DYNAMIC NASH GAMES

Mar. 26, 99 (9)

Given a dynamic system with two controllers (players) u_1 and u_2

$$\dot{x} = f(x, u_1, u_2), \qquad x(t_0) = x_0$$

and performance criteria

$$J_1(x(t_0), t_0) = \int_{t_0}^{t_f} L_1(x, u_1, u_2)\,dt + g_1(x(t_f))$$

$$J_2(x(t_0), t_0) = \int_{t_0}^{t_f} L_2(x, u_1, u_2)\,dt + g_2(x(t_f))$$

Rational ~~decision~~ behavior of each player is to minimize his/her own losses assuming that the other player does the same. Hence, both players have to formulate their own dynamic optimization problem under the assumption that the other player will use his/her Nash optimal strategy.

Nash equilibrium is defined as before

$$J_1(u_1^*, u_2^*) \le J_1(u_1, u_2^*), \qquad u_1^* = u_1^{Nash}$$

$$J_2(u_1^*, u_2^*) \le J_2(u_1^*, u_2), \qquad u_2^* = u_2^{Nash}$$

Hence the part of the above inequalities has to be satisfied at the Nash equilibrium.

If both players use their optimal Nash strategies u_1^* and u_2^* we have

$$\dot{x}^* = f(x^*, u_1^*, u_2^*), \qquad x^*(t_0) = x_0$$

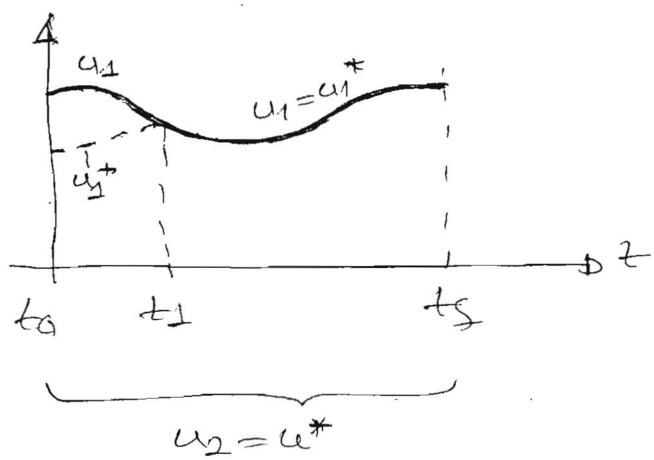
and

$$J_1^*(x(t_0),t_0) \triangleq \int_{t_0}^{t_f} L_1(x^*, u_1^*, u_2^*)\,dt + g_1(x^*(t_f))$$

$$J_2^*(x(t_0),t_0) \triangleq \int_{t_0}^{t_f} L_2(x^*, u_1^*, u_2^*)\,dt + g_2(x^*(t_f))$$

Let us first derive the B-H-J equation for P1 (player 1).

Assume that P2 uses u_2^* along the entire time interval (t_0, t_f) - the interval game is played on.



Let P1 uses u_1^* during (t_1, t_f) and any other u_1 from t_0 to t_1

$$J_1(x(t_0),t_0) = \int_{t_0}^{t_1} L_1(x, u_1, u_2^*)\,dt + \underbrace{\int_{t_1}^{t_f} L_1(x, u_1^*, u_2)\,dt + g_1(x^*(t_f))}_{J^*(x(t_1),t_1)}$$

It is obvious that

$$J_1^*(x(t_0),t_0) \le J_1(x(t_0),t_0) = \int_{t_0}^{t_1} L_1(x, u_1, u_2^*)\,dt + J^*(x(t_1),t_1)$$

$\Rightarrow$

$$J_1^*(x(t_0),t_0) - J_1^*(x(t_1),t_1) \le \int_{t_0}^{t_1} L_1(x, u_1, u_2^*)\,dt \quad \Big/ \frac{1}{t_1 - t_0}$$

$$-\lim_{t_1 \to t_0} \frac{\left(J_1^*(x(t_1),t_1) - J^*(x(t_0),t_0)\right)}{t_1 - t_0} \le \lim_{t_1 \to t_0} \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} L_1(x, u_1, u_2^*)\,dt$$

$$-\frac{d}{dt_0}\left(J_1^*(x(t_0),t_0)\right) \le L_1(x_1(t_0), u_1(t_0), u_2^*(t_0))$$

since t_0 is arbitrary, can be any t we can use t instead t_0 in the following derivations. The above total derivative can be evaluated and we get

$$-\frac{\partial}{\partial t} J_1^*(x(t),t) - \frac{\partial J_1^*}{\partial x} \cdot \frac{dx^*}{dt} \le L_1(x(t), u_1(t), u_2^*(t))$$

Note the equality hold if u_1 is u_1^* on the entire time interval, hence we have

$$(1)\quad \boxed{-\frac{\partial}{\partial t} J_1^*(x^*(t),t) - \frac{\partial J_1^*}{\partial x} \cdot f(x^*, u_1^*, u_2^*) = L_1(x^*, u_1^*, u_2^*)}$$

This is the H-B-J equation for P1. Similarly, ~~if~~ for P2

$$(2)\quad \boxed{-\frac{\partial}{\partial t} J_2^*(x^*(t),t) - \frac{\partial J_2^*}{\partial x} f(x^*, u_1^*, u_2^*) = L_2(x^*, u_1^*, u_2^*)}$$

is the H-B-J equation for P2. The second equation is obtained by optimizing $J_2 \big|_{u_1=u_1^*}$ along $\dot{x} = f(x, u_1^*, u_2)$.

The boundary conditions for the above partial differential equations are easily obtain

$$J_1(x(t_f), t_f) = \int_{t_f}^{t_f} (\,) + g_1(x(t_f)) = g_1(x(t_f))$$

and

$$J_2(x(t_f), t_f) = \int_{t_f}^{t_f} (\,) + g_2(x(t_f)) = g_2(x(t_f))$$

The H-B-J equations for Nash games can be also written in the form

$$-\frac{\partial J_1^*}{\partial t} = \min_{u_1} \left\{ L_1(x, u_1, u_2^*) + \frac{\partial J_1^*}{\partial x} f(x, u_1, u_2^*) \right\}$$

To justify this step rigorously we have to use more rigorous derivations of the H-B-J equation as given on pages 4a and 4b. Also

$$-\frac{\partial J_2^*}{\partial t} = \min_{u_2} \left\{ L_2(x, u_1^*, u_2) + \frac{\partial J_2^*}{\partial x} f(x, u_1^*, u_2) \right\}$$

Note when J_1 and J_2 do not depend explicitly on time we have $\frac{\partial J_1}{\partial t} = 0$ and $\frac{\partial J_2}{\partial t} = 0$

Let us introduce Hamiltonians for both players

$$H_1 = L_1(x, u_1, u_2^*) + \frac{\partial J_1^*}{\partial x} f(x, u_1, u_2^*)$$

$$H_2 = L_2(x, u_1^*, u_2) + \frac{\partial J_2^*}{\partial x} f(x, u_1^*, u_2)$$

Then the necessary conditions for optimality can be recorded as

$$\left.\begin{aligned} \frac{\partial H_1}{\partial u_1} = 0 &\Rightarrow \varphi_1\left(x, u_1, u_2^*, \frac{\partial J_1^*}{\partial x}\right) = 0 \\ \frac{\partial H_2}{\partial u_2} = 0 &\Rightarrow \varphi_2\left(x, u_1^*, u_2, \frac{\partial J_2^*}{\partial x}\right) = 0 \end{aligned}\right\} \Rightarrow \begin{aligned} u_1^* &= \psi_1\left(x, \frac{\partial J_1^*}{\partial x}, \frac{\partial J_2^*}{\partial x}\right) \\ u_2^* &= \psi_2\left(x, \frac{\partial J_1^*}{\partial x}, \frac{\partial J_2^*}{\partial x}\right) \end{aligned}$$

Substituting u_1^* and u_2^* (obtained above) into (1) and (2) implies the system of two partial differential equations with respect to $\frac{\partial J_1}{\partial t}$, $\frac{\partial J_1}{\partial x}$, $\frac{\partial J_2}{\partial t}$, $\frac{\partial J_2}{\partial x}$. This system of PDE is very difficult for solving (in general)

$$\mathbf{H} = \mathbf{0}, \qquad \mathbf{Q} = \begin{bmatrix} 0.25 & 0.00 \\ 0.00 & 0.05 \end{bmatrix}, \quad \text{and} \quad R = 0.05.$$

The optimal feedback gain matrix $\mathbf{F}(k)$ is shown in Fig. 3-9(a) for $N = 200$. Looking backward from $k = 199$, we observe that at $k \approx 130$ the $\mathbf{F}(k)$ matrix has reached the steady-state value

$$\mathbf{F}(k) = [-0.5522 \quad -5.9668], \qquad 0 \leq k \leq 130. \tag{3.10-24}$$

The optimal control history and the optimal trajectory for $\mathbf{x}(0) = [2 \quad 1]^T$ are shown in Fig. 3-9(b). Notice that the optimal trajectory has essentially reached $\mathbf{0}$ at $k = 100$. Thus, we would expect that insignificant performance degradation would be caused by simply using the steady-state value of $\mathbf{F}$ given in (3.10-24) rather than $\mathbf{F}(k)$ as specified in Fig. 3-9(a).

3.11 THE HAMILTON-JACOBI-BELLMAN EQUATION

In our initial exposure to dynamic programming, we approximated continuously operating systems by discrete systems. This approach leads to a recurrence relation that is ideally suited for digital computer solution. In this section we shall consider an alternative approach which leads to a nonlinear *partial* differential equation—the Hamilton-Jacobi-Bellman (H-J-B) equation. The derivation that will be given in this section parallels the development of the functional recurrence equation (3.7-18) in Section 3.7.

The process described by the state equation

$$\dot{\mathbf{x}}(t) = \mathbf{a}(\mathbf{x}(t), \mathbf{u}(t), t) \tag{3.11-1}$$

is to be controlled to minimize the performance measure

$$J = h(\mathbf{x}(t_f), t_f) + \int_{t_0}^{t_f} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau)\, d\tau, \tag{3.11-2}$$

where h and g are specified functions, t_0 and t_f are fixed, and τ is a dummy variable of integration. Let us now use the *imbedding principle* to include this problem in a larger class of problems by considering the performance measure

$$J(\mathbf{x}(t), t, \underset{t \leq \tau \leq t_f}{\mathbf{u}(\tau)}) = h(\mathbf{x}(t_f), t_f) + \int_{t}^{t_f} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau)\, d\tau, \tag{3.11-3}$$

where t can be any value less than or equal to t_f, and $\mathbf{x}(t)$ can be any admissible state value. Notice that the performance measure will depend on the

numerical values for $\mathbf{x}(t)$ and t, and on the optimal control history in the interval $[t, t_f]$.

Let us now attempt to determine the controls that minimize (3.11-3) for all admissible $\mathbf{x}(t)$, and for all $t \leq t_f$. The minimum cost function is then

$$J^*(\mathbf{x}(t), t) = \min_{\substack{\mathbf{u}(\tau) \\ t \leq \tau \leq t_f}} \left\{ \int_t^{t_f} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau)\, d\tau + h(\mathbf{x}(t_f), t_f) \right\}. \tag{3.11-4}$$

By subdividing the interval, we obtain

$$J^*(\mathbf{x}(t), t) = \min_{\substack{\mathbf{u}(\tau) \\ t \leq \tau \leq t_f}} \left\{ \int_t^{t+\Delta t} g\, d\tau + \int_{t+\Delta t}^{t_f} g\, d\tau + h(\mathbf{x}(t_f), t_f) \right\}. \tag{3.11-5}$$

The principle of optimality requires that

$$J^*(\mathbf{x}(t), t) = \min_{\substack{\mathbf{u}(\tau) \\ t \leq \tau \leq t+\Delta t}} \left\{ \int_t^{t+\Delta t} g\, d\tau + J^*(\mathbf{x}(t + \Delta t), t + \Delta t) \right\}, \tag{3.11-6}$$

where $J^*(\mathbf{x}(t + \Delta t), t + \Delta t)$ is the minimum cost of the process for the time interval $t + \Delta t \leq \tau \leq t_f$ with "initial" state $\mathbf{x}(t + \Delta t)$.

Assuming that the second partial derivatives of J^* exist and are bounded, we can expand $J^*(\mathbf{x}(t + \Delta t), t + \Delta t)$ in a Taylor series about the point $(\mathbf{x}(t), t)$ to obtain

$$\begin{aligned} J^*(\mathbf{x}(t), t) = \min_{\substack{\mathbf{u}(\tau) \\ t \leq \tau \leq t+\Delta t}} \Bigg\{ & \int_t^{t+\Delta t} g\, d\tau + J^*(\mathbf{x}(t), t) + \left[\frac{\partial J^*}{\partial t}(\mathbf{x}(t), t)\right] \Delta t \\ & + \left[\frac{\partial J^*}{\partial \mathbf{x}}(\mathbf{x}(t), t)\right]^T \left[\mathbf{x}(t + \Delta t) - \mathbf{x}(t)\right] \\ & + \text{terms of higher order} \Bigg\}. \end{aligned} \tag{3.11-7}$$

Now for small Δt

$$\begin{aligned} J^*(\mathbf{x}(t), t) = \min_{\mathbf{u}(t)} \{ & g(\mathbf{x}(t), \mathbf{u}(t), t)\, \Delta t + J^*(\mathbf{x}(t), t) \\ & + J_t^*(\mathbf{x}(t), t)\, \Delta t + J_{\mathbf{x}}^{*T}(\mathbf{x}(t), t)[\mathbf{a}(\mathbf{x}(t), \mathbf{u}(t), t)]\, \Delta t \\ & + o(\Delta t)\},\dagger \end{aligned} \tag{3.11-8}$$

Also true for any J not just optimal one

where $o(\Delta t)$ denotes the terms containing $[\Delta t]^2$ and higher orders of Δt that arise from the approximation of the integral and the truncation of the Taylor series expansion. Next, removing the terms involving $J^*(\mathbf{x}(t), t)$ and $J_t^*(\mathbf{x}(t), t)$

$\dagger$ $J_{\mathbf{x}}^* \triangleq \dfrac{\partial J^*}{\partial \mathbf{x}} = \left[\dfrac{\partial J^*}{\partial x_1} \quad \dfrac{\partial J^*}{\partial x_2} \quad \cdots \quad \dfrac{\partial J^*}{\partial x_n}\right]^T$ and $J_t^* \triangleq \dfrac{\partial J^*}{\partial t}$.

40

from the minimization [since they do not depend on $\mathbf{u}(t)$], we obtain

$$0 = J_t^*(\mathbf{x}(t), t)\,\Delta t + \min_{\mathbf{u}(t)} \{g(\mathbf{x}(t), \mathbf{u}(t), t)\,\Delta t + J_{\mathbf{x}}^{*T}(\mathbf{x}(t), t)[\mathbf{a}(\mathbf{x}(t), \mathbf{u}(t), t)]\,\Delta t + o(\Delta t)\}. \tag{3.11-9}$$

Dividing by Δt and taking the limit as $\Delta t \longrightarrow 0$ gives†

$$\boxed{0 = J_t^*(\mathbf{x}(t), t) + \min_{\mathbf{u}(t)} \{g(\mathbf{x}(t), \mathbf{u}(t), t) + J_{\mathbf{x}}^{*T}(\mathbf{x}(t), t)[\mathbf{a}(\mathbf{x}(t), \mathbf{u}(t), t)]\}.} \tag{3.11-10}$$

To find the boundary value for this partial differential equation, set $t = t_f$; from Eq. (3.11-4) it is apparent that

$$\boxed{J^*(\mathbf{x}(t_f), t_f) = h(\mathbf{x}(t_f), t_f).} \tag{3.11-11}$$

We define the Hamiltonian $\mathscr{H}$ as

$$\mathscr{H}(\mathbf{x}(t), \mathbf{u}(t), J_{\mathbf{x}}^*, t) \triangleq g(\mathbf{x}(t), \mathbf{u}(t), t) + J_{\mathbf{x}}^{*T}(\mathbf{x}(t), t)[\mathbf{a}(\mathbf{x}(t), \mathbf{u}(t), t)] \tag{3.11-12}$$

and

$$\mathscr{H}(\mathbf{x}(t), \mathbf{u}^*(\mathbf{x}(t), J_{\mathbf{x}}^*, t), J_{\mathbf{x}}^*, t) = \min_{\mathbf{u}(t)} \mathscr{H}(\mathbf{x}(t), \mathbf{u}(t), J_{\mathbf{x}}^*, t), \tag{3.11-13}$$

since the minimizing control will depend on $\mathbf{x}$, $J_{\mathbf{x}}^*$, and t. Using these definitions, we have obtained the Hamilton-Jacobi equation

$$\boxed{0 = J_t^*(\mathbf{x}(t), t) + \mathscr{H}(\mathbf{x}(t), \mathbf{u}^*(\mathbf{x}(t), J_{\mathbf{x}}^*, t), J_{\mathbf{x}}^*, t).} \tag{3.11-10a}$$

This equation is the continuous-time analog of Bellman's recurrence equation (3.7-18); therefore, we shall refer to (3.11-10a) as the "Hamilton-Jacobi-Bellman equation."

Example 3.11-1. A first-order system is described by the differential equation

$$\dot{x}(t) = x(t) + u(t); \tag{3.11-14}$$

† $\lim_{\Delta t \to 0} \left|\frac{o(\Delta t)}{\Delta t}\right| = 0.$

it is desired to find the control law that minimizes the performance measure

$$J = \tfrac{1}{4}x^2(T) + \int_0^T \tfrac{1}{4}u^2(t)\,dt. \tag{3.11-15}$$

The final time T is specified, and the admissible state and control values are not constrained by any boundaries.

Substituting $g = \frac{1}{4}u^2(t)$ and $a = x(t) + u(t)$ into Eq. (3.11-12), we find that the Hamiltonian is (omitting the arguments of J_x^*)

$$\mathscr{H}(x(t), u(t), J_x^*, t) = \tfrac{1}{4}u^2(t) + J_x^*[x(t) + u(t)], \tag{3.11-16}$$

and since the control is unconstrained, a necessary condition that the optimal control must satisfy is

$$\frac{\partial \mathscr{H}}{\partial u} = \frac{1}{2}u(t) + J_x^*(x(t), t) = 0. \tag{3.11-17}$$

Observe that

$$\frac{\partial^2 \mathscr{H}}{\partial u^2} = \frac{1}{2} > 0; \tag{3.11-18}$$

thus, the control that satisfies Eq. (3.11-17) does minimize $\mathscr{H}$. From (3.11-17)

$$u^*(t) = -2J_x^*(x(t), t), \tag{3.11-19}$$

which when substituted in the Hamilton-Jacobi-Bellman equation gives

$$\begin{aligned} 0 &= J_t^* + \tfrac{1}{4}[-2J_x^*]^2 + [J_x^*]x(t) - 2[J_x^*]^2 \\ &= J_t^* - [J_x^*]^2 + [J_x^*]x(t). \end{aligned} \tag{3.11-20}$$

The boundary value is, from (3.11-15),

$$J^*(x(T), T) = \tfrac{1}{4}x^2(T). \tag{3.11-21}$$

One way to solve the Hamilton-Jacobi-Bellman equation is to guess a form for the solution and see if it can be made to satisfy the differential equation and the boundary conditions. Let us assume a solution of the form

$$J^*(x(t), t) = \tfrac{1}{2}K(t)x^2(t), \tag{3.11-22}$$

where $K(t)$ represents an unknown scalar function of t that is to be determined. Notice that

$$J_x^*(x(t), t) = K(t)x(t), \tag{3.11-23}$$

Let us now differentiate the B-H-J equation with respect to x

$$-\frac{\partial J_1}{\partial t} = L_1(x, u_1, u_2) + \frac{\partial J_1}{\partial x} f(x, u_1, u_2) \qquad \text{(H-B-J for P}$$

$$-\frac{\partial^2 J_1}{\partial x \partial t} = \frac{\partial L_1}{\partial x} + f \frac{\partial^2 J_1}{\partial x^2} + \frac{\partial J_1}{\partial x} \cdot \frac{\partial f}{\partial x}$$

$$+ \underbrace{\frac{\partial}{\partial u_1}\left(L_1 + \frac{\partial J_1}{\partial x} f\right)}_{=0 = \frac{\partial}{\partial u_1} H_1} \frac{\partial u_1}{\partial x} + \frac{\partial}{\partial u_2}\underbrace{\left(L_1 + \frac{\partial J_1}{\partial x} f\right)}_{H_1} \frac{\partial u_2}{\partial x}$$

$$-\frac{\partial^2 J_1}{\partial x \partial t} - f \frac{\partial^2 J_1}{\partial x^2} = \frac{\partial L_1}{\partial x} + \frac{\partial J_1}{\partial x} \frac{\partial f}{\partial x} + \frac{\partial}{\partial u_2}(H_1) \frac{\partial u_2}{\partial x}$$

Introducing the co-state variable $p_1 = \left(\frac{\partial J_1}{\partial x}\right)^T$, we get

$$-\frac{d}{dt}(p_1^T) = \underbrace{\frac{\partial L_1}{\partial x} + p_1^T \frac{\partial f}{\partial x}}_{=\frac{\partial}{\partial x} H_1} + \frac{\partial}{\partial u_2}(H_1) \frac{\partial u_2}{\partial x}$$

$$\frac{d}{dt} p_1 = \boxed{\dot{p}_1 = -\left(\frac{\partial H_1}{\partial x}\right)^T - \left(\frac{\partial H_2}{\partial u_2} \cdot \frac{\partial u_2}{\partial x}\right)^T} \qquad (3)$$

Similarly

$$\boxed{\dot{p}_2 = -\left(\frac{\partial H_2}{\partial x}\right)^T - \left(\frac{\partial H_1}{\partial u_1} \frac{\partial u_1}{\partial x}\right)^T} \qquad (4)$$

Note since $J_1(x(t_f), t_f) = g_1(x(t_f))$ and $p_1 = \frac{\partial J_1}{\partial x}$

$J_2(x(t_f), t_f) = g_2(x(t_f))$ $\qquad p_2 = \frac{\partial J_2}{\partial x}$

the boundary conditions for p_1 and p_2 are

$$p_1(t_f) = \frac{\partial g_1(x(t_f))}{\partial x}$$

$$p_2(t_f) = \frac{\partial g_2(x(t_f))}{\partial x}$$

Summary of necessary conditions for Nash optimality:

Given:

$$\dot{x} = f(x, u_1, u_2), \qquad x(t_0) = x_0$$

$$J_1 = \int_{t_0}^{t_f} L_1(x, u_1, u_2)\,dt + g_1(x(t_f))$$

$$J_2 = \int_{t_0}^{t_f} L_2(x, u_1, u_2)\,dt + g_2(x(t_f))$$

Form the Hamiltonians:

$$H_1 = L_1 + p_1^T f, \qquad H_2 = L_2 + p_2^T f$$

Necessary conditions:

$$\boxed{\begin{aligned}
\dot{x} &= \frac{\partial H_i}{\partial p_i^T}, \quad i = 1, 2, \Rightarrow \dot{x} = f(x, u_1, u_2), \quad x(t_0) = x_0 \\
\dot{p}_1^T &= -\frac{\partial H_1}{\partial x} - \frac{\partial H_1}{\partial u_2}\left(\frac{\partial u_2}{\partial x}\right), \quad p_1(t_f) = \frac{\partial g_1(t_f)}{\partial x} \\
\dot{p}_2^T &= -\frac{\partial H_2}{\partial x} - \frac{\partial H_2}{\partial u_1}\left(\frac{\partial u_1}{\partial x}\right), \quad p_2(t_f) = \frac{\partial g_2(t_f)}{\partial x} \\
0 &= \frac{\partial H_1}{\partial u_1} \\
0 &= \frac{\partial H_2}{\partial u_2}
\end{aligned}}$$

In general for N-players $\Rightarrow H_i = L_i + p_i^T f$

$$\begin{cases}
\dot{x} = f(x, u_1, u_2, \dots, u_N) = \dfrac{\partial H_i}{\partial p_i}, \quad x(t_0) = x_0 \\
\dot{p}_i^T = -\dfrac{\partial H_i}{\partial x} - \displaystyle\sum_{\substack{j \neq i \\ i=1}}^{N} \left(\frac{\partial H_i}{\partial u_j}\right)\left(\frac{\partial u_j}{\partial x}\right), \quad i = 1, 2, \dots, N, \quad p_i(t_f) = \dfrac{\partial g_i(t_f)}{\partial x} \\
0 = \dfrac{\partial H_i}{\partial u_i}, \quad i = 1, 2, \dots, N
\end{cases}$$