$$\Rightarrow \hat{u} = -\operatorname{sgn} \lambda^T b = \begin{cases} 1 & ; \lambda^T b < 0 \\ -1 & ; \lambda^T b > 0 \\ ? & \lambda^T b = 0 \end{cases}$$

EXAMPLE

$$J = \frac{1}{2} \int_0^2 x^2 dt \qquad t_f = 0 \ , \ x(t_f) = 0$$

$$\dot{x} = u$$

THEN OUR HAMILTONIAN IS

$$\mathcal{H} = \frac{1}{2} x^2 + \lambda^T u \qquad ; \quad \hat{u} = -\operatorname{sgn} \lambda$$

FOR $\lambda = 0$, $\mathcal{H} = \frac{1}{2} x^2$

$$\lambda = \dot{\lambda} = \ddot{\lambda} = 0$$

$$\left.\begin{cases} \dot{x} = \dfrac{\delta \mathcal{H}}{\delta \lambda} = u \\ \dot{\lambda} = -\dfrac{\delta \mathcal{H}}{\delta x} = -x \end{cases}\right\} \text{IN GENERAL} \qquad \begin{matrix} x(0) = 1 \\ x(2) = 0 \end{matrix}$$

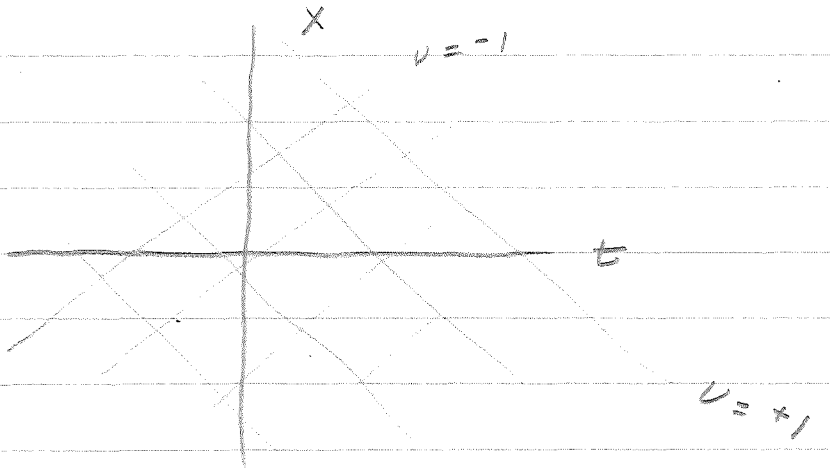
ASSUME $u = -1$ AND $\lambda > 0$

$$\Rightarrow \dot{x} = u \Rightarrow x = -t + C$$

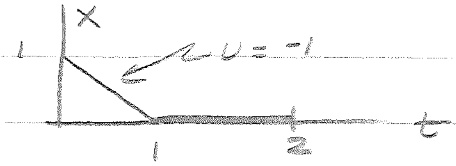
ASSUME $u = 1$ AND $\lambda < 0$

$$\Rightarrow \dot{x} = u = 1 \Rightarrow x = t + C_2$$

THUS, WE HAVE



FOR $x(0)=1$, WE HAVE



$u=-1$, $\dot{x}=-t+C$ $x(0)=1 \Rightarrow C=1$

$x=-t+1$, $\dot{\lambda}=-x=t-1 \Rightarrow \lambda(t)=\frac{1}{2}t^2-t+C'$

$C'=\lambda(0) \Rightarrow \lambda(t)=\frac{1}{2}t^2-t+\lambda(0)$

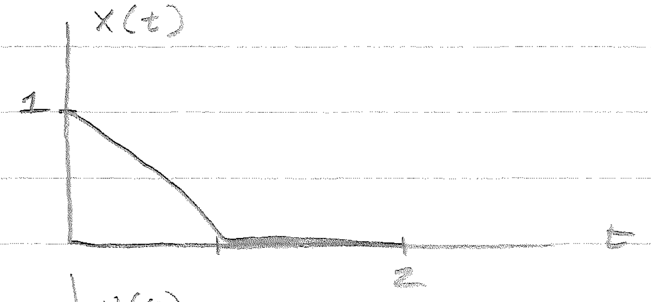
DURING THE SINGULARITY

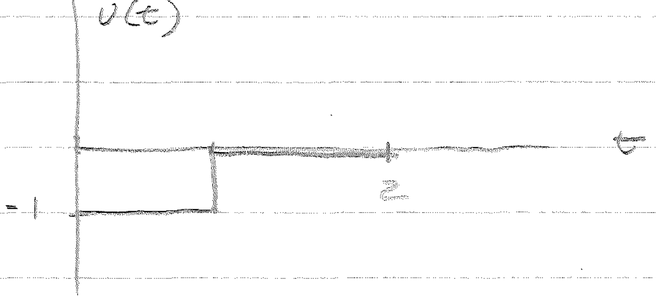
$\lambda=0 \Rightarrow \dot{\lambda}=0 \Rightarrow \dot{\lambda}=0$

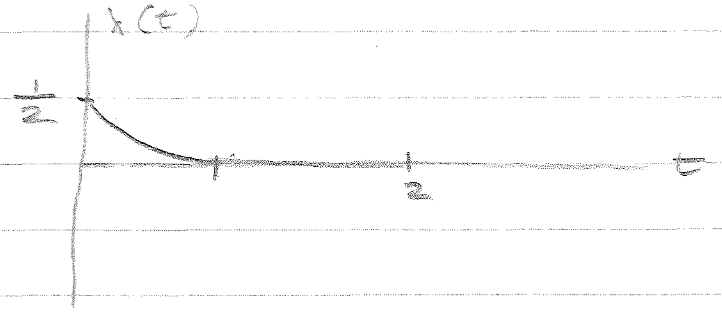
$\dot{\lambda}_1=-x=0$ $\ddot{\lambda}=0=-\dot{x}=u$

$\lambda(1)=0 \Rightarrow \lambda(0)=\frac{1}{2}$

OUR SOLUTION IS THUS:







BANG - BANG

$\dot{x} = Ax + Bu$

$J = \frac{1}{2} x^T(t_f) S x(t_f) + \int_{t_0}^{t_f} x^T Q x \, dt$

$\mathcal{H} = x^T Q x^T + \lambda^T (Ax + Bu)$

FOR OPTIMAL CONTROL $\lambda^T B \hat{u} \leq \lambda^T B u$

$\Rightarrow \hat{u} = -\mathrm{sgn}\, \lambda^T B$

FOR SINGULAR PROBLEM $\frac{\delta \mathcal{H}}{\delta u} = B^T \lambda = 0$

WHAT IS U?

EXAMPLE $J = \frac{1}{2} \int_0^{t_f} x_1^2 \, dt$; t_f IS FIXED

$$① \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{matrix} x_1 \\ x_2 \end{matrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} u \qquad \begin{matrix} x_1(0) = x_{10} \\ x_2(0) = x_{20} \end{matrix}$$

$x_1(t_f) = x_2(t_f) = 0$

$$\mathcal{H} = \frac{1}{2} x_1^2 + \underline{\lambda_1 \ \lambda_2} \begin{bmatrix} x_1 + u \\ -u \end{bmatrix}$$

$$= \frac{1}{2} x_1^2 + \lambda_1 x_2 + (\lambda_1 - \lambda_2) u$$

$\frac{\delta \mathcal{H}}{\delta u} = 0 = \lambda_1 - \lambda_2$

$$-\frac{\delta \mathcal{H}}{\delta x} = \dot{\lambda} = [-x_1 \ -\lambda_1] \Rightarrow \begin{bmatrix} \dot{\lambda}_1 \\ \dot{\lambda}_2 \end{bmatrix} = \begin{bmatrix} -x_1 \\ -\lambda_1 \end{bmatrix} ②$$

WE HAVE A SINGULAR SOLUTION

WHEN $\lambda^T B = \lambda_1 - \lambda_2 = 0$ OR $\lambda_1 = \lambda_2$ ③

THUS $\dot{\lambda}_1 = \dot{\lambda}_2$ AND $\ddot{\lambda}_1 = \ddot{\lambda}_2$

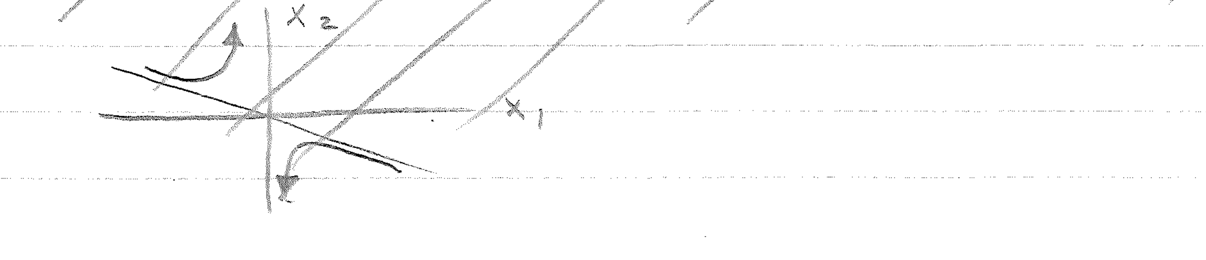
NOW, FROM ② $\dot{\lambda}_1 = -x$ & $\ddot{\lambda}_2 = -\dot{\lambda}_1 = \dot{x}_1$

THUS, FROM THE STATE EQUATIONS

$\dot{x}_1 = x_2 + u = -\dot{\lambda}_1 = -x_1 \Rightarrow u = -x_1 - x_2$

AT $\lambda_1 = \lambda_2$, $\mathcal{H} = C = \frac{1}{2} x_1^2 + \lambda_1 x_2 = \frac{1}{2} x_1^2 + x_1 x_2$

THIS IS OUR "SINGULAR ARC"



AT $\lambda_1 = \lambda_2$, $\mathcal{H}$ IS NOT AN EXPLICIT FUNCTION OF TIME. THUS

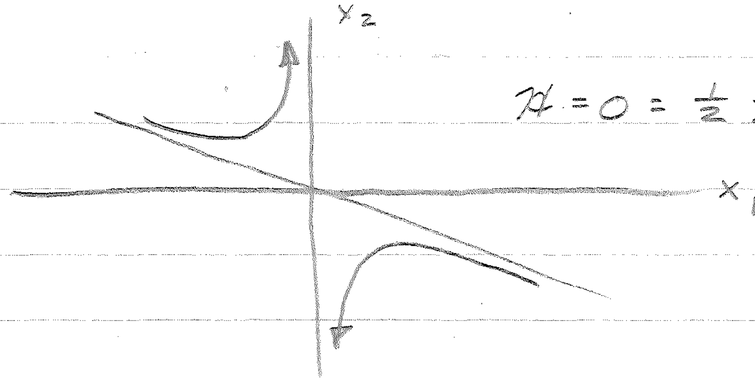
$$\mathcal{H} = C = \tfrac{1}{2} X_1^2 - \lambda_1 X_2$$
$$= \tfrac{1}{2} X_1^2 + X_1 X_2 \Longleftarrow \text{SINGULAR ARC}$$

ON THE SINGULAR ARC

$$\dot{x}_1 = x_2 + U = X_2 - (x_1 + x_2) = {}^{-}X_1$$
$$\Rightarrow x_1(t) = x_1(t_1) e^{-(t-t_1)}$$

SOLVING FOR X_2 GIVES $\dot{X}_2 = x_1 + x_2 = U$

$$\Rightarrow x_2(t) = K_1 e^{(t-t_1)} - \tfrac{1}{2} x_1(t_1) e^{-(t-t_1)}$$
$$K_1 = X_2(t_1) + \tfrac{1}{2} x_1(t_1)$$
$$\Rightarrow X_2(t) = [X_2(t_1) + \tfrac{1}{2} X_1(t_1)] e^{(t-t_1)}$$
$$- \tfrac{1}{2} X_1(t_1) e^{-(t-t_1)}]$$



$$\mathcal{H} = 0 = \tfrac{1}{2} x_1 + x_2 = 0$$

CONSIDER NOW LIMITING $U = \pm K$

$\dot{X}_1 = x_2 \pm K$ $\quad \dot{x}_2 = \mp K \Rightarrow x_2 = {}^{-}K t + X_{20}$ FOR $K>0$

GIVES $\begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix} = \begin{bmatrix} -\frac{K}{2} t^2 + (X_{20} + K) t + X_{10} \\ -Kt + X_{20} \end{bmatrix}$

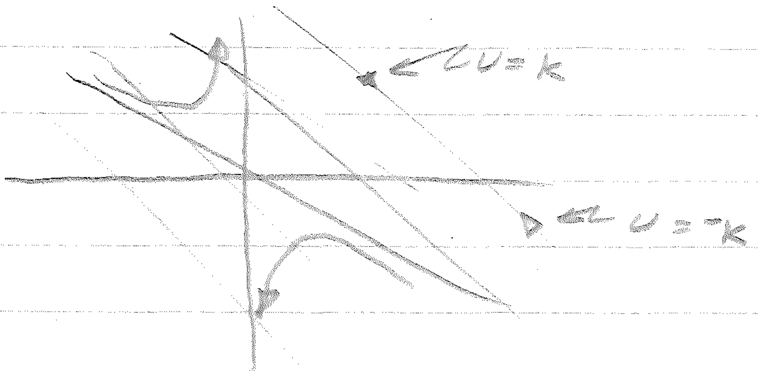
IN TERMS OF X_2, $t = {}^{X_{20} - X_2}/_{K}$

THUS

$$X_1 = \frac{-K}{2}\left(\frac{X_{20}-X_2}{K}\right)^2 + \frac{(X_{20}+K)(X_{20}-X)}{K} + X_{10}$$

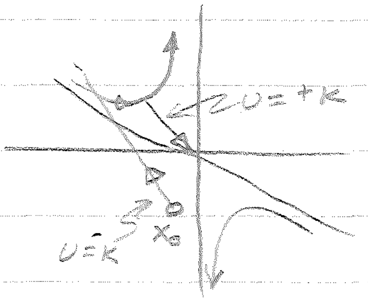
FOR LARGE K

$$X_1 = (X_{20}-X_2) + X_{10}$$

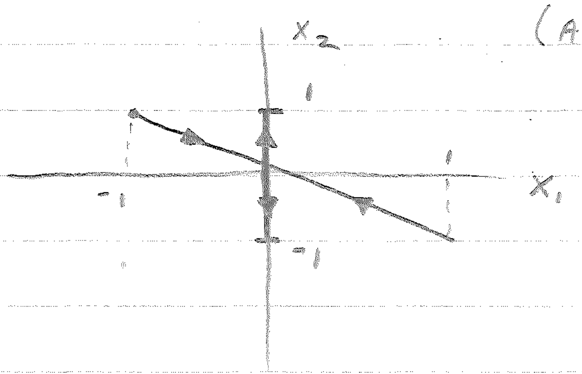


FOR U=-K,
GET SAME
SOL. WITH
DIFFERENT
DIRECTION

EXAMPLE



LET $t_f \to \infty$, OUR SINGULAR ARCS BECOME
(ASSUME K= 1)



WE WISH TO AVOID THE VERTICAL STRIP.

QUASILINEARIZATION

CONSIDER

$$\dot{x} = a_{11}x + a_{12}\lambda + e_1 \qquad x(t_0) = x_0$$

$$\dot{\lambda} = a_{21}x + a_{22}\lambda + e_2 \qquad \lambda(t_f) = \lambda_f$$

THESE ARE DERIVED FROM HAMILTONIAN:

CONSIDER HOMOGENEOUS SOLUTIONS

$$\dot{X}^H = a_{11}x^H + a_{12}\lambda^H$$

$$\dot{\lambda}^H = a_{21}x^H + a_{22}\lambda^H$$

WLOG, ASSUME $x^H(0)=0$, $\lambda^H(0)=1$

THE GENERAL SOLUTION WOULD BE:

$$X = C_1 X^H + X_p \qquad X_p(0) = X_0$$

$$\lambda = C_1\lambda_H + \lambda_p \qquad \lambda_p(0) = 0$$

NOW

$$\lambda(t_f) = \lambda_f = C_1\lambda_H(t_f) + \lambda_p(t_f)$$

$$\Longrightarrow C_1 = \frac{\lambda_f - \lambda_p(t_f)}{\lambda_H(t_f)}$$

ALSO $X(t_0) = X_0 = C_1 \underset{0}{\cancel{X^H}} + X_p(t_0)$.

SO X B.C. IS SATISFIED $\forall$ C

NOTE: USING

$$X(t_0) = X_0 \qquad \lambda(t_f) = \lambda_f$$

$$X_H(t_0) = 0 \qquad \lambda_H(t_0) = 1$$

$$X_p(t_0) = X_0 \qquad \lambda_p(t_0) = 0$$

WE CAN STRAIGHTFORWARDLY SOLVE FOR x & λ USING PRINCIPLE OF SUPERPOSITION

LINEARIZATION OF REDUCED STATE/COSTATE EQUATIONS

$$\dot{x} = f[x, \lambda, t] \Longleftarrow \text{STATE} \qquad x(t_0) = x_0$$

$$\dot{\lambda} = g[x, \lambda, t] \Longleftarrow \text{COSTATE} \qquad \lambda(t_f) = \lambda_f$$

LET $x^0(t)$ ∮ $\lambda^0(t)$ BE AN INITIAL GUESS, TAYLOR SERIES ABOUT THEM GIVES

$$\dot{x}^{(1)}(t) = f[x^0, \lambda^0, t] + \frac{\delta f}{\delta x}\Big|_{x^0}[x^{(1)} - x^{(0)}] + \frac{\delta f}{\delta \lambda}\Big|_{\lambda_0}[\lambda^{(1)} - \lambda^{(0)}] + HOT$$

$$\dot{\lambda}^{(1)}(t) = g(x^0, \lambda^0, t) + \frac{\delta g}{\delta x}\Big|_{x_0}[x^{(1)} - x^{(0)}] + \frac{\delta g}{\delta \lambda}\Big|_{\lambda_0}[\lambda^{(1)} - \lambda^{(0)}] + HOT$$

REARRANGING

$$\dot{x}^{(1)}(t) = \left[\frac{df}{\delta x}\Big|_{x_0} x^{(1)}\right] + \left[\frac{\delta f}{\delta \lambda}\Big|_{\lambda_0} \lambda^{(1)}\right] + \left[f^0 - \frac{\delta f}{\delta x}\Big|_{x^0} x^0 - \frac{\delta f}{\delta \lambda}\Big|_{\lambda_0} \lambda^0\right]$$

$$\dot{\lambda}^{(1)}(t) = \left[\frac{\delta g}{\delta x}\Big|_{x_0} x^{(1)}\right] + \left[\frac{\delta g}{\delta \lambda}\Big|_{\lambda_0} \lambda^{(1)}\right] + \left[g^0 - \frac{\delta g}{\delta x}\Big|_{x_0} x^0 - \frac{\delta g}{\delta \lambda}\Big|_{\lambda_0} \lambda^0\right]$$

NOTE THESE ARE OF THE FORM

$$\begin{bmatrix} \dot{x}' \\ \dot{\lambda}' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x' \\ \lambda' \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

WE CAN SOLVE FOR $x^{(1)}$ AND $\lambda^{(1)}$ AND USE THEM TO FIND $x^{(2)}$ AND $\lambda^{(2)}$ ETC.

STOP ITERATION WHEN

$$\left\| \begin{bmatrix} x^{(K+1)} \\ \lambda^{(K+1)} \end{bmatrix} - \begin{bmatrix} x^{(K)} \\ \lambda^{(K)} \end{bmatrix} \right\| \leq \epsilon$$

EXAMPLE

$\dot{x} = x^2 + u \qquad x(0) = 3$

$J = \int_0^1 (2x^2 + u^2)\,dt$

$\Rightarrow \mathcal{H} = 2x^2 + u^2 + \lambda(x^2 + u)$

$\dot{\lambda} = -\frac{\delta \mathcal{H}}{\delta x} = -4x - 2\lambda x$

$\frac{\delta \mathcal{H}}{\delta u} = 0 = 2u + \lambda = 0 \Rightarrow u = -\lambda/2$

$$\Rightarrow \begin{cases} \dot{x} = x^2 - \lambda/2 \neq f; & x(0) = 3 \\ \dot{\lambda} = -4x - 2\lambda x = g; & \lambda(1) = 0 \end{cases}$$

$\dot{x}^{(1)} = x^{(0)^2}(t) + 2x^0[x^{(1)} - x^{(0)}] - \frac{1}{2}[\lambda^{(1)} - \lambda^{(0)}]$

$\dot{\lambda}^{(1)} = \lambda^{(0)} - [4 + 2\lambda^0][x^{(1)} - x^{(0)}] - 2x^0[\lambda^{(1)} - \lambda^{(0)}]$

OR,

$\dot{x}^{(1)} = 2x^{(0)}x^{(1)} - \frac{1}{2}p^{(1)} - x^{(0)^2}$

$\dot{\lambda}^{(1)} = -[4 + 2\lambda^0]x^{(1)} - 2x^0\lambda^{(1)} + 2x^0p^{(0)}$

QUASI-LINEARIZATION GENERALIZATION

BOTH x $\&$ λ ARE NOW n VECTORS.

THE STRAIGHTFORWARD GENERALIZATION IS

$$\dot{x}^{(i+1)} = f^{i} + \left.\frac{\delta f}{\delta x}\right|_{x^{(i)}} \left[x^{(i+1)} - x^{(i)}\right] + \left.\frac{\delta f}{\delta \lambda}\right|_{\lambda^{(i)}} \left[\lambda^{(i+1)} - \lambda^{(i)}\right]$$

$$\dot{\lambda}^{(i+1)} = - \left.\frac{\delta \mathcal{H}}{\delta x}\right|_{i} - \left.\frac{\delta^2 \mathcal{H}}{\delta x^2}\right|_{i} \left[x^{(i+1)} - x^{(i)}\right] - \frac{\delta^2 \mathcal{H}}{\delta x \delta \lambda} \left[\lambda^{(i+1)} - \lambda^{(i)}\right]$$

OR
$$\begin{bmatrix} \dot{x}^{(i+1)} \\ \dot{\lambda}^{(i+1)} \end{bmatrix} = A(t) \begin{bmatrix} \dot{x}^{(i+1)} \\ \dot{\lambda}^{(i+1} \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

THE HOMOGENEOUS BOUNDRY CONDITIONS ARE

$$\left.\begin{bmatrix} x_{h_i} \\ \lambda_{h_i} \end{bmatrix}\right|_{t=0} = \left[\begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \\ \hline 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{array}\right] \begin{array}{l} \uparrow \\ n \\ \downarrow \\ \\ \\ \\ \leftarrow i\text{TH POSITION FROM TOP} \\ \\ \\ \end{array}$$

THE GENERAL SOLUTION IS OF THE FORM

$$\begin{bmatrix} x \\ \lambda \end{bmatrix} = c_1 \begin{bmatrix} x_{h_1} \\ \lambda_{h_1} \end{bmatrix} + c_2 \begin{bmatrix} x_{h_2} \\ \lambda_{h_2} \end{bmatrix} + \cdots + c_n \begin{bmatrix} x_{h_n} \\ \lambda_{h_n} \end{bmatrix} + \begin{bmatrix} x_p \\ \lambda_p \end{bmatrix}$$

WHERE THE B.C. OF THE PARTICULAR SOLN' ARE

$$\begin{bmatrix} x_p(0) \\ \lambda_p(0) \end{bmatrix} = \left[\begin{array}{c} x_0 \\ \hline 0 \end{array}\right]$$

NOW, @ $t = t_0$:

$$\begin{bmatrix} x(0) \\ \lambda(0) \end{bmatrix} = c_1 \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ \hline 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{array}\right] + c_2 \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ \hline 0 \\ 1 \\ 0 \\ \vdots \end{array}\right] + \cdots + c_n \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ \hline 0 \\ 0 \\ \vdots \\ 1 \end{array}\right] + \begin{bmatrix} x_0 \\ 0 \end{bmatrix}$$

$$= \left[\begin{array}{c} x_0 \\ \hline c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{array}\right]$$

@ $t = t_f$,

$$\lambda(t_f) = c_1 \lambda_{h_1}(t_f) + \ldots + c_n \lambda_{h_n}(t_f) + \lambda_p(t_f)$$

EACH λ ENTRY IS AN n COLUMN.

$$\lambda(t_f) = \underbrace{\left[\lambda_{h_1}(t_f) \; \lambda_{h_2}(t_f) \cdots \lambda_{h_n}(t_f) \right]}_{\substack{\text{A SQUARE MATRIX} \\ \text{OF } \lambda = \Lambda}} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \lambda_p(t_f)$$

THUS

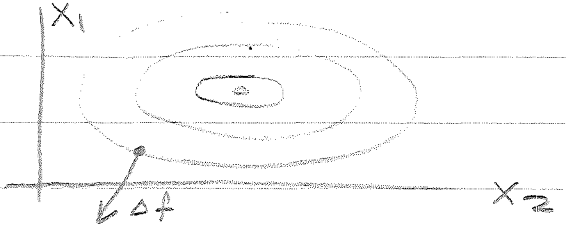
$$c = \Lambda_h^{-1} \left[\lambda(t_f) - \lambda_p(t_f) \right]$$

THE METHOD OF STEEPEST DESCENT

CONSIDER CALCULUS PROBLEM

MINIMIZE $Y = f(x_1, x_2)$

$$\frac{\delta Y}{\delta X_1} = \frac{\delta Y}{\delta X_2} = 0$$



GRADIANT : $\frac{\delta f}{\delta X} = \begin{bmatrix} \delta f / \delta x_1 \\ \delta f / \delta X_2 \end{bmatrix} = \Delta f$

WE WANNA GO IN OPPOSITE DIRECTION.

DEFINE THE UNIT VECTOR:

$$Z = \frac{\left.\frac{\delta f}{\delta x}\right|_{x_0}}{\sqrt{\left(\left.\frac{\delta f}{\delta x_1}\right|_{x_0}\right)^2 + \left(\left.\frac{\delta f}{\delta x_2}\right|_{x_0}\right)^2}}$$

THEN $\Delta Y = Y^{(1)} - Y^{(0)} = -\gamma Z[Y^{(0)}]$

OR $Y^{(1)} = Y^{(0)} - \gamma Z[Y^{(0)}]$

γ = STEP SIZE

MINIMIZATION OF FUNCTIONALS BY
STEEPEST DESCENT

PROBLEM IS

$$\dot{x}^{(i)}(t) = f\left[x^{(i)}, u^{(i)}, t\right]$$

$$\lambda^{(i)}(t) = -\delta \mathcal{H}/\delta x \Big|_i$$

BOUNDRY CONDITIONS:

$$x^{(i)}(t_0) = x_0 \qquad \lambda^{(i)}(t_f) = \frac{\delta h}{\delta x}_i$$

ASSUME THESE CONDITIONS HOLD, BUT

$$\delta \mathcal{H}/\delta u \Big|_i = 0$$ DOES <u>NOT</u>

NOW, IN GENERAL

$$\delta J_a = \frac{\delta h^T}{\delta x}\delta x(t_f) + \int_{t_0}^{t}\left[\left[\dot{\lambda} + \frac{\delta \mathcal{H}}{\delta x}\right]^T \delta x + \frac{\delta \mathcal{H}^T}{\delta u}\delta u + f\,\delta\lambda\right]dt$$

IF ALL BUT $\delta \mathcal{H}/\delta u$ ARE SATISFIED, THEN

$$\delta J_a = \int_{t_0}^{t_f} \delta \mathcal{H}/\delta u \; \delta u \; dt$$

LET

$$\delta u = u^{(i+1)} - u^{(i)} = -\tau\, \delta \mathcal{H}^{(i)}/\delta u \quad ; \tau > 0$$

THEN

$$\delta J_a = -\tau \int_{t_0}^{t_f} \frac{\delta \mathcal{H}}{\delta u}^{T(i)} \frac{\delta \mathcal{H}}{\delta u}^{(i)} dt \leq 0$$

EQUALITY ONLY IF $\frac{\delta \mathcal{H}}{\delta u}^{(i)} = 0$

STEEPEST DESCENT ALGORITHM

1. CHOOSE $U^{(0)}$

2. COMPUTE $X^{(0)}$ FROM

$$\dot{X}^{(0)} = f(X^{0}, U^{0}, t) \quad ; \quad X^{0}(0) = X_0$$

3. FORM $X^{0}(t_f)$

CALCULATE

$$\lambda(t_f) = \lambda_f^{0} = \left.\frac{\delta\Theta}{\delta X}\right|_{X^{0}(t_f)}$$

4. INTEGRATE COSTATE (BACKWARDS)

$$\dot{\lambda}^{(0)} = -\delta \mathcal{H}/\delta X \quad ; \quad \lambda(t_f) = \lambda_f$$

5. EVALUATE $\delta \mathcal{H}/\delta U$

IS $\| \cdot \|^2 \leq \gamma^2$. IF SO, WE'RE DONE. IF NOT...

(6). COMPUTE

$$U^{(1)} = U^{(0)} - \tau \left(\frac{\delta \mathcal{H}}{\delta U}\right)^{0}$$

GO TO STEP 2.

EXAMPLE

$\dot{x} = -x + u \qquad x(0) = 4$

$J = x^2(1) + \int_0^1 \frac{1}{2} u^2(t)\,dt$

(A LINEAR REGULATOR TYPE PROBLEM)

$\dot{\lambda}(t) = \lambda(t) = -\frac{\delta \mathcal{H}}{\delta x}$

$\mathcal{H} = \frac{1}{2}u^2 + \lambda((-x) + u)$

$\lambda(t_f) = \frac{\delta \Theta}{\delta x}\Big|_{t_f} = 2x(t_f) \Rightarrow \lambda(1) = 2x(1)$

$\frac{\delta \mathcal{H}}{\delta u} = u + \lambda = 0$

i. ASSUME $u^{(0)} = 1$

ii. $\dot{x}^0 = -x^0 + u^0 \quad ; x(0) = 4$

$\Rightarrow x^0 = 3e^{-t} + 1$

iii. $\lambda^0(t_f) = 2x(t_f) \Rightarrow \lambda^0(1) = 2(3e^{-1} + 1)$

iv. $\dot{\lambda}^0 = \lambda^0$

$\lambda^0(1) = 2(3e^{-1} + 1)$

$\Rightarrow \lambda^0 = ke^t = 2e^{-1}(3e^{-1} + 1)e^t$

v. $\frac{\delta \mathcal{H}}{\delta u}^0 = u^0 + \lambda^0$

$= 1 + 2e^{-1}(3e^{-1} + 1)e^t$

vi. $u^{(1)} = u^0 - \tau(\delta\mathcal{H}/\delta u)^0$

$= 1 - \tau[1 - 2e^{-1}(3e^{-1} + 1)e^t]$

$\tau = 1$

ETC

DIFFERENTIAL APPROXIMATION

$\dot{x} = f(x, u, a, t)$

ASSUME WE KNOW x AND WANT TO APPROXIMATE

a. PICK a TO MINIMIZE

$$J(a) = \frac{1}{2} \int_{t_0}^{t_f} \| \dot{x} - f(x, u, a, t) \|_R^2 dt$$

THUS

$$\frac{\delta J}{\delta a} = \int \frac{1}{2} \frac{\delta}{\delta a} \left[(\dot{x} - f(a))^T R (\dot{x} - f(a)) dt \right.$$

$$= - \int_{t_0}^{t_f} \frac{\delta f^T}{\delta a} R (\dot{x} - f) dt = 0$$

THUS, BY FUNDEMENTAL LEMMA OF CALCULUS OF VARIATIONS:

$$\int_{t_0}^{t_f} \frac{\delta f^T}{\delta a} R \dot{x} dt = \int_{t_0}^{t_f} \left(\frac{\delta f}{\delta a} \right)^T f dt$$

EXAMPLE

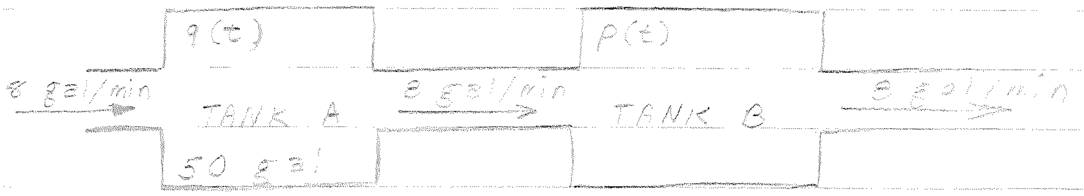
$\dot{x} = -a x^2 \qquad t \in [0, 4]$

$$\Rightarrow \int_0^4 (-x^2) \dot{x} dt = \int_0^4 (-x^2)(-a x^2) dt$$

$$\Rightarrow a = \frac{\int_0^4 (-x^2) \dot{x} dt}{\int_0^4 x^4 dt}$$

BOB MARKS
EE 5312
DUE 9/10/76

PG. 23 165
(1-1)



$q(t)$ = AMOUNT OF SALT IN TANK A ; $q(0) = 60$ lb
$p(t)$ = " " " " " B ; $p(0) = 0$

(a) $\frac{dq}{dt} = \frac{dq}{dV}\frac{dV}{dt}$

BUT $\frac{dV}{dt} = {}^-8$ gal/min

AND $\frac{dq}{dV} = {}^-q(t) \big/ 50$ gal (ie, ~~UNIFORM~~ SALT DENSITY)

THUS: $\dot{q} = \frac{-8}{50}q \quad \left(\frac{gal}{min}\right)$ ①

THE RATE OF CHANGE OF SALT IN TANK B IS THE SUM OF THAT ENTERING MINUS THAT LEAVING. THUS:

$$\dot{p} = \frac{-8}{50}p - \frac{dq}{dt}$$
$$= \frac{8}{50}q - \frac{8}{50}p$$

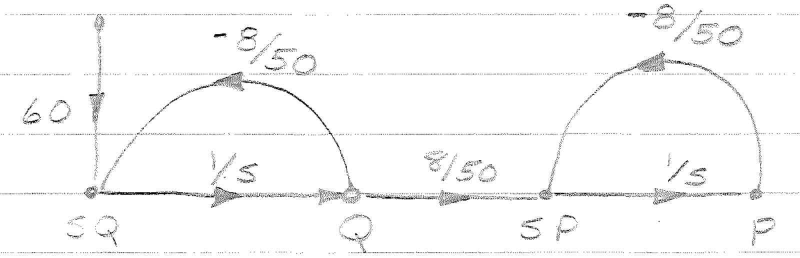
OUR STATE EQUATIONS ARE THUS:

$$\begin{bmatrix}\dot{q}\\ \dot{p}\end{bmatrix} = \begin{bmatrix}\frac{-8}{50} & 0\\ \frac{8}{50} & \frac{-8}{50}\end{bmatrix}\begin{bmatrix}q\\ p\end{bmatrix} \qquad \begin{matrix}q(t_0) = 60\\ p(t_0) = 0\end{matrix}$$

(b) SINCE WE HAVE MODELED THE SYSTEM AS LINEAR TIME INVARIANT, LET'S USE A SIGNAL FLOW GRAPH. LAPLACE TRANSFORMING THE STATE EQUATIONS:

$$\begin{bmatrix} sQ \\ sP \end{bmatrix} = \begin{bmatrix} -8/50 & 0 \\ \frac{8}{50} & \frac{-8}{50} \end{bmatrix} \begin{bmatrix} Q \\ P \end{bmatrix} + \begin{bmatrix} 60 \\ 0 \end{bmatrix}$$

WHICH GIVES:



(c) AGAIN, LET'S USE LAPLACE:

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

$$= \left\{ \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} - \begin{bmatrix} -8/50 & 0 \\ 8/50 & -8/50 \end{bmatrix} \right\}^{-1}$$

$$= \begin{bmatrix} s + 8/50 & 0 \\ -8/50 & s + 8/50 \end{bmatrix}^{-1}$$

$$\det[D] = (s + 8/50)^2 \ni D = \begin{bmatrix} s + \frac{8}{50} & 0 \\ -8/50 & s + \frac{8}{50} \end{bmatrix}$$

$$\text{adj } D = \begin{bmatrix} s + 8/50 & +8/50 \\ 0 & s + 8/50 \end{bmatrix}$$

$$[\text{adj } D]^T = \begin{bmatrix} s + 8/50 & 0 \\ +8/50 & s + 8/50 \end{bmatrix}$$

$$\therefore D^{-1} = \mathcal{L}[\Phi(t)] = \begin{bmatrix} \frac{1}{s + 8/50} & 0 \\ +\frac{8}{50} \frac{1}{(s + 8/50)^2} & \frac{1}{s + 8/50} \end{bmatrix}$$

SINCE

$$\mathcal{L}^{-1}\left[\frac{1}{s-a}\right] = e^{+at}$$

AND

$$\mathcal{L}^{-1}\left[\frac{1}{(s-a)^2}\right] = te^{at}$$

WE HAVE

$$\Phi(t) = \begin{bmatrix} e^{-\frac{8}{50}t} & 0 \\ +\frac{8}{50}te^{-\frac{8}{50}t} & e^{-\frac{8}{50}t} \end{bmatrix} = e^{At}$$

(d) NOW, FOR THE LINEAR TIME INVARIANT SYSTEM WITH NO INPUTS

$$x(t) = \Phi(t-t_0)\,x(t_0)$$

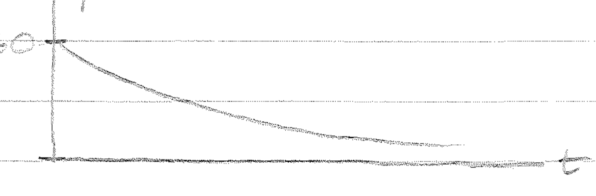
SETTING $t_0 = 0$, WE THUS HAVE

$$\begin{bmatrix} q(t) \\ p(t) \end{bmatrix} = \begin{bmatrix} e^{-8/50t} & 0 \\ +\frac{8}{50}te^{-\frac{8}{50}t} & e^{-\frac{8}{50}t} \end{bmatrix} \begin{bmatrix} 60 \\ 0 \end{bmatrix}$$

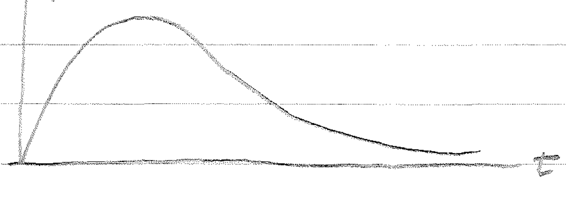
OR

$q(t) = 60e^{-8/50t}$ POUNDS OF SALT

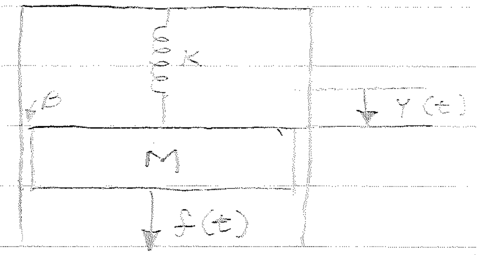
(t IN MINUTES)



$p(t) = \frac{+48}{5}te^{-\frac{8}{50}t}$ POUNDS OF SALT



pg. 24
(1-3)



(a) FORCE FROM MASS = $M \frac{d^2Y}{dt^2}$

FORCE FROM VISCOUS FRICTION = $B \frac{dY}{dt}$

FORCE FROM SPRING = KY

THUS

$$M \frac{d^2Y}{dt^2} + B\frac{dY}{dt} + KY = f(t)$$

OR $$\frac{d^2Y}{dt^2} = -\frac{B}{M}\frac{dY}{dt} - \frac{K}{M}Y + \frac{1}{M}f(t)$$

DEFINE THE STATE VARIABLES:

$$X_1 = Y \qquad X_2 = \frac{dY}{dt}$$

THEN

$$\begin{bmatrix} \dot{X}_1 \\ \dot{X}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{K}{M} & -\frac{B}{M} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{M} \end{bmatrix} f(t)$$

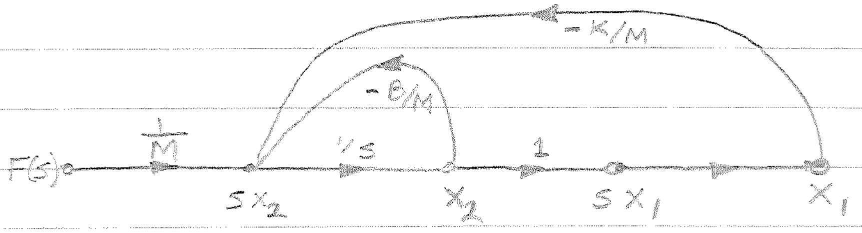
WHERE OUR CONTROL INPUT IS

$$U_1 = f(t)$$

(b) USING LAPLACE TRANSFORMS:

$$\begin{bmatrix} S X_1 \\ S X_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{K}{M} & -\frac{B}{M} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{M} \end{bmatrix} F(s)$$

THE SIGNAL-FLOW GRAPH IS THEN



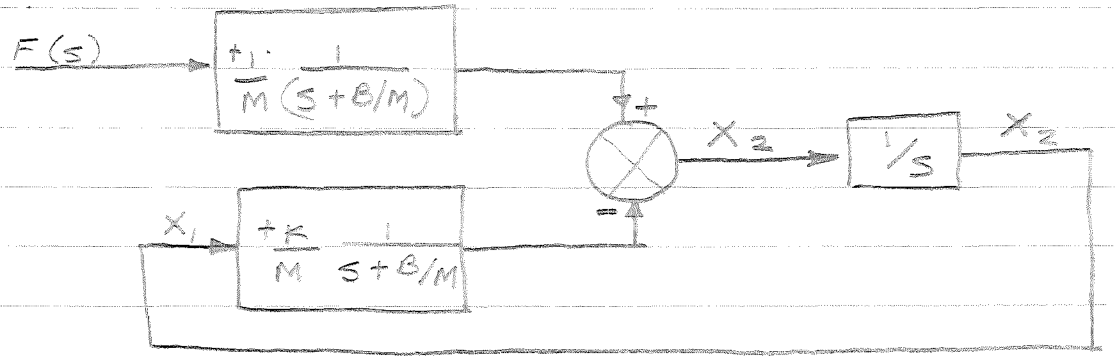
IF WE WISHED TO DO THE BLOCK DIAGRAM:

$$SX_1 = X_2 \Rightarrow X_1 = \frac{1}{S}X_2$$

$$SX_2 = -\frac{K}{M}X_1 - \frac{B}{M}X_2 + \frac{1}{M}F(S)$$

OR $$X_2\left(S + \frac{B}{M}\right) = -\frac{K}{M}X_1 + \frac{1}{M}F(S)$$

OR $$X_2 = -\frac{K}{M}\frac{X_1}{S+B/M} + \frac{1}{M}\frac{1}{S+B/M}F(S)$$



(c) $M = 1k \quad K = 2N/m \quad B = 2N/m/sec$

$$\Rightarrow \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} f(t)$$

$$D = SI - A = \begin{bmatrix} S & -1 \\ +2 & S+2 \end{bmatrix}$$

$$\det D = S(S+2) + 2 = S^2 + 2S + 2$$

$$= (s+1+j)(s+1-j)$$

$$\operatorname{adj} D = \begin{bmatrix} s+2 & -2 \\ +1 & s \end{bmatrix}$$

$$[\operatorname{adj} D]^T = \begin{bmatrix} s+2 & 1 \\ -2 & s \end{bmatrix}$$

$$\Rightarrow (sI - A)^{-1} = \begin{bmatrix} \frac{s+2}{(s+1+j)(s+1-j)} & \frac{1}{(s+1+j)(s+1-j)} \\ \frac{-2}{(s+1+j)(s+1-j)} & \frac{s}{(s+1+j)(s+1-j)} \end{bmatrix}$$

NOW

① $$\frac{s+2}{(s+1+j)(s+1-j)} = \frac{-(1+j)+2}{[-(1+j)+1-j]} \frac{1}{s+1+j} + \frac{-(1-j)+2}{[-(1-j)+1+j]} \frac{1}{s+j-1}$$

$$= \frac{1-j}{-j2} \frac{1}{s+1+j} + \frac{1+j}{j2} \frac{1}{s+j-1}$$

$$\frac{1+j}{j2} = \frac{1}{2}\sqrt{2}\, e^{j\frac{\pi}{4}} e^{-j\frac{\pi}{2}} = \frac{1}{\sqrt{2}} e^{-j\frac{\pi}{4}}$$

$$\Rightarrow \frac{s+2}{(s+1+j)(s+1-j)} = \frac{1}{\sqrt{2}} e^{+j\frac{\pi}{4}} \frac{1}{s+1+j} + \frac{1}{\sqrt{2}} e^{-j\frac{\pi}{4}} \frac{1}{s+j-1}$$

NOW

$$\mathcal{L}\left[2Me^{-at}\cos(\omega t-\theta)\right] = \frac{Me^{j\theta}}{s+\alpha+j\omega} + \frac{Me^{-j\theta}}{s+\alpha-j\omega}$$

THUS

$$\mathcal{L}^{-1}\left[\frac{s+2}{(s^2+2s+2)}\right] = \sqrt{2}\,e^{-t}\cos\left(t-\frac{\pi}{4}\right)$$

② $$\frac{1}{(s+1+j)(s+1-j)} = \frac{1}{[-(1-j)+1+j](s+1-j)} + \frac{1}{[-(1+j)+1-j](s+1+j)}$$

$$= \frac{1}{j2}\,\frac{1}{s+1-j} + \frac{1}{-j2}\,\frac{1}{s+1+j}$$

$$= \frac{1}{2}e^{j\pi/2}\frac{1}{s+1+j} + \frac{1}{2}e^{-j\pi/2}\frac{1}{s+1-j}$$

$$\mathcal{L}^{-1}\left[\frac{1}{(s+1+j)(s+1-j)}\right] = e^{-t}\cos(t-\pi/2)$$

$$= e^{-t}\sin t$$

③ $$\frac{-2}{(s+1+j)(s+1-j)} = -2\left[\frac{1}{(s+1+j)(s+1-j)}\right]$$

$$\Rightarrow \mathcal{L}^{-1}\left[\frac{-2}{(s+1+j)(s+1-j)}\right] = -2e^{-t}\sin t$$

④ $$\frac{s}{(s+1+j)(s+1-j)} = \frac{-(1+j)}{[-(1+j)+1-j](s+1+j)} + \frac{-(1-j)}{[-(1-j)+1+j](s+1-j)}$$

$$= \frac{1+j}{j2}\,\frac{1}{s+1+j} + \frac{1-j}{-j2}\,\frac{1}{s+1-j}$$

$$= \frac{1}{\sqrt{2}}e^{-j\pi/4}\frac{1}{s+1+j} + \frac{1}{\sqrt{2}}e^{j\frac{\pi}{4}}\frac{1}{s+1-j}$$

$$\therefore \mathcal{L}^{-1}\left[\frac{s}{s^2+2s+2}\right] = \sqrt{2}\,e^{-t}\cos t+\frac{\pi}{4}$$

THUS

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

$$= \begin{bmatrix} \sqrt{2}\,e^{-t}\cos\left(t-\frac{\pi}{4}\right) & e^{-t}\sin t \\ -2e^{-t}\sin t & \sqrt{2}\,e^{-t}\cos\left(t+\frac{\pi}{4}\right) \end{bmatrix}$$

(d) $Y(0) = 2/10$ $\qquad \dot{Y}(0) = 0$

$f(t) = 2e^{-2t}\mu(t)$

FIND $Y(t)$ AND $\dot{Y}(t)$ $\forall$ $t \geq 0$

FOR $t_0 = 0$

$$x(t) = \Phi(t)x_0 + \int_0^t \Phi(t-\tau)B \quad u(\tau)\,d\tau$$

$$= \begin{bmatrix} \sqrt{2}\,e^{-t}\cos\left(t-\frac{\pi}{4}\right) & e^{-t}\sin t \\ -2e^{-t}\sin t & \sqrt{2}\,e^{-t}\cos\left(t+\frac{\pi}{4}\right) \end{bmatrix} \begin{bmatrix} 2/10 \\ 0 \end{bmatrix}$$

$$+\int_0^t \left[\begin{array}{c|c} \sqrt{2}\,e^{-(t-\tau)}\cos\left(t-\tau-\frac{\pi}{4}\right) & e^{-(t-\tau)}\sin(t-\tau) \\ -2e^{-(t-\tau)}\sin(t-\tau) & \sqrt{2}\,e^{-(t-\tau)}\cos\left(t-\tau+\frac{\pi}{4}\right) \end{array}\right] \begin{bmatrix} 0 \\ \frac{1}{M} \end{bmatrix} 2e^{-2\tau}\,d\tau$$

$$= \begin{bmatrix} \frac{2\sqrt{2}}{10}e^{-t}\cos\left(t-\frac{\pi}{4}\right) \\ -\frac{4}{10}e^{-t}\sin t \end{bmatrix}$$

$$+ \int_0^t \begin{bmatrix} \frac{2}{M}e^{-t}e^{-\tau}\sin(t-\tau) \\ \frac{2\sqrt{2}}{M}e^{-t}e^{-\tau}\cos\left(t-\tau+\frac{\pi}{4}\right) \end{bmatrix} d\tau$$

WE MUST NOW EVALUATE THESE INTEGRALS $\Longrightarrow$

NOW $\int e^{ax} \sin(x-\xi)\,dx = \frac{e^{ax}}{a^2+1}\left[a\sin(x-\xi) - \cos(x-\xi)\right]$

THUS

$$\int_0^t e^{-\tau}\sin(t-\tau)\,d\tau = \int_t^0 e^{-\tau}\sin(\tau-t)\,d\tau$$

$$= \frac{-e^{-\tau}}{2}\left[+\sin(\tau-t) + \cos(\tau-t)\right]_t^0$$

$$= -\frac{1}{2}\left[\, e^0\{\sin(-t) + \cos(-t)\} - e^{-t}\{\sin 0 + \cos 0\}\right]$$

$$= -\frac{1}{2}\left[-\sin t + \cos t - e^{-t}\right]$$

$$= \frac{1}{2}e^{-t} + \frac{1}{2}\sin t - \frac{1}{2}\cos t$$

THUS

$$X_1(t) = \frac{2\sqrt{2}}{10}e^{-t}\cos\left(t-\frac{\pi}{4}\right) + \frac{2}{M}e^{-t}\left[\frac{1}{2}e^{-t} + \frac{1}{2}\sin t - \frac{1}{2}\cos t\right]$$

$$= \frac{2\sqrt{2}}{10}e^{-t}\cos\left(t-\frac{\pi}{4}\right) + \frac{1}{M}e^{-2t} + \frac{1}{M}e^{-t}\sin t - \frac{1}{M}e^{-t}\cos t$$

BUT $M=1$

$$\Rightarrow x_1(t) = \frac{2\sqrt{2}}{10}e^{-t}\cos\left(t-\frac{\pi}{4}\right) + e^{-2t} + e^{-t}\left[\sin t - \cos t\right]$$

NOW

$$\sin t - \cos t = \sqrt{2}\left[\sin t \sin\left(+\frac{3\pi}{4}\right) + \cos t \cos\frac{3\pi}{4}\right]$$

$$= \sqrt{2}\cos\left(t - \frac{3\pi}{4}\right)$$

THUS

$$X_1(t) = \frac{2\sqrt{2}}{10}e^{-t}\cos\left(t-\frac{\pi}{4}\right) + e^{-2t} + \sqrt{2}\cos\left(t-\frac{3\pi}{4}\right)e^{-t}$$

WE ALSO MUST EVALUATE

$$\int_0^t e^{-\tau} \cos(t - \tau + \tfrac{\pi}{4}) d\tau$$

$$= \int_0^t e^{-\tau} \cos(\tau - t - \tfrac{\pi}{4}) d\tau$$

SINCE $\cos(\xi - \frac{\pi}{2}) = \sin \xi$; WE EQUIVAL. HAVE

$$\int_0^t e^{-\tau} \sin(\tau - t + \tfrac{\pi}{4}) d\tau$$

$$= \tfrac{-1}{2} e^{-t} \left[+\sin(\tau - t + \tfrac{\pi}{4}) + \cos(\tau - t + \tfrac{\pi}{4}) \right]_0^t$$

$$= -\tfrac{1}{2} \Big[e^{-t} \{ \sin \tfrac{\pi}{4} + \cos \tfrac{\pi}{4} \}$$

$$- e^{0} \{ \sin(-t + \tfrac{\pi}{4}) + \cos(-t + \tfrac{\pi}{4}) \} \Big]$$

$$= -\tfrac{1}{2} e^{-t} \tfrac{2}{\sqrt{2}} + \tfrac{1}{2} \{ -\sin(t - \tfrac{\pi}{4}) + \cos(t - \tfrac{\pi}{4}) \}$$

$$= -\tfrac{1}{\sqrt{2}} e^{-t} + \tfrac{1}{2} \{ \cos(t - \tfrac{\pi}{4}) - \sin(t - \tfrac{\pi}{4}) \}$$

BUT

$$\cos(t - \tfrac{\pi}{4}) - \sin(t - \tfrac{\pi}{4})$$

$$= \sqrt{2} \Big[\cos \tfrac{-\pi}{4} \cos(t - \tfrac{\pi}{4})$$

$$+ \sin \tfrac{-\pi}{4} \sin(t - \tfrac{\pi}{4}) \Big]$$

$$= \sqrt{2} \cos\left[\tfrac{-\pi}{4} - (t - \tfrac{\pi}{4}) \right]$$

$$= \sqrt{2} \cos \bar{t}$$

THUS

$$X_2(t) = \tfrac{-4}{10} e^{-t} \sin t + 2\sqrt{2} e^{-t} \Big[-\tfrac{1}{\sqrt{2}} e^{-t}$$

$$+ \tfrac{1}{\sqrt{2}} \cos t \Big]$$

$$= \tfrac{-4}{10} e^{-t} \sin t + 2e^{-2t} + 2e^{-t} \cos t$$

IN SUMMARY:

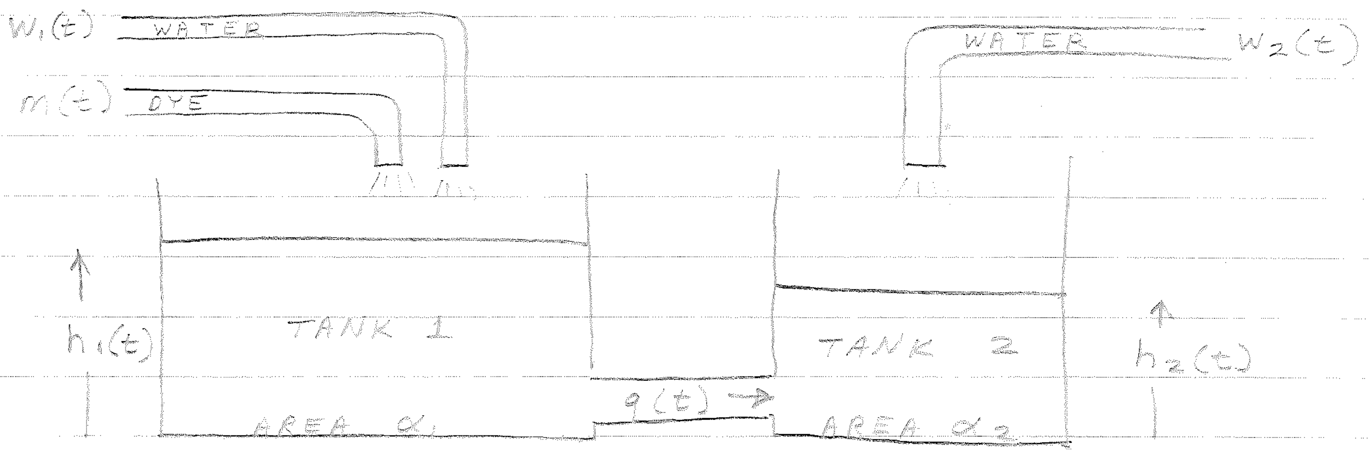
$$Y(t) = X_1(t) = \tfrac{2\sqrt{2}}{10} e^{-t} \cos(t - \tfrac{\pi}{4}) + e^{-2t}$$

$$+ \sqrt{2} \cos(t - \tfrac{3\pi}{4}) e^{-t}$$

$$\dot{Y}(t) = X_2(t) = \tfrac{-4}{10} e^{-t} \sin t - 2e^{-2t}$$

$$+ 2e^{-t} \cos t$$

pg. 25 16.5
(1-6)



ASSUME $q(t) = k(h_1 - h_2)$ ①

DYE DENSITY IN BOTH TANKS IS HOMOGENEOUS.

THE VOLUME IN TANK 1 IS

$$\text{VOL } 1 = \alpha_1 h_1 \checkmark$$

THUS

$$\frac{d}{dt} \text{VOL } 1 = \alpha_1 \frac{dh_1}{dt} = w_1(t) + m(t) - q(t)$$

USING ①:

$$\alpha_1 \frac{dh_1}{dt} = kh_2 - kh_1 + w_1(t) + m(t) \checkmark \quad ②$$

SIMILARLY, FOR TANK 2:

$$\text{VOL } 2 = \alpha_2 h_2$$

$$\frac{d}{dt} \text{VOL } 2 = \alpha_2 \frac{dh_2}{dt} = +q(t) + w_2(t)$$

$$= kh_1 - kh_2 + w_2 \quad ③ \checkmark$$

ASSUMING THE DENSITY IN BOTH TANKS IS THE SAME $\forall$ t:

$$\rho(t) = \text{DENSITY} = \frac{V_1(t)}{\text{VOL } 1} = \frac{V_2(t)}{\text{VOL } 2} \quad ④$$

$$= \frac{V_1}{\alpha_1 h_1} = \frac{V_2}{\alpha_2 h_2}$$

OR $h_2/h_1 = \alpha_1 V_2/\alpha_2 V_1$ ⑤

SIMILARLY

$$\frac{dV_1}{dt} = m(t) - \rho(t)\, q(t)$$

$$= m + k\rho(h_2 - h_1)$$

$$= m + k\frac{V_1}{\alpha_1 h_1}(h_2 - h_1)$$

$$= m + \frac{KV_1 h_2}{\alpha_1 h_1} - \frac{KV_1}{\alpha_1}$$

SUBSTITUTING ⑤ GIVES

$$dV_1/dt = m + \frac{KV_1}{\alpha_1}\frac{\alpha_1 V_2}{\alpha_2 V_1} - \frac{KV_1}{\alpha_1}$$

$$= m + \frac{KV_2}{\alpha_2} - \frac{KV_1}{\alpha_1} \quad ⑥$$

SINCE

$$\frac{dV_2}{dt} = q(t)\rho(t)$$

WE MAY WRITE:

$$dV_2/dt = \frac{KV_1}{\alpha_1} - \frac{KV_2}{\alpha_2} \quad ⑦$$

REWRITING EQS. ②, ③, ⑥ AND ⑦:

$$\dot{h}_1(t) = -\frac{K}{\alpha_1} h_1(t) + \frac{K}{\alpha_1} h_2(t) + \frac{1}{\alpha_1} W_1(t) + \frac{1}{\alpha_1} m(t)$$

$$\dot{h}_2(t) = \frac{K}{\alpha_2} h_1(t) - \frac{K}{\alpha_2} h_2(t) + \frac{1}{\alpha_2} W_2(t)$$

$$\dot{V}_1(t) = -\frac{K}{\alpha_1} V_1(t) + \frac{K}{\alpha_2} V_2(t) + m(t)$$

$$\dot{V}_2(t) = \frac{K}{\alpha_1} V_1(t) - \frac{K}{\alpha_2} V_2(t)$$

OR, IN MATRIX FORM

$$\begin{bmatrix} \dot{h}_1(t) \\ \dot{h}_2(t) \\ \dot{V}_1(t) \\ \dot{V}_2(t) \end{bmatrix} = \begin{bmatrix} -\frac{K}{\alpha_1} & \frac{K}{\alpha_1} & 0 & 0 \\ \frac{K}{\alpha_2} & -\frac{K}{\alpha_2} & 0 & 0 \\ 0 & 0 & -\frac{K}{\alpha_1} & \frac{K}{\alpha_2} \\ 0 & 0 & \frac{K}{\alpha_1} & -\frac{K}{\alpha_2} \end{bmatrix} \begin{bmatrix} h_1(t) \\ h_2(t) \\ V_1(t) \\ V_2(t) \end{bmatrix} + \begin{bmatrix} \frac{1}{\alpha_1} & 0 & \frac{1}{\alpha_1} \\ 0 & \frac{1}{\alpha_2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} W_1(t \\ W_2(t \\ m(t \end{bmatrix}$$

NOTE

THE ASSUMPTION THAT THE (UNIFORM) DYE DENSITY IN BOTH TANKS IS THE SAME SEEMS SOMEWHAT UNREALISTIC. AS SUCH, LET'S REWORK THE PROBLEM ASSUMING A (UNIFORM) DYE DENSITY $\rho_1(t)$ IN TANK 1 AND $\rho_2(t)$ IN TANK 2. OBVIOUSLY:

$$\rho_1(t) = \frac{V_1(t)}{VOL\ 1} = \frac{V_1(t)}{\alpha_1 h_1(t)} \quad (8a)$$

AND

$$\rho_2(t) = \frac{V_2(t)}{\alpha_2 h_2(t)} \quad (8b)$$

THE VOLUMES ARE STILL

$$VOL\ 1 = \alpha_1 h_1(t)$$
$$VOL\ 2 = \alpha_2 h_2(t)$$

AND

$$\frac{d}{dt} VOL\ 1 = \alpha_1 \frac{dh_1(t)}{dt} = W_1(t) + m(t) - q(t)$$
$$= W_1(t) + m(t) - k\left[h_1(t) - h_2(t)\right]$$

OR

$$\frac{dh_1}{dt} = \frac{1}{\alpha_1} W_1(t) + \frac{1}{\alpha_1} m(t) + \frac{K}{\alpha_1} h_2(t) - \frac{K}{\alpha_2} h_1(t) \quad ⑨$$

SIMILARLY

$$\frac{dh_2}{dt} = \frac{K}{\alpha_2} h_1(t) - \frac{K}{\alpha_2} h_2(t) + \frac{1}{\alpha_2} W_2(t) \quad ⑩$$

THESE ARE THE SAME AS FOR THE PREVIOUS CASE (Eqs ② + ③)

THE RATE OF CHANGE OF DYE IN TANK #1 DEPENDS ON WHETHER IT IS SUPPLYING DYE TO TANK #2 ($q(t) > 0$) OR IS GETTING DYE FROM TANK #2 ($q(t) < 0$). THUS

$$\frac{dv_1}{dt} = m(t) - \rho_1(t)\,q(t)\,\mu[q(t)] - \rho_2(t)\,q(t)\,\mu[-q(t)]$$

$$= m(t) - q(t)\left[\rho_1(t)\,\mu\{q(t)\} + \rho_2(t)\,\mu\{-q(t)\}\right]$$

WHERE $\mu(t)$ IS THE UNIT STEP:

$$\mu(x) = \begin{cases} 1 & ;\ x \geq 0 \\ 0 & ;\ x < 0 \end{cases}$$

SINCE $\mu[q(t)] = \mu[h_1(t) - h_2(t)]$:

$$\frac{dv_1}{dt} = m(t) - k[h_1(t) - h_2(t)]\left[\rho_1(t)\,\mu\{h_1(t) - h_2(t)\} + \rho_2(t)\,\mu\{h_2(t) - h_1(t)\}\right]$$

SUBSTITUTING ⑧a & ⑧b AND DROPPING TIME ARGUMENTS:

$$\frac{dv_1}{dt} = m + k[h_2 - h_1]\left[\frac{v_1}{\alpha_1 h_1}\mu\{h_1 - h_2\} + \frac{v_2}{\alpha_2 h_2}\mu\{h_2 - h_1\}\right] \qquad ⑪$$

WE CAN ALSO WRITE

$$\frac{dV_2}{dt} = \rho_1(t)q(t)\mu(q(t)) + \rho_2(t)q(t)\mu(-q(t))$$

$$= q(t)\left[\rho_1(t)\mu(q(t)) + \rho_2(t)\mu(-q(t))\right]$$

$$= k\left[h_1 - h_2\right]\left[\frac{v_1(t)}{\alpha_1 h_1(t)}\mu\{h_1 - h_2\}\right.$$

$$\left. + \frac{v_2}{\alpha_2 h_2}\mu\{h_2 - h_1\}\right] \quad \text{⑫}$$

THE STATE EQUATIONS FOR THIS TIME-INVARIANT NON-LINEAR MODEL, FROM ⑨, ⑩, ⑪, AND ⑫, ARE THUS

$$\frac{dh_1(t)}{dt} = \frac{-K}{\alpha_1}h_1(t) + \frac{K}{\alpha_1}h_2(t) + \frac{1}{\alpha_1}w_1(t) + \frac{1}{\alpha_1}m(t)$$

$$\frac{dh_2(t)}{dt} = \frac{K}{\alpha_2}h_1(t) - \frac{K}{\alpha_2}h_2(t) + \frac{1}{\alpha_2}w_2(t)$$

$$\frac{dv_1(t)}{dt} = -k\{h_1(t) - h_2(t)\}\left[\frac{v_1(t)}{\alpha_1 h_1(t)}\mu\{h_1(t) - h_2(t)\}\right.$$

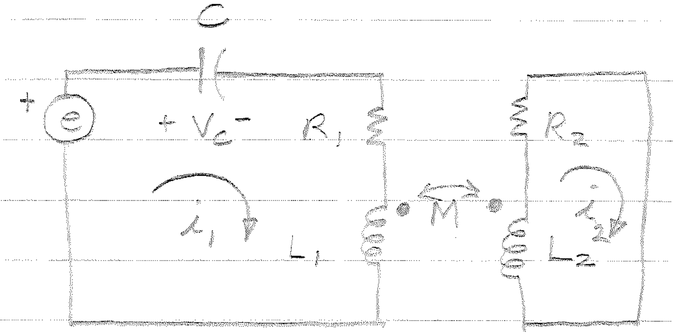
$$\left. + \frac{v_2(t)}{\alpha_2 h_2(t)}\mu\{h_2(t) - h_1(t)\}\right] + m(t)$$

$$\frac{dv_2(t)}{dt} = k\{h_1(t) - h_2(t)\}\left[\frac{v_1(t)}{\alpha_1 h_1(t)}\mu\{h_1(t) - h_2(t)\}\right.$$

$$\left. + \frac{v_2(t)}{\alpha_2 h_2}\mu\{h_2(t) - h_1(t)\}\right]$$

PG. 27
(1-9)

l(65



USE V_c, i_1, AND i_2 AS STATE VARIABLES

$$-e + V_c + R_1 i_1 + L_1 \frac{di_1}{dt} - M \frac{di_2}{dt} = 0 \quad ①$$

$$L_2 \frac{di_2}{dt} + R_2 i_2 - M \frac{di_1}{dt} = 0 \quad ②$$

$$\frac{dV_c}{dt} = \frac{1}{C} i_1 \quad ③$$

SOLVING FOR $\frac{di_1}{dt}$ IN ②:

$$M \frac{di_1}{dt} = L_2 \frac{di_2}{dt} + R_2 i_2$$

$$\frac{di_1}{dt} = L_2/M \frac{di_2}{dt} + R_2/M\, i_2$$

SUBSTITUTING INTO ①:

$$-e + V_c + R_1 i_1 + L_1 \left[\frac{L_2}{M}\frac{di_2}{dt} + \frac{R_2}{M} i_2\right] - M \frac{di_2}{dt} = 0$$

$$-e + V_c + R_1 i_1 + \frac{R_2 L_1}{M} i_2 + \left[\frac{L_1 L_2}{M} - M\right] \frac{di_2}{dt} = 0$$

$$-e + V_c + R_1 i_1 + \frac{R_2 L_1}{M} i_2 + \frac{1}{M}\left[L_2 L_1 - M^2\right] \frac{di_2}{dt} = 0$$

LET $K^2 = L_1 L_2 - M^2$ ④

$$\Rightarrow -e + V_c + R_1 i_1 + \frac{R_2 L_1}{M} i_2 + \frac{K^2}{M}\frac{di_2}{dt}$$

OR

$$\frac{K^2}{M}\frac{di_2}{dt} = -R_1 i_1 - \frac{R_2 L_1}{M} i_2 - V_c + e$$

OR

$$\frac{di_2}{dt} = \frac{-MR_1}{K^2} i_1 - \frac{\cancel{M}}{K^2}\frac{R_2 L_1}{\cancel{M}} i_2 - \frac{M}{K^2} V_c + \frac{M}{K^2} e$$

$$= \frac{-MR_1}{K^2} i_1 - \frac{R_2 L_1}{K^2} i_2 - \frac{M}{K^2} V_c + \frac{M}{K^2} e \quad ⑤$$

SOLVING FOR $\frac{di_2}{dt}$ IN ①:

$$M\frac{di_2}{dt} = -e + V_c + R_1 i_1 + L_1\frac{di_1}{dt}$$

$$\frac{di_2}{dt} = \frac{-R_1}{M} i_1 + \frac{1}{M} V_c - \frac{1}{M} e + \frac{L_1}{M}\frac{di_1}{dt}$$

SUBSTITUTING INTO ②:

$$L_2\left[\frac{R_1}{M} i_1 + \frac{1}{M} V_c - \frac{1}{M} e + \frac{L_1}{M}\frac{di_1}{dt}\right] + R_2 i_2 - M\frac{di_1}{dt} = 0$$

$$\frac{L_2R_1}{M} i_1 + R_2 i_2 + \frac{L_2}{M} V_c - \frac{L_2}{M} e + \left[\frac{L_1L_2}{M} - M\right]\frac{di_1}{dt} = 0$$

USING ④:

$$\frac{L_2R_1}{M} i_1 + R_2 i_2 + \frac{L_2}{M} V_c - \frac{L_2}{M} e + \frac{K^2}{M}\frac{di_1}{dt} = 0$$

OR

$$\frac{K^2}{M}\frac{di_1}{dt} = \frac{-L_2R_1}{M} i_1 - R_2 i_2 - \frac{L_2}{M} V_c + \frac{L_2}{M} e$$

$$\frac{di_1}{dt} = \frac{-\cancel{M}}{K^2}\frac{L_2R_1}{\cancel{M}} i_1 - \frac{M}{K^2} R_2 i_2 - \frac{\cancel{M}}{K^2}\frac{L_2}{\cancel{M}} V_c + \frac{\cancel{M}}{K^2}\frac{L_2}{\cancel{M}} e$$

$$= \frac{-L_2R_1}{K^2} i_1 - \frac{MR_2}{K^2} i_2 - \frac{L_2}{K^2} V_c + \frac{L_2}{K^2} e \quad ⑥$$

OUR STATE EQUATIONS ARE ③, ⑤, & ⑥:

$$\begin{bmatrix} dV_c/dt \\ di_2/dt \\ di_1/dt \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1/C \\ -M/K^2 & \frac{-R_2L_1}{K^2} & \frac{-MR_1}{K^2} \\ -L_2/K^2 & -MR_2/K^2 & \frac{-L_2R_1}{K^2} \end{bmatrix} \begin{bmatrix} V_c \\ i_2 \\ i_1 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{M}{K^2} \\ \frac{L_2}{K^2} \end{bmatrix} e(t)$$

$$\Rightarrow K^2 \triangleq L_1L_2 - M^2$$

pg. 27 15.5

(1-12)

(a) $\frac{Y(s)}{U(s)} = \frac{5}{s[s+1]} = \frac{5}{s^2+s}$

$= \frac{Y'}{W} \frac{W}{U}$

$\frac{W}{U} = \frac{5}{s^2+s}$; $\frac{Y}{W} = 1$

$s^2 W + sW = 5U$

$\ddot{W} + \dot{W} = 5U$

$\ddot{W} = -\dot{W} + 5U$; $Y = W$

$$\begin{bmatrix} \dot{W} \\ \ddot{W} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} W \\ \dot{W} \end{bmatrix} + \begin{bmatrix} 0 \\ 5 \end{bmatrix} U$$

$X_1 = W$, $\dot{W} = X_2$

$$\begin{bmatrix} \dot{X}_1 \\ \dot{X}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 5 \end{bmatrix} U$$

$Y = X_1$

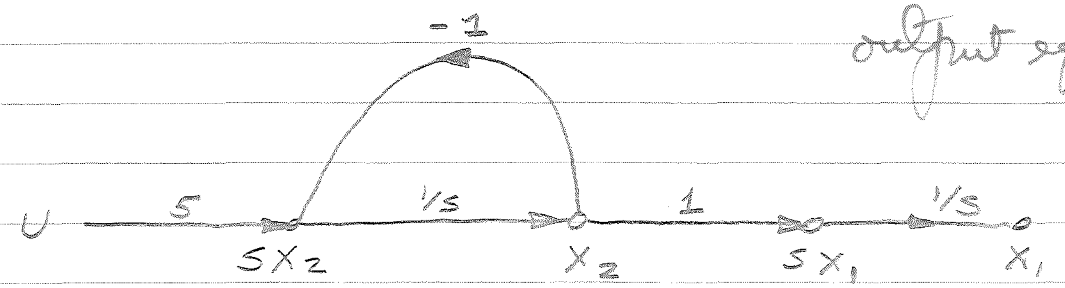
IN LAPLACE FORM

$SX_1 = X_2$

$SX_2 = -X_2 + 5U$

SIGNAL FLOW GRAPH: $y = X_1(t)$

output eqn.

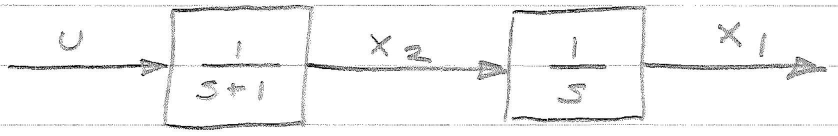


REARRANGING:

$$X_1 = \frac{1}{s} X_2$$

$$(s+1) X_2 = 5U \Rightarrow X_2 = \frac{1}{s+1} U$$

BLOCK DIAGRAM IS THEN:



$$(f)\quad \frac{Y(s)}{U(s)} = \frac{(s+1)(s+2)}{s^2} = \frac{Y(s)}{W(s)}\,\frac{W(s)}{U(s)}$$

$$\frac{Y(s)}{W(s)} = (s+1)(s+2) \qquad \frac{W(s)}{U(s)} = \frac{1}{s^2}$$

$$Y(s) = [s^2+3s+2]\,W(s) \qquad U(s) = s^2 W$$

$$Y = \ddot{W} + 3\dot{W} + 2W \qquad U = \ddot{W} \Rightarrow$$

$$\begin{bmatrix} \dot{W} \\ \ddot{W} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} W \\ \dot{W} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} U$$

NOTE THAT WE CAN WRITE

$$Y = 3\dot{W} + 2W + U$$

$$\Rightarrow Y = [2 \quad 3]\begin{bmatrix} W \\ \dot{W} \end{bmatrix} + U$$

LET $x_1 = W \quad x_2 = \dot{W}$

$$\Rightarrow \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} U$$

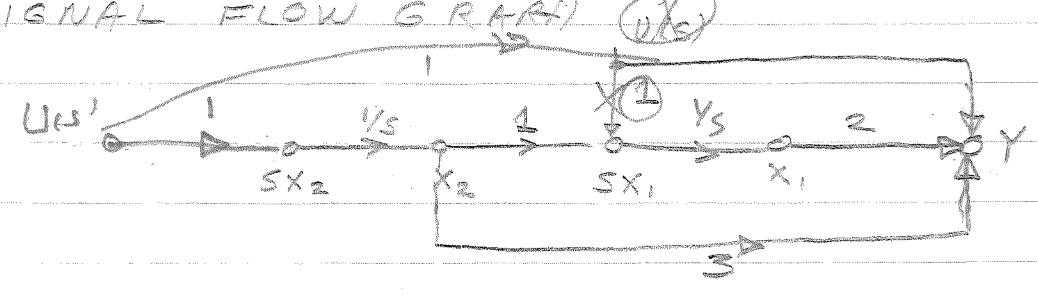
$$Y = [2 \quad 3]\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + U$$

LAPLACE TRANSFORMING

$$\begin{bmatrix} sX_1 \\ sX_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} U$$

$$Y = [2 \quad 3]\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + U$$

SIGNAL FLOW GRAPH

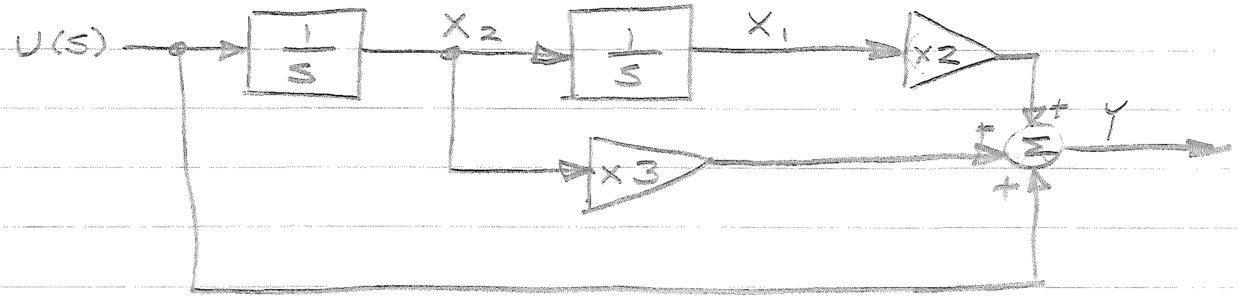


TO DRAW BLOCK DIAGRAM, WE WRITE

$$sX_1 = X_2 \Longrightarrow X_1 = \frac{1}{s} X_2$$

$$sX_2 = U \Longrightarrow X_2 = \frac{1}{s} U$$

$$Y = 2X_1 + 3X_2 + U$$



(i) $\frac{Y(s)}{U(s)} = \frac{s^2+7s+12}{s(s+1)(s+2)} = \frac{Y(s)}{W} \frac{W}{U}$

$\frac{Y(s)}{W} = s^2+7s+12$

$\Rightarrow Y = \ddot{w} + 7\dot{w} + 12W$

$\frac{W}{U} = \frac{1}{s(s^2+3s+2)}$

$\Rightarrow U = \dddot{w} + 3\ddot{w} + 2\dot{w}$

$\Rightarrow \dddot{w} = -3\ddot{w} - 2\dot{w} + U$

$$\begin{bmatrix} \dot{w} \\ \ddot{w} \\ \dddot{w} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & -3 \end{bmatrix} \begin{bmatrix} w \\ \dot{w} \\ \ddot{w} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} U$$

$$Y = \begin{bmatrix} 12 & 7 & 1 \end{bmatrix} \begin{bmatrix} w \\ \dot{w} \\ \ddot{w} \end{bmatrix}$$

OR, WITH $x_1 = w$ $x_2 = \dot{w}$ $x_3 = \ddot{w}$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} U$$

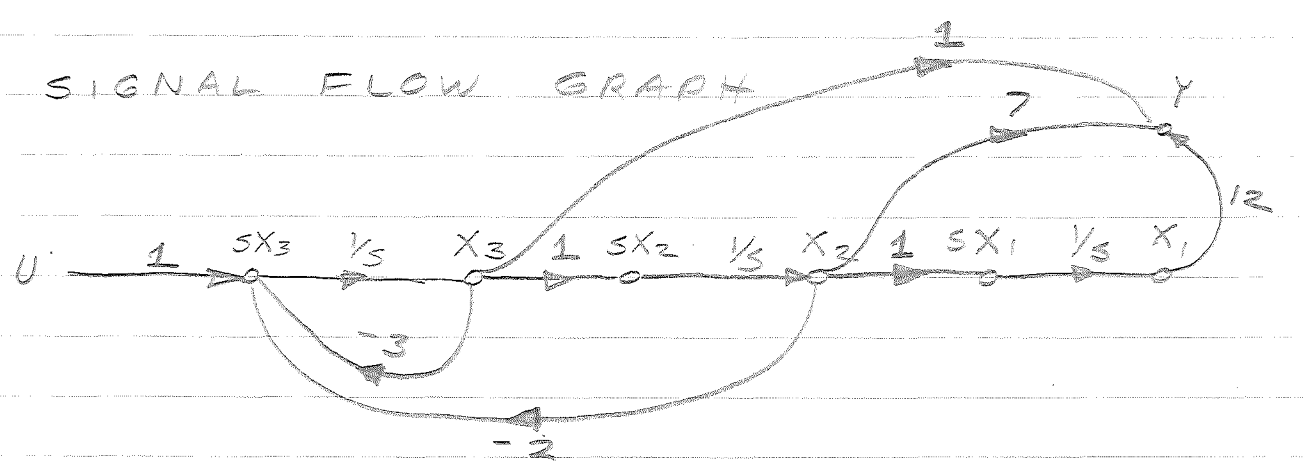
$$Y = \begin{bmatrix} 12 & 7 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

LAPLACE

XFORMING:

$$\begin{bmatrix} sX_1 \\ sX_2 \\ sX_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & -3 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} U$$

SIGNAL FLOW GRAPH

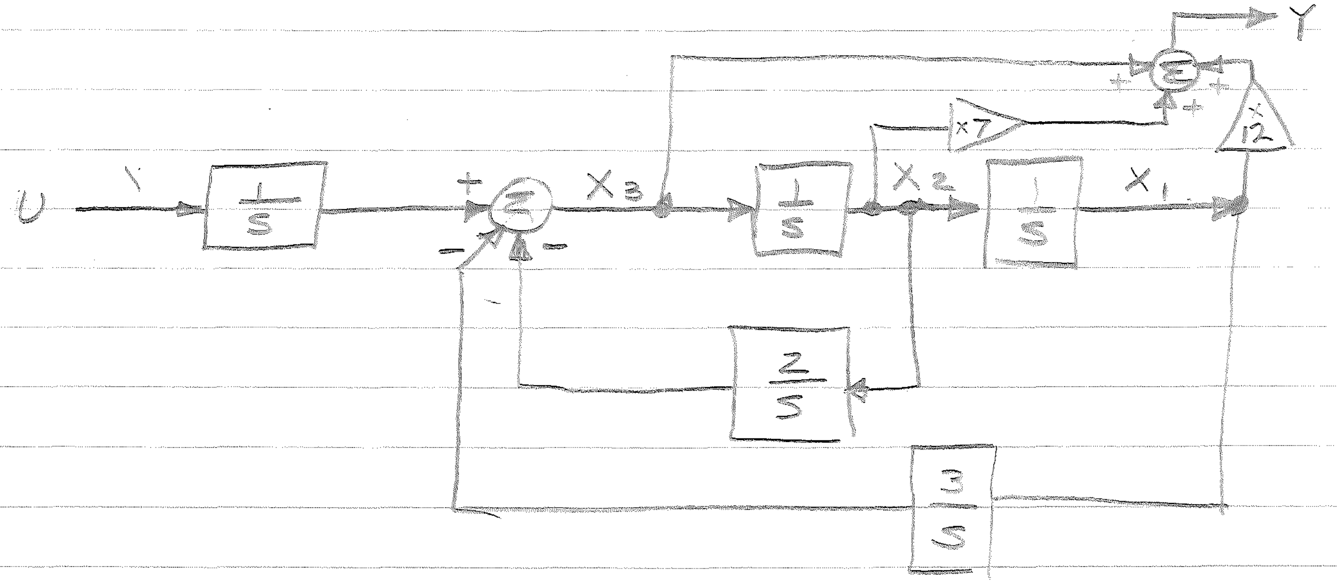


TO DRAW BLOCK DIAGRAM, CONSIDER

$sX_1 = X_2 \Longrightarrow X_1 = \frac{1}{s}X_2$

$sX_2 = X_3 \Longrightarrow X_2 = \frac{1}{s}X_3$

$sX_3 = -2X_2 - 3X_3 + U \Longrightarrow X_3 = \frac{-2}{s}X_2 - \frac{3}{s}X_3 + \frac{U}{s}$



Pg. 28
(1-14)

(a) $\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} U(t) \; ; \; Y(t) = X_1(t)$

5.5

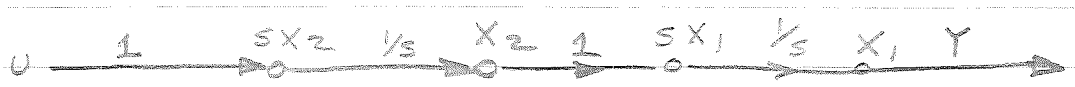
OR

$SX_1 = X_2 \Rightarrow X_1 = \frac{1}{S} X_2$

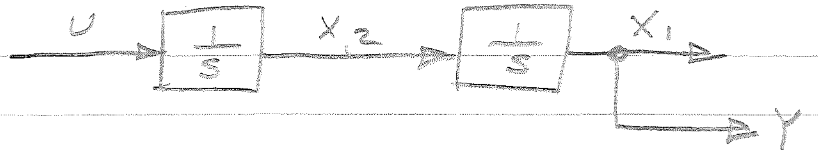
$SX_2 = U \Rightarrow X_2 = \frac{1}{S} U$

$Y = X_1 \Rightarrow Y = X_1$

SIGNAL FLOW GRAPH



BLOCK DIAGRAM



CONTROLLABILITY TEST

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$AB = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow \left[B \,\vdots\, AB \right] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{(NONSINGULAR)}$$

$\Rightarrow$ SYSTEM IS CONTROLLABLE

OBSERVABILITY TEST

$$Y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Rightarrow C^T = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$A^T = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \Rightarrow A^T C^T = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\left[C^T \,\vdots\, C^T A^T \right] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{(NONSINGULAR)}$$

$\Rightarrow$ SYSTEM IS OBSERVABLE

(c) FOR $M=0$, THE ANSWER TO PROBLEM 9 BECOMES

$$\begin{bmatrix} \dot{X}_1 \\ \dot{X}_2 \\ \dot{X}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1/C \\ 0 & -\frac{R_2}{L_2} & 0 \\ -1/L_1 & 0 & -\frac{R_1}{L_1} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L_1} \end{bmatrix} e(t)$$

$\ni \; X_1 = V_c \; , \; X_2 = i_2 \; , \; X_3 = i_1$

TEST FOR CONTROLLABILLITY

$$AB = \begin{bmatrix} 0 & 0 & 1/C \\ 0 & -R_2/L_2 & 0 \\ -1/L_1 & 0 & -R_1/L_1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1/L_1 \end{bmatrix} = \begin{bmatrix} 1/L_1C \\ 0 \\ -R_1/L_1^2 \end{bmatrix}$$

$$A^2 = \begin{bmatrix} 0 & 0 & 1/C \\ 0 & -R_2/L_2 & 0 \\ -1/L_1 & 0 & -\frac{R_1}{L_1} \end{bmatrix} \begin{bmatrix} 0 & 0 & 1/C \\ 0 & -R_2/L_2 & 0 \\ -1/L_1 & 0 & -\frac{R_1}{L_1} \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{1}{L_1C} & 0 & -\frac{R_1}{L_1C} \\ 0 & R_2^2/L_2^2 & 0 \\ R_1/L_1^2 & 0 & R_1^2/L_1^2 \end{bmatrix}$$

$$A^2B = \begin{bmatrix} -\frac{R_1}{L_1^2C} \\ 0 \\ R_1^2/L_1^3 \end{bmatrix}$$

$$\left[B \,\vdots\, AB \,\vdots\, A^2B\right] = \begin{bmatrix} 0 & 1/L_1C & -R_1/L_1^2C \\ 0 & 0 & 0 \\ 1/L_1 & -R_1/L_1^2 & R_1^2/L_1^3 \end{bmatrix}$$

$\Rightarrow$ SYSTEM IS NOT CONTROLLABLE SINCE THIS MATRIX IS SINGULAR (ie det( ) = 0), OR, EQUIVALENTLY, IT ONLY HAS RANK 2

GIVEN

$$Y(t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} X(t)$$

$$\Rightarrow C^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$A^T = \begin{bmatrix} 0 & 0 & -1/L_1 \\ 0 & -R_2/L_2 & 0 \\ \frac{1}{C} & 0 & -R_1/L_1 \end{bmatrix}$$

$$A^T C^T = \begin{bmatrix} 0 & 0 \\ 0 & -R_2/L_2 \\ 1/C & 0 \end{bmatrix}$$

$$(A^T)^2 = \begin{bmatrix} 0 & 0 & -1/L_1 \\ 0 & -R_2/L_2 & 0 \\ 1/C & 0 & -\frac{R_1}{L_1} \end{bmatrix} \begin{bmatrix} 0 & 0 & -\frac{1}{L_1} \\ 0 & -\frac{R_2}{L_2} & 0 \\ \frac{1}{C} & 0 & -\frac{R_1}{L_1} \end{bmatrix} = \begin{bmatrix} \frac{-1}{L_1C} & 0 & \frac{R_1}{L_1^2} \\ 0 & \frac{R_2^2}{L_2^2} & 0 \\ \frac{-R_1}{L_1C} & 0 & R_1^2/L_1^2 \end{bmatrix}$$

$$(A^T)^2 C^T = \begin{bmatrix} -\frac{1}{L_1C} & 0 \\ 0 & \frac{R_2^2}{L_2^2} \\ -R_1/L_1C & 0 \end{bmatrix}$$

THUS:

$$\left[C^T \vdots A^TC^T \vdots (A^T)^2C^T\right] = \begin{array}{c} \begin{matrix} A & B & C & D & E & F \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & -\frac{1}{L_1C} & 0 \\ 0 & 1 & 0 & \frac{-R_2}{L_2} & 0 & \frac{R_2^2}{L_2^2} \\ 0 & 0 & \frac{1}{C} & 0 & \frac{-R_1}{L_1C} & 0 \end{bmatrix} \end{array}$$

COLUMNS B D & F ARE LINEARLY DEPENDENT

$\Rightarrow$ RANK = 4 ≥ 3 Rank = 3

∴ SYSTEM IS OBSEVABLE

(e) $\dot{x}(t) = \begin{bmatrix} -2 & 0 & 1 \\ 0 & -1 & 0 \\ -3 & -4 & -2 \end{bmatrix} x(t) + \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix} U(t)$

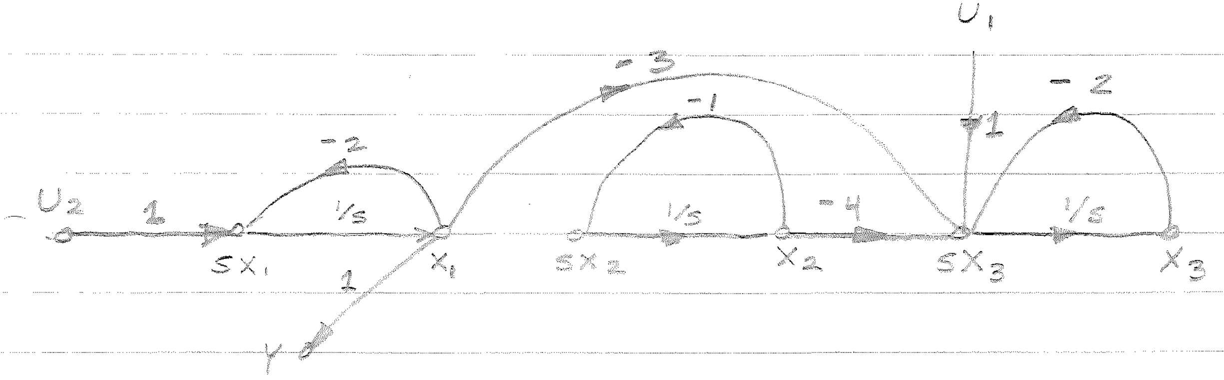
$Y(t) = x_1(t)$

$S X_1 = -2X_1 + X_3 + U_2$

$S X_2 = -X_2$

$S X_3 = -3X_1 - 4X_2 - 2X_3 + U_1$

SIGNAL FLOW GRAPH:



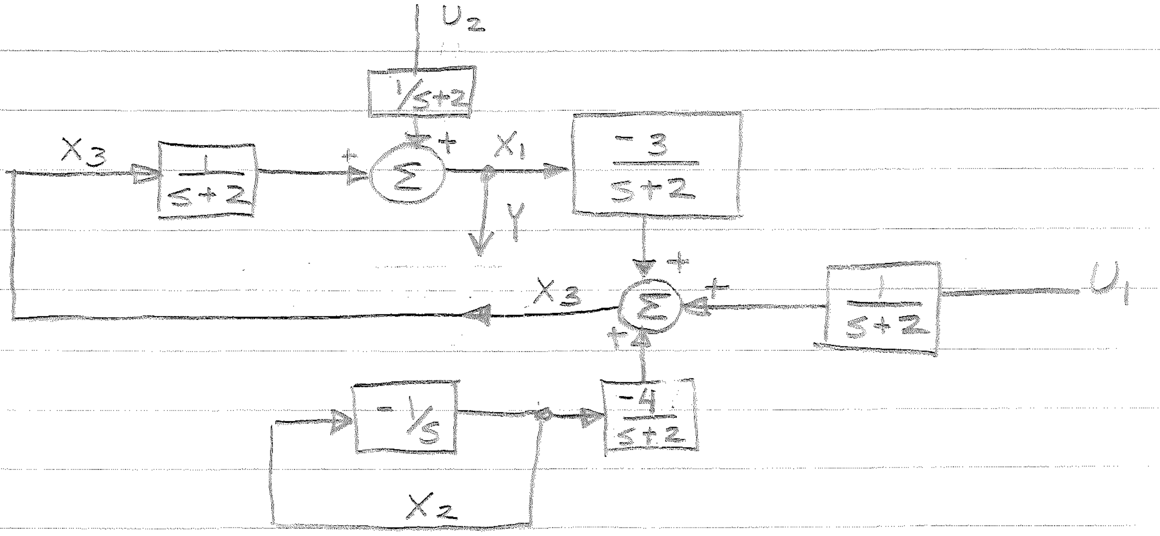
NOW

$X_1(s+2) = X_3 + U_2 \Rightarrow X_1 = \frac{1}{s+2} X_3 + \frac{1}{s+2} U_2$

$X_2 = -\frac{1}{s} X_2$

$X_3(s+2) = -3X_1 - 4X_2 + U_1 \Rightarrow X_3 = \frac{-3X_1}{s+2} - \frac{4X_2}{s+2} + \frac{U_1}{s+2}$

THE BLOCK DIAGRAM IS THUS



TESTING FOR CONTROLLABILITY:

$$B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}$$

$$A = \begin{bmatrix} -2 & 0 & 1 \\ 0 & -1 & 0 \\ -3 & -4 & -2 \end{bmatrix}$$

$$AB = \begin{bmatrix} 1 & -2 \\ 0 & 0 \\ -2 & -3 \end{bmatrix}$$

$$A^2 = \begin{bmatrix} -2 & 0 & 1 \\ 0 & -1 & 0 \\ -3 & -4 & -2 \end{bmatrix} \begin{bmatrix} -2 & 0 & 1 \\ 0 & -1 & 0 \\ -3 & -4 & -2 \end{bmatrix} = \begin{bmatrix} 1 & -4 & -4 \\ 0 & 1 & 0 \\ 12 & 12 & 1 \end{bmatrix}$$

$$\Rightarrow A^2B = \begin{bmatrix} -4 & 1 \\ 0 & 0 \\ 1 & 12 \end{bmatrix}$$

$$\left[B \,\vdots\, AB \,\vdots\, A^2B\right] = \left[\begin{array}{cc|cc|cc} 0 & 1 & 1 & -2 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -2 & -3 & 1 & 12 \end{array}\right]$$

ENTIRE SECOND ROW IS ZERO!

$\Rightarrow$ <u>NOT</u> CONTROLLABLE

The Rank = 2

TESTING FOR OBSERVABILITY

$$Y(t) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} x(t)$$

$$C^T = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$A^T C^T = \begin{bmatrix} -2 & 0 & -3 \\ 0 & -1 & -4 \\ 1 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

$$(A^T)^2 C^T = \begin{bmatrix} 1 & 0 & 12 \\ -4 & 1 & 12 \\ -4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -4 \\ -4 \end{bmatrix}$$

$$\left[C^T \,\vdots\, A^T C^T \,\vdots\, A^{T^{2}} C^T \right] = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & -4 \\ 0 & 1 & -4 \end{bmatrix}$$

RANK = 3

$\Longrightarrow$ SYSTEM <u>IS</u> OBSERVABLE

BOB MARKS
DUE 9/27/76

pg. 47.

2-1. (a) A PERFORMANCE MEASURE TO

20 KEEP $v_2(t)$ AS CLOSE TO M AS POSSIBLE OVER A 24 HOUR PERIOD IS

$$J = \int_{t_0}^{t_0 + 1\,\text{DAY}} [v_2(t) - M]^2 \, dt$$

WE WOULD OF COURSE WANT TO MINIMIZE J.

(b) THE INPUT FLOWS MUST BE UPPER BOUNDED AND CANNOT BE NEGATIVE. THUS

$$0 \le m(t) \le M_{MAX}$$
$$0 \le w_1(t) \le W_{1MAX}$$
$$0 \le w_2(t) \le W_{2MAX}$$

THE TWO STATES, h_1 AND h_2, CANNOT BE NEGATIVE. FURTHERMORE, SINCE TANKS 1 & 2 HAVE FINITE HIEGHTS:

$$0 \le h_1(t) \le H_{1MAX}$$
$$0 \le h_2(t) \le H_{2MAX}$$

SIMILARLY, THE AMOUNT OF DIE IN THE TANKS MUST BE $\ni$

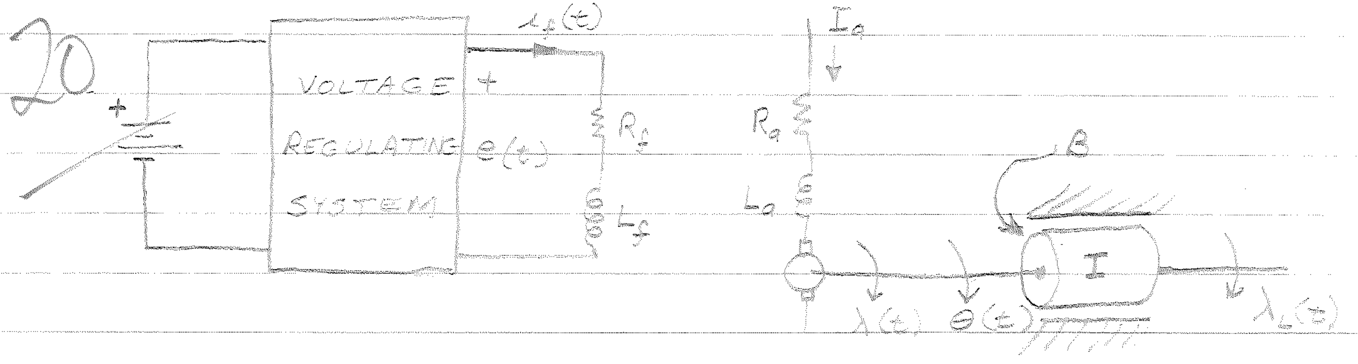
$$0 \le v_1(t) \le V_{1MAX}$$
$$0 \le v_2(t) \le V_{2MAX}$$

A POSSIBLE ASSIGNMENT IS

$$V_{1MAX} = \alpha_1 H_{1MAX}$$
$$V_{2MAX} = \alpha_2 H_{2MAX}$$

PG. 47
(2-3)

20



(a) FROM THE CIRCUIT

$$e(t) = R_f i_f(t) + L_f \frac{d i_f(t)}{dt}$$

$$\Rightarrow \quad di_f/dt = \frac{1}{L_f} e(t) - \frac{R_f}{L_f} i_f(t) \quad ①$$

THE MECHANICAL TORQUES ARE

$\lambda(t) = K_t i_f(t)$; DEVELOPED TORQUE

$B\omega = B\frac{d\theta}{dt}$; DUE TO VISCOUS FRICTION

λ_L ; LOAD TORQUE (OPPOSES DEVELOPED TORQ)

$I\alpha = I\frac{d\omega}{dt} = I\frac{d^2\theta}{dt^2}$; INERTIAL TORQUE (OPPOSES DEV. TORQUE)

SUMMING TORQUES:

$$\lambda(t) + B\omega - \lambda_L(t) - I\alpha = 0$$

OR $K_t i_f(t) + B\omega - \lambda_L - I\frac{d\omega}{dt} = 0$

OR $\frac{d\omega}{dt} = \frac{K_t}{I} i_f(t) \mp B/I\, \omega(t) - \lambda_L(t)/I$ ②

USING STATE VARIABLES $\omega = \frac{d\theta}{dt}$ AND $i_f(t)$
WITH INPUTS $\lambda_L(t)$ AND $e(t)$ GIVES,
FROM ① AND ②:

$$\begin{bmatrix} \dot{i}_f \\ \dot{\omega} \end{bmatrix} = \begin{bmatrix} 0 & -R_f/L_f \\ K_t/I & B/I \end{bmatrix} \begin{bmatrix} i_f \\ \omega \end{bmatrix} + \begin{bmatrix} 1/L_f & 0 \\ 0 & -1/I \end{bmatrix} \begin{bmatrix} e \\ \lambda_L \end{bmatrix}$$

(b) INPUT CONTROL CONSTRAINTS

$e(t)$ MUST OBVIOUSLY BE BOUNDED. AS SUCH, LET THE CONSTRAINT BE

$$|e(t)| \leq E_{MAX}$$

LIKEWISE, $\lambda_L(t)$ WILL HAVE BOUNDS. LET

$$|\lambda_L(t)| \leq \Lambda_{MAX}$$

STATE VARIABLE CONSTRAINTS

$i_f(t)$ WILL BE BOUNDED:

$$|i_f(t)| \leq I_{MAX}$$

SIMILARLY, THE ANGULAR VELOCITY

$$|\omega(t)| \leq \Omega_{MAX}$$

(c) i. $L_f = 0$

(THIS CONDITION, BY THE WAY, INVALIDATES THE STATE EQUATIONS)

FIRST OFF, ASSUME THAT THE ANGULAR VELOCITY IS TRANSFORMED LINEARLY TO LINEAR VELOCITY, $v(t)$. IF, FOR EXAMPLE, k IS THE RADIUS OF A WHEEL, THEN

$$v(t) = k\,\omega(t)$$

WE WISH TO KEEP THE SPEED AS CLOSE AS POSSIBLE TO 5 mph. THE PERFORMANCE MEASURE FOR THIS WOULD BE

$$J_1 = \int_0^{t_f} [k\,\omega(t) - 5]^2 dt \quad ③$$

WHERE t_f IS THE MISSION TIME, K HAS UNITS OF MILES, AND $\omega(t)$ HAS UNITS OF REVOLUTIONS PER HOUR. FOR $L_f = 0$, THE TOTAL ENERGY EXPENDED BY THE CONTROL DURING THE MISSION IS

$$E = \int_0^{t_f} i_f(t) \times e(t)\,dt = \frac{1}{R_f}\int_0^{t_f} e^2(t)\,dt \quad ④$$

THUS A SECOND PERFORMANCE MEASURE IS

$$J_2 = \frac{1}{R_f}\int_0^{t_f} e^2(t)\,dt \quad ⑤$$

THE COMPOSITE PERFORMANCE MEASURE IS THE SUM OF ③ AND ⑤

$$J = \int_0^{t_f}\left[\mu e^2(t) + \{k\omega(t) - 5\}^2\right]dt$$

WHERE μ IS A WEIGHTING FACTOR INTRODUCED TO ALLOW TRADEOFF BETWEEN THE TWO PERFORMANCE REQUIREMENTS. ($1/R_f$ IS ABSORBED INTO μ).

(ii) FOR $L_f \neq 0$, Eq. ③ IS STILL VALID. THE TOTAL EXPENDED ENERGY, HOWEVER, MUST NOW BE WRITTEN IN THE MORE GENERAL FORM

$$J_2 = E = \int_0^{t_f} e(t)\, i(t)\, dt$$

THE TOTAL PERFORMANCE MEASURE THUS BECOMES THE SUM OF THIS AND ③:

$$J = \int_0^{t_f} \left[\mu e(t)\, i(t) + \{ k\omega(t) - 5 \}^2 \right] dt$$

AGAIN, μ IS A WEIGHTING FACTOR.

pg. 48

(2-4)(a) THE TORQUE PRODUCED BY THE GAS JETS MUST OBVIOUSLY BY BOUNDED, SO THE INPUT CONSTRAINT WOULD BE

$$|u(t)| \leq \overline{\underline{U}}_{MAX}$$

THE STATE $X_1(t) \triangleq \Theta(t)$ IS SPECIFIED AT $t = t_f$. THUS

$$14.9^\circ \leq \Theta(t_f = 30) \leq 15.1^\circ$$

SINCE THERE IS NO ANGULAR VELOCITY REQUIREMENT, WE WILL NOT PLACE ANY CONSTRAINTS ON $X_2(t) = {}^{d\Theta}/_{dt}$

(b) IN THAT WE REQUIRE 30 SEC TO ATTAIN THE STATE CONSTRAINT UNDER THE INPUT CONSTRAINT, AND SINCE THE CONTROL ENERGY IS HERE PROPORTIONAL TO $U^2(t)$, LET THE PERFORMANCE MEASURE BE

$$J = \int_0^{30\,SEC} U^2(t)\,dt$$

$$J = \int_0^{30} |u|\,dt.$$

rate of fuel expenditure $\propto |u(t)|$.

pg 48

(2-5) (a) THE INPUT AND STATE CONSTRAINTS HERE ARE THE SAME AS IN THE PREVIOUS PROBLEM:

20

$$|u(t)| \leq U_{MAX}$$

$$14.9° \leq \theta(t_f) \leq 15.1°$$

(b) THE BEST PERFORMANCE MEASURE FOR MINIMUM TIME WOULD BE

$$J = t_f - t_0$$

OR, SETTING $t_0 = 0$:

$$J = t_f$$

Pg. 48

(2-6) (a) INPUT CONSTRAINTS

20

- THE THRUST MUST BE POSITIVE AND UPPER BOUNDED

$$\Rightarrow 0 \leq u_1 \leq T_{max}$$

- THE THRUST ANGLE MUST BE $\Rightarrow$

$$0 \leq u_2 \leq 2\pi$$

(ACTUALLY, THIS DOESN'T MATTER SINCE u_2 ALWAYS APPEARS AS THE ARGUMENT OF A TRIG. FUNCTION)

- STATE CONSTRAINTS

- IF WE ASSUME X=0 IS THE GROUND:

$$x_1 \geq 0$$

- IN PRINCIPLE, THE VELOCITY, x_2, IS UNBOUNDED
- x_3 = HORIZONTAL DISTANCE IS SIMILARLY UNBOUNDED
- AS IS THE HORIZONTAL VELOCITY, x_4
- THE MASS, $m(t)$, IS LOWER BOUNDED BY THE ROCKET'S MASS AND UPPER BOUNDED BY THE ROCKET'S MASS PLUS FULL FUEL LOAD:

$$0 < M_{min} \leq x_5 \leq M_{max}$$

ALSO, SINCE THE ROCKET STARTS AT X=0, Y=0:

$$x_1(t_0) = 0 \qquad x_3(t_0) = 0$$

SINCE IT STARTS FROM REST:

$$X_2(t_0)=0 \qquad X_4(t_0)=0$$

(b) AN ADDITIONAL CONSTAINT IS OBVIOUSLY

$$Y(t_f) = X_3(t_f) = 3 \text{ MILES}$$

IF t_f IS SPECIFIED, OUR PERFORMANCE MEASURE IS

$$J = -X(t_f) = -X_1(t_f)$$

WHERE WE USE THE MINUS SIGN TO DENOTE THAT WE WISH TO MINIMIZE J.

(c) ADDITIONAL CONSTRAINTS ARE

$$X(t_f) = 500 \text{ MILES} = X_1(t_f)$$

$$Y(t_f) = 3 \text{ MILES} = X_3(t_f)$$

WHERE $t_f = 2.5$ MIN.

A PERFORMANCE MEASURE TO MAXIMIZE $m(t) = X_5$ IS

$$J = - X_5(t_f)$$

AGAIN, WE WISH TO MINIMIZE J.

[illegible] [illegible] Nov. 5, 1976

(1)(a) Find an $\phi(x, \dot{x}, t)$ such that the Euler-Lagrange equation associated with

$$J = \int_0^1 \phi(x, \dot{x}, t)\, dt$$

is

$$\ddot{x} + \alpha(x + \varepsilon x^3) = 0 .$$

(b) For $\varepsilon = 0$, find the extremum trajectory $\hat{x}(t)$ of (a) subject to the boundary conditions $x(0) = x(1) = 0$ and show that your answer obeys the Weierstrass (E-function) condition.

(2) Determine the extremals for the functional

$$J(x) = \int_0^4 (\dot{x}(t) - 1)^2 (\dot{x}(t) + 1)^2 dt$$

which have only one corner. The boundary conditions are $x(0) = 0$, $x(4) = 2$.

(3) Consider the linear regulator problem in which the plant is described by

$$\dot{x} = Ax + Bu$$

and the performance measure to be minimized is

$$J = \frac{1}{2} x^T(t_f) S x(t_f) + \frac{1}{2}\int_{t_0}^{t_f} (x^T Q x + u^T R u)\, dt$$

where S and Q are symm. semidefinite matrices, $R(t)$ is symm. positive definite.

(a) Show that one may solve the optimal feedback control law by solving a matrix Ricati equation

$$\dot{P} = -PA - A^T P + PBR^{-1}B^T P - Q$$

where the symm. matrix P is related to the Lagrange multiplier λ by $\lambda = Px$.

(b) Determine the optimal control law for a 2nd-order system with

$$A = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad S = 0, \quad Q = I, \quad R = 1, \quad t_f = \infty$$

such that the resulting closed-loop system is stable.

(1) a. $J = \int_0^1 \phi(x, \dot{x}, t)\, dt$

26

Euler - Lagrange:

$$\frac{\delta\phi}{\delta x} - \frac{d}{dt}\frac{\delta\phi}{\delta\dot{x}} = 0 = \ddot{x} + \alpha(x + \xi x^3)$$

$$\phi = -\frac{1}{2}\dot{x}^2 + \alpha\left(\frac{x^2}{2} + \frac{\xi}{4}x^4\right)$$

b. $\xi = 0$

$$\Rightarrow \ddot{x} + \alpha x = 0 \quad \Leftarrow \text{Euler Lagrange}$$

$$\ddot{x} = -\alpha x$$

$$\Rightarrow x = A\cos\sqrt{\alpha}\, t + B\sin\sqrt{\alpha}\, t$$

$$x(0) = 0 = A \Rightarrow A = 0$$

$$\therefore x = B\sin\sqrt{\alpha}\, t_0$$

$$0 = x(1) = B\sin\sqrt{\alpha} \Rightarrow [illegible] = 0$$

① If α is fixed, we must require that $B = 0 \Rightarrow x(t) = 0$

② If we may choose α, then let $\sqrt{\alpha} = n\pi\sqrt{3}$ leave B unspecified (~~[illegible]~~

We will choose ① to work. b $\Rightarrow$

Weierstrass E function

$$E = \Phi(\dot{X}) - \Phi(\dot{x}) + (\dot{X} - \dot{x})\frac{\delta\Phi}{\delta t} \geq 0$$

$$\Phi(X) = -\frac{1}{2}\dot{X}^2 + \alpha\frac{X^2}{2}$$

The weierstrass E function is obviously satisfied for the ~~[illegible]~~ 0 solution

$$x = 0. \Rightarrow \acute{x} = 0. \qquad \Phi(\acute{x} = 0) = 0$$

$$E = \Phi(\acute{X}) - 0 + \dot{X}\cancel{\frac{\delta\Phi}{\delta t}}$$

$$= \Phi(\acute{X}) \geq 0^{0}$$

This is true since $\Phi(\acute{X})$ is always non-negative.

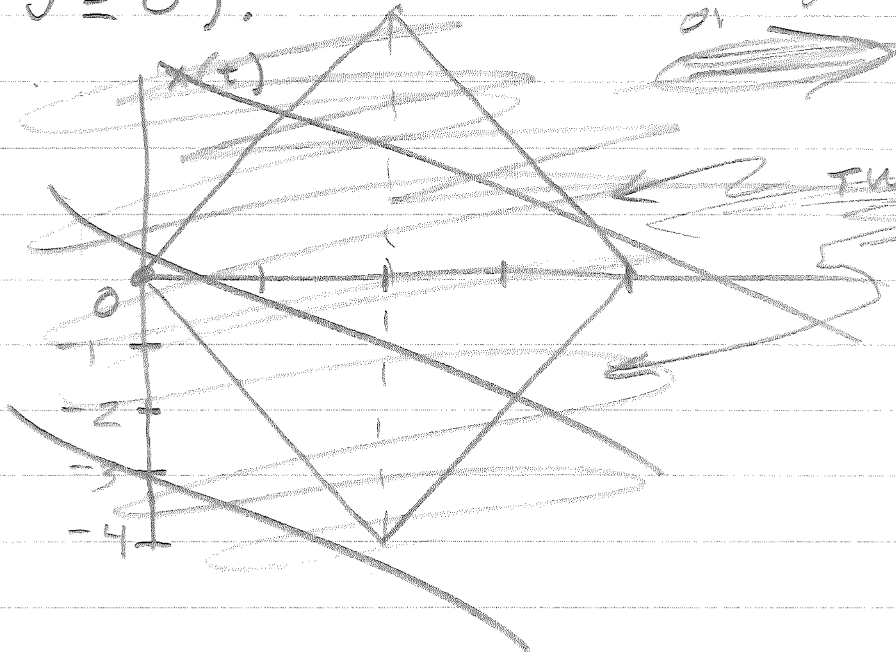
2. $J = \int_0^4 (\dot{x}-1)^2(\dot{x}+1)^2 dt$

MUST OBEY Euler Lagrange:

$$\frac{\delta \phi}{\delta x} - \frac{d}{dt}\frac{\delta \phi}{\delta \dot{x}} = 0$$

28

But, by inspection, $J = 0$ for $\dot{x} = 1, -1$ (Note, in general, $J \geq 0$).

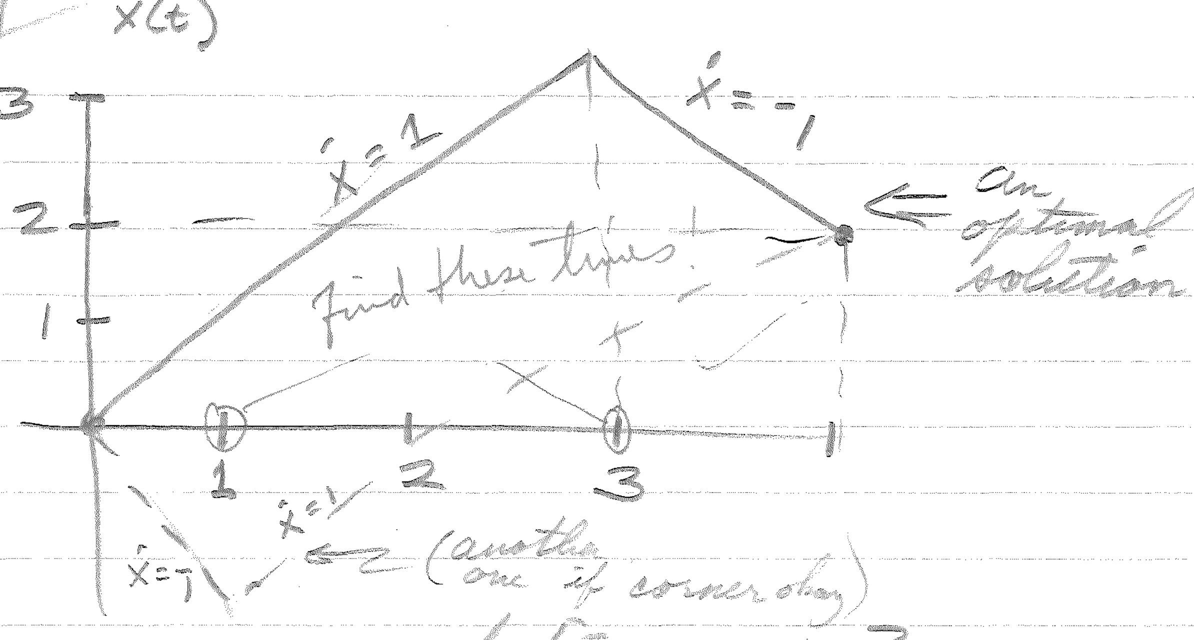


$x(0) = 0$
$x(4) = 2$

TWO SOLUTIONS

$\dot{x} = (1, -1)$ is what we'd of gotten by applying ~~var~~ variational principle. $\Longrightarrow$

By inspection



Are corner conditions ok?

$$\left.\frac{\delta\Phi}{\delta\dot{x}}\right|_{3-} \overset{?}{=} \left.\frac{\delta\Phi}{\delta\dot{x}}\right|_{3+} \quad ①$$

$$\phi = (\dot{x}-1)^2(\dot{x}+1)^2$$

$$\frac{\delta\Phi}{\delta\dot{x}} = 2(\dot{x}-1)(\dot{x}+1)^2 + 2(\dot{x}-1)^2(\dot{x}+1)$$

$$\Longrightarrow \left.\frac{\delta\Phi}{\delta\dot{x}}\right|_{3+} = \left.\frac{\delta\Phi}{\delta\dot{x}}\right|_{3-} = 0$$

$$\left.\phi - \underset{0}{\cancel{\dot{x}\frac{\delta\Phi}{\delta\dot{x}}}}\right|_{3+} \overset{?}{=} \left.\phi - \underset{0}{\cancel{\dot{x}\frac{\delta\Phi}{\delta\dot{x}}}}\right|_{3-} \quad (2$$

$$\phi|_{3-} = 0 = \phi|_{3+}$$

Thus, corner conditions are satisfied.

3. $H = \frac{1}{2}x^T Q x + \frac{1}{2}u^T R u + \lambda^T (Ax + Bu)$

20

$\frac{\delta H}{\delta u} = 0 = Ru + B^T\lambda \Rightarrow u = -R^{-1}B^T\lambda$

$\dot{\lambda} = -\frac{\delta H}{\delta x} = -(Qx + A^T\lambda)$

$\dot{x} = Ax + Bu$ $\quad \lambda(t_f) = Sx(t_f)$.

assume

$\lambda = Px$

$\Rightarrow \dot{\lambda} = P\dot{x} + \dot{P}x$

$-(Qx + A^T\lambda) = P[Ax + Bu] + \dot{P}x$

$-(Qx + A^TPx) = PAx + PBu + \dot{P}x$

$-(Qx + A^TPx) = PAx + PBR^{-1}B^T\lambda + \dot{P}x$

$-(\underline{Qx} + A^T\underline{Px}) = \underline{PAx} - \underline{PBR^{-1}B^TPx} + \underline{\dot{P}x}$

$\Rightarrow [\dot{P} + PA + A^TP - PBR^{-1}B^TPx + Q]x = 0$

~~Gives Riccati~~ =

$\begin{cases} \dot{P} + PA + A^TP - PBR^{-1}B^TPx + Q = 0 \\ P(t_f) = S \end{cases}$

b) We should have

$$u = -R^{-1}B^T\lambda = -B^T\lambda = [1, 0]\begin{bmatrix}\lambda_1 \\ \lambda_2\end{bmatrix}$$

$$\Rightarrow u = \lambda_1 \qquad = \lambda_1$$

$$\dot{\lambda} = -Qx + A^T\lambda$$

$$= -x + \begin{bmatrix}-1 & 0 \\ 1 & 0\end{bmatrix}^T \lambda$$

$$= -x + \begin{bmatrix}-1 & 1 \\ 0 & 0\end{bmatrix}\lambda$$

$$\begin{bmatrix}\dot{\lambda}_1 \\ \dot{\lambda}_2\end{bmatrix} = -\begin{bmatrix}x_1 \\ x_2\end{bmatrix} + \begin{bmatrix}-\lambda_1 + \lambda_2 \\ 0\end{bmatrix}$$

$$\dot{\lambda}_2 = -x_2$$

$$\dot{\lambda}_1 = -x_1 - \lambda_1 + \lambda_2$$

$$\dot{x} = Ax + Bu$$

$t_f = \infty \Rightarrow$

$$\overset{0}{\cancel{\dot{P}}} = -PA - A^TP + PBR^{-1}B^TP - Q$$

$$-P\begin{bmatrix}-1 & 0\\ 1 & 0\end{bmatrix} - \begin{bmatrix}-1 & 1\\ 0 & 0\end{bmatrix}P$$

$$+ \underbrace{P\begin{bmatrix}1\\ 0\end{bmatrix}(1)\begin{bmatrix}1 & 0\end{bmatrix}P}_{P\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}P} - I$$

$$(-P)\begin{bmatrix}-1 & 0\\ 1 & 0\end{bmatrix} - \begin{bmatrix}-1 & 1\\ 0 & 0\end{bmatrix})P + P^2 - \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

$$\begin{bmatrix}P_{11} & P_{12}\\ P_{21} & P_{22}\end{bmatrix}\begin{bmatrix}-1 & 0\\ 1 & 0\end{bmatrix} - \begin{bmatrix}-1 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}P_{11} & P_{12}\\ P_{21} & P_{22}\end{bmatrix}$$

$P_{12} = P_{21} = \hat{P}$

$$\begin{bmatrix}P_{11} & P_{12}\\ P_{21} & P_{22}\end{bmatrix}\begin{bmatrix}P_{11} & P_{12}\\ P_{21} & P_{22}\end{bmatrix}^2 - \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} = 0$$

$$\begin{bmatrix}-P_{11}+\hat{P}_{12} & 0\\ -\hat{P}+P_{22} & 0\end{bmatrix} - \begin{bmatrix}-P_{11}+\hat{P} & -\hat{P}+P_{22}\\ 0 & 0\end{bmatrix}$$

$$+ \begin{bmatrix}P_{11}^2+\hat{P}^2 & P_1\hat{P} + \hat{P}P_2\\ \hat{P}P_1+\hat{P}P_2 & \hat{P}^2+P_2^2\end{bmatrix} - \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} = 0$$

$-P_{11} + \hat{P} + P_{11} - \hat{P} + P_{11}^2 + \hat{P}^2 - 1 = 0$

$\hat{P} - P_{22} + P_1\hat{P} + \hat{P}P_2 = 0$

$-\hat{P} + P_{22} + \hat{P}P_1 + \hat{P}P_2 = 0$

$\hat{P}^2 + P_2^2 - 1 = 0$

We could in principle solve them -

(This is too hard)

(1) (a) Suppose it is desired to extremize the performance measure

$$J = \int_{t_0}^{t_f} F(x, \dot{x}, t)\, dt$$

subject to the isoperimetric constraint

$$\int_{t_0}^{t_f} e(x, \dot{x}, t)\, dt = c.$$

Convert the isoperimetric constraint into the form of differential equation constraints and show that the corresponding Lagrange multipliers are constant.

(b) In the circuit shown in Fig. 1, design an input voltage source $V_s(t)$ over the interval $0 \le t \le T$ such that the average value of the output voltage $V_c(t)$ is maximized over the time interval $[0, T]$ subject to the constraint that the input energy E_{in} is fixed over the same interval.

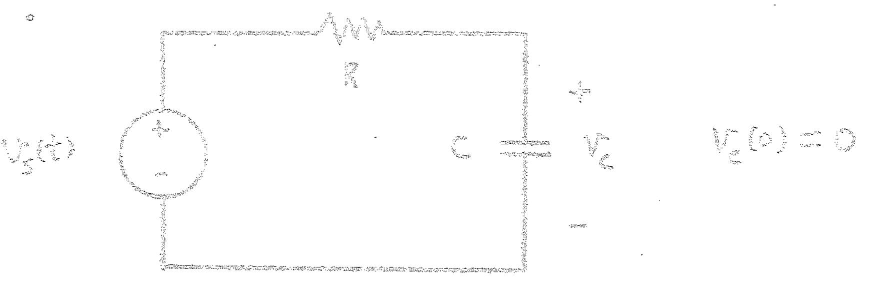


Fig. 1

(2) Consider the system

$$\dot{x} = ax + u, \qquad a > 0$$

which is to be transferred from an initial state x_0 to the origin. The admissible controls are constrained by $|u(t)| \le 1$. Assume that x_0 is such that the origin can be reached by applying admissible controls.

(a) Determine the time-optimal control law.

(b) Determine the fuel-optimal control law.

(c) Compare the two optimal control laws and explain the difference between this problem and the case where $a < 0$.

[illegible] the [illegible]
nonlinear two point boundary value problem:

$$\dot{x} = f(x, \lambda, t) \qquad x(t_0) = x_0$$

$$\dot{\lambda} = -\frac{\partial H(x, \lambda, t)}{\partial x} \qquad \lambda(t_f) = \lambda_f$$

[illegible] Discuss the singular solution possibility for the system

$$\dot{x}_1 = x_2 + u(t), \qquad x_1(0) = x_{10} \qquad x_1(t_f) = 0$$

$$\dot{x}_2 = -u(t), \qquad x_2(0) = x_{20} \qquad x_2(t_f) = 0$$

where $|u(t)| \leq 1$ and the performance index to be minimized is given by

$$J = \frac{1}{2}\int_0^\infty x_1^2 \, dt.$$