Chapter 2
Deterministic and Stochastic Differential Games

This chapter introduces the theory of deterministic and stochastic differential games, including the dynamic optimization techniques, (stochastic) differential games and their solution concepts, which will lay a foundation for later study.

2.1 Dynamic Optimization Techniques

Consider the dynamic optimization problem in which the single decision-maker:

$$\max_{u}\left\{\int_{t_0}^{T} g[s,x(s),u(s)]ds + q(x(T))\right\}, \tag{2.1}$$

Subject to the vector-valued differential equation:

$$\dot{x}(s) = f[s,x(s),u(s)]ds, \quad x(t_0) = x_0, \tag{2.2}$$

where $x(s)\in X\subset\mathbb{R}^n$ denotes the state variables of game, and $u\in\mathcal{U}$ is the control. The functions $f[s,x,u]$, $g[s,x,u]$ and $q(x)$ are differentiable functions.

Dynamic programming and optimal control are used to identify optimal solutions for the problem (2.1)–(2.2).

2.1.1 Dynamic Programming

A frequently adopted approach to dynamic optimization problems is the technique of dynamic programming. The technique was developed by Bellman (1957). The technique is given in Theorem 2.1.1 below.

C.-k. Zhang et al., *Non-cooperative Stochastic Differential Game Theory of Generalized Markov Jump Linear Systems*, Studies in Systems, Decision and Control 67, DOI 10.1007/978-3-319-40587-2_2

Theorem 2.1.1 (Bellman's Dynamic Programming) *A set of controls $u^*(t) = \phi^*(t,x)$ constitutes an optimal solution to the control problem (2.1)–(2.2) if there exist continuously differentiable functions $V(t,s)$ defined on $[t_0, T] \times \mathbb{R}^n \to \mathbb{R}$ and satisfying the following Bellman equation*:

$$
\begin{aligned}
-V_t(t,x) &= \max_u \{g[t,x,u] + V_x(t,x)f[t,x,u]\} \\
&= \{g[t,x,\phi^*(t,x)] + V_x(t,x)f[t,x,\phi^*(t,x)]\}, \\
V(T,x) &= q(x).
\end{aligned}
$$

Proof Define the maximized payoff at time t with current state x as a value function in the form:

$$
\begin{aligned}
V(t,x) &= \max_u \left[\int_t^T g(s,x(s),u(s))ds + q(x(T)) \right] \\
&= \int_t^T g[s,x^*(s),\phi^*(s,x^*(s))]ds + q(x^*(T)).
\end{aligned}
$$

Satisfying the boundary condition

$$V(T,x^*(T)) = q(x^*(T)),$$

and

$$\dot{x}^*(s) = f[s,x^*(s),\phi^*(s,x^*(s))], \quad x^*(t_0) = x_0.$$

If in addition to $u^*(s) \equiv \phi^*(s,x)$, we are given another set of strategies, $u(s) \in \mathcal{U}$, with the corresponding terminating trajectory $x(s)$, then Theorem 2.1.1 implies

$$
\begin{aligned}
g(t,x,u) + V_x(t,x)f(t,x,u) + V_t(t,x) &\le 0, \\
g(t,x^*,u^*) + V_{x^*}(t,x^*)f(t,x^*,u^*) + V_t(t,x^*) &= 0.
\end{aligned}
$$

Integrating the above expressions from t_0 to T, we obtain

$$
\begin{aligned}
&\int_{t_0}^T g(s,x(s),u(s))ds + V(T,x(T)) - V(t_0,x_0) \le 0, \\
&\int_{t_0}^T g(s,x^*(s),u^*(s))ds + V(T,x^*(T)) - V(t_0,x_0) \le 0.
\end{aligned}
$$

Elimination of $V(t_0, x_0)$ yields

$$\int_{t_0}^{T} g(s, x(s), u(s))ds + q(x(T)) \leq \int_{t_0}^{T} g(s, x^*(s), u^*(s))ds + q(x^*(T)).$$

From which it readily follows that u^* is the optimal strategy.

Upon substituting the optimal strategy $\phi^*(t, x)$ into (2.2) yields the dynamics of optimal state trajectory as:

$$\dot{x}(s) = f[s, x(s), \phi^*(s, x(s))]ds, \quad x(t_0) = x_0. \tag{2.3}$$

Let $x^*(t)$ denote the solution to (2.3). The optimal trajectory $\{x^*(t)\}_{t=t_0}^{T}$ can be expressed as:

$$x^*(t) = x_0 + \int_{t_0}^{t} f[s, x^*(s), \phi^*(s, x^*(s))]ds. \tag{2.4}$$

For notational convenience, we use the terms $x^*(t)$ and x_t^* interchangeably. The value function $V(t, x)$ where $x = x_t^*$ can be expressed as

$$V(t, x) = \int_{t}^{T} g[s, x^*(s), \phi^*(s)]ds + q(x^*(T)).$$

2.1.2 Optimal Control

The maximum principle of optimal control was developed by Pontryagin (details in Pontryagin et al (1962)). Consider again the dynamic optimization problem (2.1)–(2.2).

Theorem 2.1.2 (Pontryagin's Maximum Principle) *A set of controls* $u^*(s) = \zeta^*(s, x_0)$ *provides an optimal solution to control problem (2.1)–(2.2), and* $\{x^*(s), t_0 \leq s \leq T\}$ *is the corresponding state trajectory, if there exist costate functions* $\Lambda(s) : [t_0, T] \to \mathbb{R}^m$ *such that the following relations are satisfied*:

$$\begin{aligned}
\zeta^*(s, x_0) &\equiv u^*(s) = \arg\max\{g[s, x^*(s), u(s)] + \Lambda(s)f[s, x^*(s), u(s)]\}, \\
\dot{x}^*(s) &= f[s, x^*(s), u^*(s)], \quad x^*(t_0) = x_0, \\
\dot{\Lambda}(s) &= -\frac{\partial}{\partial x}\{g[s, x^*(s), u^*(s)] + \Lambda(s)f[s, x^*(s), u^*(s)]\}, \\
\Lambda(T) &= \frac{\partial}{\partial x^*} q(x^*(T)).
\end{aligned}$$

Proof First define the function (Hamiltonian)

$$H(t,x,u) = g(t,s,u) + V_x(t,x)f(t,x,u).$$

From Theorem 2.1.2, we obtain

$$-V_t(t,x) = \max_u H(t,x,u).$$

This yields the first condition of Theorem 2.1.2. Using u^* to denote the payoff maximizing control, we obtain

$$H(t,x,u^*) + V_t(t,x){=}0.$$

Which is an identity in x. Differentiating this identity partially with respect to x yields

$$\begin{aligned} &V_{tx}(t,x) + g_x(t,x,u^*) + V_x(t,x)f_x(t,x,u^*) + V_{xx}(t,x)f(t,x,u^*) \\ &+ [g_u(t,s,u) + V_x(t,x)f_u(t,x,u^*)]\frac{\partial u^*}{\partial x} = 0. \end{aligned}$$

If u^* is an interior point, then $[g_u(t,x,u^*) + V_x(t,x)f_u(t,x,u^*)] = 0$ according to the condition $-V_t(t,x) = \max_u H(t,x,u)$. If u^* is not an interior point, then it can be shown that

$$[g_u(t,x,u^*) + V_x(t,x)f_u(t,x,u^*)]\frac{\partial u^*}{\partial x} = 0.$$

(because of optimality, $[g_u(t,x,u^*) + V_x(t,x)f_u(t,x,u^*)]$ and $\frac{\partial u^*}{\partial x}$ are orthogonal; and for specific problems we may have $\frac{\partial u^*}{\partial x} = 0$). Moreover, the expression $V_{tx}(t,x) + V_{xx}(t,x)f(t,x,u^*) \equiv V_{tx}(t,x) + V_{xx}(t,x)\dot{x}$ can be written as $[dV_x(t,x)](dt)^{-1}$. Hence, we obtain:

$$\frac{dV_x(t,x)}{dt} + g_x(t,x,u^*) + V_x(t,x)f_u(t,x,u^*) = 0.$$

By introducing the costate vector, $\Lambda(t) = V_{x^*}(t,x^*)$, where $x*$ denotes the state trajectory corresponding to u^*, we arrive at

$$\frac{dV_x(t,x^*)}{dt} = \dot{\Lambda}(s) = -\frac{\partial}{\partial x}\{g[s,x^*(s),u^*(s)] + \Lambda(s)f[s,x^*(s),u^*(s)]\}.$$

Finally, the boundary condition for $\Lambda(t)$ is determined from the terminal condition of optimal control in Theorem 2.1.2 as

$$\Lambda(T) = \frac{\partial V(T, x^*)}{\partial x} = \frac{\partial q(x^*)}{\partial x}.$$

Then,we obtain Theorem 2.1.2.

2.1.3 Stochastic Control

Consider the dynamic optimization problem in which the single decision maker

$$\max_u \ \mathbf{E}_{t_0}\left\{\int_{t_0}^{T} g[s, x(s), u(s)]ds + q(x(T))\right\}, \tag{2.5}$$

Subject to the vector-valued stochastic differential equation:

$$dx(s) = f[s, x(s), u(s)]ds + \sigma[s, x(s)]dw(s), \quad x(t_0) = x_0, \tag{2.6}$$

where $\mathbf{E}_{t_0}$ denotes the expectation operator performed at time t_0, and $\sigma[s, x(s)]$ is a $n \times \Theta$ matrix and $w(s)$ is a Θ dimensional Brownian motion and the initial state x_0 is given. Let $\Omega[s, x(s)] = \sigma[s, x(s)]\sigma[s, x(s)]'$ denote the covariance matrix with its element in row h and column ζ denoted by $\Omega^{h\zeta}[s, x(s)]$.

The technique of stochastic control developed by Fleming (1969) can be applied to solve the problem.

Theorem 2.1.5 *A set of controls $u^*(t) = \phi^*(t, x)$ constitutes an optimal solution to the problem (2.5)–(2.6), if there exist continuously differentiable functions $V(t, s)$ $[t_0, T] \times \mathbb{R}^n \to \mathbb{R}$, satisfying the following partial differential equation*:

$$- V_t(t,x) - \frac{1}{2}\sum_{h,\zeta=1}^{n} \Omega^{h\zeta}(t,x)V_{x^h x^\zeta}(t, s) = \max_u\{g\phi[t, x,u] + V_x f[t, x, u]\},$$
$$V(T, x) = q(x).$$

Proof Substitute the optimal control $\phi^*(t, x)$ into the (2.6) to obtain the optimal state dynamics as

$$\begin{aligned} dx(s) &= f[s, x(s), \phi^*(s, x(s))]ds + \sigma[s, x(s)]dw(s), \\ x(t_0) &= x_0. \end{aligned} \tag{2.7}$$

The solution to (2.7), denoted by $x^*(t)$, can be expressed as

$$x^*(t) = x_0 + \int_{t_0}^{t} f[s, x^*(t), \phi^*(s, x^*(t))]ds + \int_{t_0}^{T} \sigma[s, x^*(t)]dw(s). \tag{2.8}$$

We use X_t^* to denote the set of realizable values of x_t^* at time t generated by (2.8). The term x_t^* is used to denote an element in the set x_t^*.

Define the maximized payoff at time t with current state x_t^* as a value function in the form

$$V(t, x_t^*) = \max_u \mathbf{E}_{t_0}\left\{\int_t^T g(s, x(s), u(s))ds + q(x(T))|x(t) = x_t^*\right\}$$
$$= \mathbf{E}_{t_0} \int_t^T g[s, x^*(s), \phi^*(s, x^*(s))]ds + q(x^*(T)).$$

Satisfying the boundary condition

$$V(T, x^*(T)) = q(x^*(T)).$$

One can express $V(t, x_t^*)$ as

$$V(t, x_t^*) = \max_u \mathbf{E}_{t_0}\left\{\int_t^T g(s, x(s), u(s))ds + q(x(T))|x(t) = x_t^*\right\}$$
$$= \max_u \mathbf{E}_{t_0}\left\{\int_t^{t+\Delta t} g(s, x(s), u(s))ds + V(t + \Delta t, x_t^* + \Delta x_t^*)|x(t) = x_t^*\right\}. \tag{2.9}$$

where

$$\Delta x_t^* = f[t, x_t^*, \phi^*(t, x_t^*)]\Delta t + \sigma[t, x_t^*]\Delta z_t + o(\Delta t),$$
$$\Delta w_t = w(t + \Delta t) - w(t).$$

With $\Delta t \to 0$, applying Ito's lemma Eq. (2.9) can be expressed as:

$$\begin{aligned} V(t, x_t^*) = \max_u E_{t_0}\{&g[t, x_t^*, u]\Delta t + V(t, x_t^*) + V(t, x_t^*)\Delta t \\ &+ V_{x_t}(t, x_t^*) f[t, x_t^*, \phi^*(t, x_t^*)]\Delta t + V_{x_t}(t, x_t^*)\sigma[t, x_t^*]\Delta w \\ &+ \frac{1}{2}\sum_{h,\zeta=1}^{m} \Omega^{h\zeta}(t, x) V_{x^h x^\zeta}(t, x)\Delta t + o(\Delta t)\}. \end{aligned} \tag{2.10}$$

Dividing (2.10) throughout by Δt, with $\Delta t \to 0$, and taking expectation yields

$$
\begin{aligned}
& -V\left(t, x_t^*\right)-\frac{1}{2} \sum_{h, \zeta=1}^m \Omega^{h \zeta}(t, x) V_{x^h x^\zeta}(t, x)= \\
& =\max_u\left\{g\left[t, x_t^*, u\right]+V_{x_t}\left(t, x_t^*\right) f\left[t, x_t^*, \phi^*\left(t, x_t^*\right)\right] \Delta t+V_{x_t}\left(t, x_t^*\right)\right\} .
\end{aligned}
$$

With boundary condition

$$V\left(T, x^*(T)\right)=q\left(x^*(T)\right).$$

2.2 Differential Games and Their Solution Concepts

Firstly we introduce the definition of differential game briefly:

Definition 2.2.1 If the time difference between each phase of the game narrowed to the minimum limit, differential games can be considered as continuous-time dynamic games. A continuous-time infinite dynamic games of the initial state x_0 and continuous time $T - t_0$, and can be expressed as $\Gamma(x_0, T - t_0)$.

In particular, in the general n-person differential game, Player i seek to:

$$\max_{u_i} \int_{t_0}^T g^i\left[s, x(s), u_1(s), \cdots, u_n(s)\right] d s+q^i(x(T)). \tag{2.11}$$

For $i \in N = \{1, 2, \cdots, n\}$, where $g^i(\cdot) \geq 0$ and $q^i(\cdot) \geq 0$.
Subject to the deterministic dynamics

$$\dot{x}(s)=f\left[s, x(s), u_1(s), \ldots, u_n(s)\right], x\left(t_0\right)=x_0. \tag{2.12}$$

The functions $f[s, x(s), u_1(s), \ldots, u_n(s)]$, $g^i[s, x(s), u_1(s), u_2(s), \cdots, u_n(s)]$ and $q^i(\cdot)$, for $i \in N$, $s \in [t_0, T]$ are differentiable functions.

2.2.1 Open-Loop Nash Equilibria

If the players choose to commit their strategies from the outset, the players' information structure can be seen as an open-loop pattern in which $\eta^i(s) = \{x_0\}, s \in [t_0, T]$. Their strategies become functions of the initial state x_0 and time s, and can be expressed as $\{\mu_i(s) = \vartheta_i(s, x_0)\}$, for $i \in N$. An open-loop Nash equilibrium for the game is characterized as follows.

Theorem 2.2.1 *For the differential game (2.11) and (2.12), a set of strategies* $\{u_i^*(s) = \zeta_i^*(s, x_0), i \in N\}$ *provides an open-loop Nash equilibrium, an* $\{x^*(s),\ t_0 \le s \le T\}$ *is the corresponding state trajectory, if there exist n costate functions* $\Lambda^i(s) : [t_0, T] \to \mathbb{R}^n$, *for* $i \in N$, *such that the following relations are satisfied*:

$$\zeta_i^*(s, x_0) \equiv u_i^*(s) = \arg\max_{u_i \in \mathcal{U}^i} \{g^i[s, x^*(s), u_1^*(s), \cdots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s) \cdots, u_n^*(s)]$$
$$+ \Lambda^i(s) f[s, x^*(s), u_1^*(s), \cdots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s) \cdots, u_n^*(s)]\},$$
$$\dot{x}^*(s) = f[s, x^*(s), u_1^*(s), \cdots, u_n^*(s)], x^*(t_0) = x_0,$$
$$\dot{\Lambda}^i(s) = -\frac{\partial}{\partial x^*}\{g^i[s, x^*(s), u_1^*(s), \cdots, u_n^*(s)] + \Lambda^i(s) f[s, x^*(s), u_1^*(s), \cdots, u_n^*(s)]\}.$$

According to the analysis above, we know that:

First, given the optimal strategies of players, they should maximize the sum of the instantaneous payment and integration of state variation and covariate function in current time at every time point. That is, not only the instantaneous payment but also the whole payment influenced by state variation should be considered when one player chooses the optimal strategy. Second, the variation of optimal state depends on the optimal strategies of all the players, current time and state, and the optimal state of the beginning consistent with the initial state of the game. Third, given the optimal strategies of players $i \in N$ which only depend on current time and initial state, the variation of covariate functions depend on current instantaneous payment, variation of current state and current covariate functions. The value of covariate function equal to the marginal impact of optimal state at the end of game. Therefore, covariate functions of players reflect the impacts on future payment by the variation of optimal state.

2.2.2 Closed-Loop Nash Equilibria

After discussing the necessary conditions of open-loop Nash Equilibria, then we study the necessary conditions of closed-loop Nash Equilibria.

The players' information structures follow the pattern $\eta^i(s) = \{x_0, x(s)\}, s \in [t_0, T]$, for $i \in N$. The players' strategies become functions of the initial state x_0, current state $x(s)$ and current time s, and can be expressed as $\{u_i(s) = \vartheta_i(s, x(s), x_0), i \in N\}$. The following theorem provides a set of necessary conditions for any closed-loop no-memory Nash equilibrium solution to satisfy.

Theorem 2.2.2 *A set of strategies* $\{u_i(s) = \vartheta_i(s, x, x_0), i \in N\}$ *provides a closed-loop no memory Nash equilibrium solution to the game* (2.11)−(2.12), *and* $\{x^*(s), t_0 \le s \le T\}$ *is the corresponding state trajectory, if there exist n costate functions* $\Lambda^i(s) : [t_0, T] \to \mathbb{R}^n$, *for* $i \in N$, *such that the following relations are satisfied*:

$$\vartheta_i^*(s, x^*, x_0) \equiv u_i^*(s) = \arg\max_{u_i \in \mathcal{U}^i} \{g^i[s, x^*(s), u_1^*(s), \cdots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s) \cdots, u_n^*(s)]$$
$$+ \Lambda^i(s) f[s, x^*(s), u_1^*(s), \cdots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s) \cdots, u_n^*(s)]\},$$
$$\dot{x}^*(s) = f[s, x^*(s), u_1^*(s), \cdots, u_n^*(s)], x^*(t_0) = x_0,$$
$$\dot{\Lambda}^i(s) =$$
$$-\frac{\partial}{\partial x^*}\{g^i[s, x^*(s), \vartheta_1^*(s, x^*, x_0), \cdots, \vartheta_{i-1}^*(s, x^*, x_0),$$
$$u_i^*(s), \vartheta_{i+1}^*(s, x^*, x_0), \cdots \vartheta_n^*(s, x^*, x_0)] + \Lambda^i(s) f[s, x^*(s), \vartheta_1^*(s, x^*, x_0), \cdots,$$
$$\vartheta_{i-1}^*(s, x^*, x_0), u_i^*(s), \vartheta_{i+1}^*(s, x^*, x_0), \cdots \vartheta_n^*(s, x^*, x_0)]\},$$
$$\Lambda^i(T) = \frac{\partial}{\partial x^*} q^i(x^*(T)), i \in N.$$

Then a set of strategies $\{u_i(s) = \vartheta_i(s, x, x_0), i \in N\}$ *provides a closed-loop no memory Nash equilibrium.*

According to Theorem 2.2.2, similar to the open-loop situation, in closed-loop Nash equilibrium solution, we know that:

First, given the optimal strategies of players, they should maximize the sum of the instantaneous payment and integration of state variation and covariate function in current time at every time point. That is, not only the instantaneous payment but also the whole payment influenced by state variation should be considered when one player chooses the optimal strategy. Second, the variation of optimal state depends on the optimal strategies of all the players, current time and state, and the optimal state of the beginning consistent with the initial state of the game. Third, given the optimal strategies of players $i \in N$ which only depend on current time and initial state, the variation of covariate functions depend on current instantaneous payment, variation of current state and current covariate functions. The value of covariate function equal to the marginal impact of optimal state at the end of game. Therefore, covariate functions of players reflect the impacts on future payment by the variation of optimal state. Note that the partial derivatives of covariate function on optimal state depend on strategies of other players.

2.2.3 Feedback Nash Equilibria

The set of equations of closed-loop Nash Equilibria in general admits of an uncountable number of solutions, which correspond to "informationally non-unique" Nash equilibrium solutions of differential games under memoryless perfect state information pattern. Derivation of nonunique closed-loop Nash equilibria can be found in Mehlmann and Willing (1984). To eliminate information nonuniqueness in the derivation of Nash equilibria, one can constrain the Nash solution further by requiring it to satisfy the feedback Nash equilibrium property. In particular, the players' information structures follow either a closed-loop perfect state (CLPS) pattern in which $\eta^i(s) = \{x(t), t_0 \leq t \leq s\}$ or amemoryless perfect state (MPS) pattern in which $\eta^i(s) = \{x_0, x(s)\}$. Moreover, we require the following feedback Nash equilibrium condition to be satisfied.

Definition 2.3 For the n-person differential game (2.11)−(2.12), with MPS or CLPS information, an n-tuple of strategies $\{u_i^*(s) = \phi_i^*(s,x) \in \mathcal{U}^i, i \in N\}$ constitutes a feedback Nash equilibrium solution if there exist functionals $V^i(t,x), i \in N$ defined on $[t_0, T] \times \mathbb{R}^n$ and satisfying the following relations:

$$
\begin{aligned}
&V^i(t,x) = \int_t^T g^i[s, x^*(s), \phi_1^*(s,\eta_s), \cdots, \phi_n^*(s,\eta_s)]ds + q^i(x^*(T)) \geq \\
&\int_t^T g^i[s, x^{[i]}(s), \phi_1^*(s,\eta_s), \cdots, \phi_{i-1}^*(s,\eta_s), \phi_i(s,\eta_s), \phi_{i+1}^*(s,\eta_s), \cdots, \phi_n^*(s,\eta_s)]ds \\
&+ q^i(x^{[i]}(T)), \forall \phi_i(\cdot,\cdot) \in \mathcal{U}^i, x \in \mathbb{R}^n, \\
&V^i(T,x) = q^i(x)
\end{aligned}
$$

where on the interval $[t_0, T]$,

$$x^{[i]}(t) = x,$$

$$\dot{x}^*(s) = f[s, x^*(s), \phi_1^*(s,\eta_s), \cdots, \phi_n^*(s,\eta_s)], x(s) = x,$$

$\eta(s)$ stands for either the data set $\{x(s), x_0\}$ or $\{x(\tau), \tau \leq s\}$, depending on whether the information pattern is MPS or CLPS. Therefore the players' strategies can be expressed as $\{u_i^*(s) = \phi_i^*(s,x) \in U^i, i \in N\}$.

The following theorem provides a set of necessary conditions characterizing a feedback Nash equilibrium solution for the game (2.11)−(2.12) is characterized as follows:

Theorem 2.2.3 *An n-tuple of strategies* $\{u_i^*(t) = \phi_i^*(t,x) \in \mathcal{U}^i, i \in N\}$ *provides a feedback Nash equilibrium solution to the game* (2.11)−(2.12) *if there exist continuously differentiable functions* $V^i(t,x) : [t_0, T] \times \mathbb{R}^n \to \mathbb{R}$, $i \in N$, *satisfying the following set of partial differential equations*:

$$
\begin{aligned}
-V_t^i(t,x) &= \max_{u_i} \{ g^i[t, x, \phi_1^*(t,x), \ldots, \phi_{i-1}^*(t,x), u_i(t,x), \phi_{i+1}^*(t,x) \ldots, \varphi_n^*(t,x)] \\
&\quad + V_x^i(t,x) f[t, x, \phi_1^*(t,x), \ldots, \phi_{i-1}^*(t,x), u_i(t,x), \phi_{i+1}^*(t,x), \ldots, \phi_n^*(t,x)] \} \\
&= \{ g^i[t, x, \phi_1^*(t,x), \ldots, \varphi_n^*(t,x)] + V_x^i(t,x) f[t, x, \phi_1^*(t,x), \ldots, \phi_n^*(t,x)] \}, \\
V^i(T,x) &= q^i(x), i \in N.
\end{aligned}
$$

Theorem 2.2.4 *A pair of strategies* $\{\phi_i^*(t,x); i = 1,2\}$ *provides a feedback saddle-point solution to the zero-sum version of the game (2.11)−(2.12) if there exists a function* $V : [t_0, T] \times \mathbb{R}^n \to \mathbb{R}$ *satisfying the partial differential equation*:

$$
\begin{aligned}
-V_t(t,x) &= \min_{u_1\in\mathcal{U}^1}\max_{u_2\in\mathcal{U}^2}\{g[t,x,u_1(t),u_2(t)] + V_x f[t,x,u_1(t),u_2(t)]\} \\
&= \max_{u_2\in\mathcal{U}^2}\min_{u_1\in\mathcal{U}^1}\{g[t,x,u_1(t),u_2(t)] + V_x f[t,x,u_1(t),u_2(t)]\} \\
&= \{g[t,x,\phi_1^*(t,x),\phi_2^*(t,x)] + V_x f[t,x,\phi_1^*(t,x),\phi_2^*(t,x)]\}, \\
V(T,x) &= q(x).
\end{aligned}
$$

According to the necessary condition of feedback Nash equilibrium solution, there are two points should to note,

First, the value of the value functions of each player will change as time when they choose the optimal strategies under current time and state. Second, the payments of each player at the last time point are equal to that in the end of game.

2.3 Stochastic Differential Games and Their Solutions

We introduce the deterministic differential games and their solutions with stochastic factors.

2.3.1 The Model of Stochastic Differential Game

One way to incorporate stochastic elements in differential games is to introduce stochastic dynamics. A stochastic formulation for quantitative differential games of prescribed duration involves a vector-valued stochastic differential equation

$$
\begin{aligned}
&dx(s) = f[s,x(s),u_1(s),u_2(s),\ldots,u_n(s)]ds + \sigma[s,x(s)]dw(s), \\
&x(t_0) = x_0.
\end{aligned}
\tag{2.13}
$$

which describes the evolution of the state and N objective functionals

$$
\mathbf{E}_{t_0}\left\{\int_{t_0}^{T} g^i[s,x(s),u_1(s),u_2(s),\ldots,u_n(s)]ds + q^i(x(T))\right\}, i\in N \tag{2.14}
$$

with $\mathbf{E}_{t_0}\{\cdot\}$ denoting the expectation operation taken at time t_0, $\sigma[s,x(s)]$ is a $n\times\Theta$ matrix and $w(s)$ is a Θ dimensional Brownian motion and the initial state x_0 is given. Let $\Omega[s,x(s)] = \sigma[s,x(s)]\sigma[s,x(s)]'$ denote the covariance matrix with its element in row h and column ζ denoted by $\Omega^{h\zeta}[s,x(s)]$. Moreover, $\mathbf{E}[dw_\varpi] = 0$, $\mathbf{E}[dw_\varpi dt] = 0$, and $\mathbf{E}\left[(dw_\varpi)^2\right] = dt$, for $\varpi\in[1,2,\ldots,\Theta]$; $\mathbf{E}[dw_\varpi dw_\omega] = 0$, for $\varpi\in[1,2,\ldots,\Theta]$, $\omega\in[1,2,\ldots,\Theta]$ and $\varpi\neq\omega$. Given the stochastic nature, the information structures must follow the MPS pattern or CLPS pattern or the feedback perfect state (FB) pattern in which $\eta^i(s) = \{x(s)\}$, $s\in[t_0,T]$.

2.3.2 The Solutions of Stochastic Differential Game

The character of stochastic differential game is the state changes with the stochastic dynamic system in every moment. Therefore, stochastic differential game is closer to reality compared with the deterministic differential game. Based on this, the following section only discuss the feedback solutions which are more realistic then the open-loop solution. A Nash equilibrium of the stochastic game (2.13)−(2.14) can be characterized as:

Theorem 2.3.1 *An n-tuple of feedback strategies* $\{\phi_i^*(t,x) \in \mathcal{U}^i; i \in N\}$ *provides a Nash equilibrium solution to the game (2.13)−(2.14) if there exist suitably smooth functions* $V^i : [t_0, T] \times \mathbb{R}^n \to \mathbb{R}$, *satisfying the semilinear parabolic partial differential equations*

$$\begin{aligned} -V_t^i - \frac{1}{2}\sum_{h,\zeta}^{n} \Omega^{h\zeta}(t,x) V^i_{x_h x_\zeta} &= \max_{u_i}\{g^i[t,x,\phi_1^*(t,x),\ldots,\phi_{i-1}^*(t,x),u_i(t),\phi_{i+1}^*(t,x),\ldots,\phi_n^*(t,x)] \\ &\quad + V_x^i(t,x) f[t,x,\phi_1^*(t,x),\ldots,\phi_{i-1}^*(t,x),u_i(t),\phi_{i+1}^*(t,x),\ldots,\phi_n^*(t,x)]\} \\ &= \{g^i[t,x,\phi_1^*(t,x),\phi_2^*(t,x),\ldots,\phi_n^*(t,x) + V_x^i(t,x) f[t,x,\phi_1^*(t,x),\ldots,\phi_n^*(t,x)]\}, \\ V^i(T,x) &= q^i(x), i \in N. \end{aligned}$$

Proof This result follows readily from the definition of Nash equilibrium and from Theorem 2.1.2, since by fixing all players' strategies, except the ith one's, at their equilibrium choices (which are known to be feedback by hypothesis), we arrive at a stochastic optimal control problem of the type covered by Theorem 2.3.1 and whose optimal solution (if it exists) is a feedback strategy.

Consider the two-person zero-sum version of the game (2.13)–(2.14) in which the payoff of Player 1 is the negative of that of Player 2. Under either MPS or CLPS information pattern, a Nash equilibrium solution can be characterized as follows.

Theorem 2.3.2 *A pair of strategies* $\{\phi_i^*(t,x) \in \mathcal{U}^i; i = 1,2\}$ *provides a feedback saddle-point solution to the two-person zero-sum version of the game (2.13)–(2.14) if there exists a function* $\Lambda(s) : [t_0, T] \to \mathbb{R}^n$ *satisfying the partial differential equation*:

$$\begin{aligned} -V_t - \frac{1}{2}\sum_{h,\zeta}^{n} \Omega^{h\zeta}(t,x) V_{x_h x_\zeta} &= \min_{u_1 \in \mathcal{U}^1} \max_{u_2 \in \mathcal{U}^2} \{g[t,x,u_1,u_2] + V_x f[t,x,u_1,u_2]\} \\ &= \max_{u_2 \in \mathcal{U}^2} \min_{u_1 \in \mathcal{U}^1} \{g[t,x,u_1,u_2] + V_x f[t,x,u_1,u_2]\} \\ &= \{g[t,x,\phi_1^*(t,x),\phi_2^*(t,x)] + V_x f[t,x,\phi_1^*(t,x),\phi_2^*(t,x)]\}, \\ V(T,x) &= q(x). \end{aligned}$$

Proof This result follows as a special case of Theorem 2.3.1 by taking $n = 2$, $g^1(\cdot) = -g^2(\cdot) \equiv g(\cdot)$, and $q^1(\cdot) = -q^2(\cdot) \equiv q(\cdot)$, in which case $V^1 = -V^2 \equiv V$

and existence of a saddle point is equivalent to interchangeability of the min max operations.

According to the necessary condition of feedback Nash Equilibria, there are two points we should to know,

First, the value functions in stochastic differential game (2.13)–(2.14) change with time when all the players (include i) determine the optimal strategies depend on current time and state. Second, the value function of player $i \in N$ in last point equals to his final payment in the game.

http://www.springer.com/978-3-319-40586-5

Non-cooperative Stochastic Differential Game Theory of Generalized Markov Jump Linear Systems
Zhang, C.-k.; Zhou, H.-y.; Zhu, H.-n.; Bin, N.
2017, XV, 187 p. 6 illus., Hardcover
ISBN: 978-3-319-40586-5