Time-Consistent Distribution Mechanism in the Game of Environmental Management

Shuqing Liu¹, Hongwei Gao¹,², Lei Wang¹, Mengke Zhen¹ and Haiqiang Ji¹
¹ School of Mathematics and Statistics, Qingdao University, Qingdao 266071, China; ² Institute of Applied Mathematics of Shandong, Qingdao 266071, China; *Corresponding author

Abstract—For the differential game about the pollution emissions control problem among countries, we apply Shapley value to allocate the total payoff of all countries. The important contribution of this paper is the establishment of a time-consistent imputation distribution procedure so that the initial agreement remains valid for the duration of the whole game.

Keywords-differential games; environmental pollution; shapley value; time-consistency; imputation distribution procedure

I. INTRODUCTION

Global environmental problems such as global warming, ozone layer destruction and marine pollution have brought a series of challenges to human beings, including the frequent occurrence of extreme weather, rising sea levels, etc. Although there are many agreements related to environmental protection, global environmental governance is still unsatisfactory. One of the important reasons is that the implementation of global environmental governance agreement is a long process. During the long course, the cooperation between countries is difficult to maintain consistent throughout. To solve this problem, we establish the linear state differential game model on pollution control and construct the imputation distribution procedure to ensure that the initial allocation agreement is effective on each period of the game, so as to ensure all countries wouldn’t deviate from the initial agreement.

Differential games provide an effective tool to study pollution control problems and analyze the interactions between the participants’ strategic behaviors. Applications of noncooperative differential games in environmental control are presented in [1-3]. Cooperative differential games in environmental control can be found in [2, 4-8]. We adopt the cooperative differential game to study the pollution control problem. Each player in the model represents a country and national industrial production would discharge pollution. Players choose emission strategies to maximize their payoffs. The state variable is the current pollutant stock. This model work has fundamentally solved the problem of how to ensure long-term cooperation between countries.

The time-consistent solution for differential games was initially proposed in [9]. In 1982 [10] introduced the concept of imputation distribution procedure for cooperative solution which becoming the most breakthrough result in this field. In recent years, the research on the time-consistency can refer to literatures [5, 8, 11].

II. THE MODEL

Consider an $n$-player differential game played on the time interval $[0, \infty)$. Let $I=\{1,...,n\}$ denote the set of players, where each player represents a country.

National industrial production would discharge pollution. Emission of player $i$ at time $t \in [0, \infty)$ is denoted by $E_i(t)$. Let $S(t)$ denote the stock of accumulated pollution by time $t$. The evolution of this stock is governed by the following differential equation:

$$\dot{S}(t) = \sum_{i=1}^{n} E_i(t) - \delta S(t), \quad S(0) = S_0$$

where $\delta > 0$ denotes the natural absorption rate of pollution and $S_0$ denotes the initial stock of pollution.

We assume that the emissions are proportional to the production and hence the revenue from production can be expressed as a function of emissions. From now we omit on the time argument when no ambiguity may arise. Let $R_i(E_i)$ denote the concave increasing revenue function and let $D_1(S)$ be the liner damage cost. These functions are specified as follows:

$$R_i(E_i) = \ln E_i, \quad D_1(S) = \varphi S,$$

where $\varphi > 0$ represents the unit damage cost of player $i$. Player $i$’s payoff is given by

$$W_i = \int_0^\infty (\ln E_i - \varphi S)e^{-\rho t} dt,$$  (2)
where \( \rho \) is a positive discount rate.

By (1)-(2), we have defined a linear-state differential game on pollution control in which the state equation and the payoff functions are polynomials of degree 1 with respect to the state variables.

### III. The Solution of the Model

We adopt a cooperative differential game methodology to deal with the problem of payoff allocation problems associated with emissions of the players. The definition and the computation of a time-consistent distribution scheme are dealt with below after introducing some notations.

Let the state of the game be defined by the pair \((S, t)\) and denote by \(\Gamma(S, t)\) the subgame starting at date \(t\) with stock of pollution \(S\). Denote by \(S'(t)\) the trajectory of accumulated pollution under full cooperation. The characteristic function value for a coalition \(K \subseteq N\) in subgame \(\Gamma(S, t)\) is defined to be its maximal payoff while left-out players’ strategies are determined and is denoted by \(\nu(K, S, t)\). This approach is different from the traditional definition that coalition \(K\) and coalition \(N \setminus K\) played two players zero-sum game. Indeed, in dealing with global environmental issues, there is almost no mutual confrontation between two coalitions. From this perspective, our approach is more in line with reality, but the characteristic function may not meet the super-additivity.

Let \(\phi(v, S, t) = (\phi_i(v, S, t), \ldots, \phi_n(v, S, t))\) denote the Shapley value in subgame \(\Gamma(S, t)\). Denote by \(\beta_i(t)\) the immediate payoff allocated to player \(i\) at instant of time \(t\) and let \(\beta(t) = (\beta_1(t), \ldots, \beta_n(t))\).

**Definition 1** The vector \(\beta(t) = (\beta_1(t), \ldots, \beta_n(t))\) is an imputation distribution procedure (IDP) if

\[
\phi_i(v, S, 0) = \int_0^\infty e^{-\rho t} \beta_i(t) dt, i = 1, \ldots, n. \tag{3}
\]

Definition 1 shows that imputation distribution procedure consists of payment of players at each stage of the game such that the sum of discounted instantaneous payment to a player \(\int_0^\infty e^{-\rho t} \beta_i(t) dt\) is equal to his payoff \(\phi_i(v, S, 0)\) according to the cooperative solution concept on which he has agreed before the game starts.

**Definition 2** The vector \(\beta(t) = (\beta_1(t), \ldots, \beta_n(t))\) is a time-consistent IDP if at \((S', t)\), \(\forall t \in [0, \infty)\), the following condition holds

\[
\phi_i(v, S, 0) = \int_0^\infty e^{-\rho t} \beta_i(t) dt + e^{-\rho t} \phi_i(v, S', t), i = 1, \ldots, n. \tag{4}
\]

Definition 2 shows that along the cooperation state trajectory, supposed player \(i, \forall i \in I\) wants to renegotiate the agreement reached in game \(\Gamma(S, 0)\) in any state \((S', t)\), he has received payoff \(\int_0^\infty e^{-\rho t} \beta_i(t) dt\) at this moment. If all players still maintain cooperation in subgame \(\Gamma(S', t)\), player \(i\) would get \(e^{-\rho t} \phi_i(v, S', t)\) in this subgame. If player \(i\)’s payoff \(\int_0^\infty e^{-\rho t} \beta_i(t) dt\) determined by imputation distribution procedure up to this moment together with his remaining payoff \(e^{-\rho t} \phi_i(v, S', t)\) from this moment till the end will coincide with the payoff \(\phi_i(v, S, 0)\) on which he agreed at the beginning of the game according to the cooperative solution concept, then this renegotiation would leave the original agreement unaltered. If one can find an IDP such that (4) holds true, then this IDP is time-consistent. The following is the establishment of an algorithm for computing the time-consistent IDP.

#### A. An Algorithm of Time-Consistent IDP

**Step 1** Maximizing the total payoff of the grand coalition.

To maximize the total payoff of all players, we need to solve a standard dynamic programming problem, that is:

\[
\max \sum_{i \in I} W_i = \sum_{i \in I} \int_t^\infty (\ln E_i - \phi_i(S) e^{-\rho (t+s)} d\tau
\]

s.t. \(E_i = \delta S, S(t) = S'(t)\).

Denote by \(V(I, S, t)\) the value function of this problem. The outcome of this optimization is a vector of emission strategy \(E'(S') = (E'_1(S'), \ldots, E'_n(S'))\), where \(S'\) refers to the accumulated pollution under the scenario of full cooperation. Assuming the value function \(V(I, S, t)\) is differentiable, then the Hamilton-Jacobi-Bellman (HJB) equations is

\[
\rho V(I, S, t) = \max \left\{ \sum_{i \in I} (\ln E_i - \phi_i(S) + V'(I, S, t) \left[ \sum_{i \in I} E_i - \delta S \right] \right\}
\]

Differentiating (5) with respect to \(E_i\) and equating to zero, then we get that

\[
E_i' = -\frac{1}{V'(I, S, t)}.
\]

From the structure of objective function and state equation in dynamic programming problem, we can see that a linear form for \(V(I, S, t)\) satisfies (5). Therefore we obtain
\begin{equation}
V(I,S,t) = -\sum_{i=1}^{n} \phi_i \frac{1}{\rho + \delta} S + \frac{n}{\rho} (\ln \rho + \delta - 1), \tag{6}
\end{equation}

\begin{equation}
E_i^t = \frac{\rho + \delta}{\sum_{i} \phi_i} \forall i \in I. \tag{7}
\end{equation}

According to (7) we know that all players have the same emission level \( \frac{\rho + \delta}{\sum_{i} \phi_i} \), which is a positive constant.

To get accumulated stock of pollution \( S^i(t) \) at time \( t \), we put \( E_i^t \) into (1) and solve to obtain

\begin{equation}
S^i(t) = S_0 e^{-\delta t} + \left( \frac{\rho + \delta}{\sum_{i} \phi_i} \right) n (1 - e^{-\delta t}), \tag{8}
\end{equation}

**Step 2** Computation of feedback Nash equilibrium

Denote by \( V_i(S,t) \) the player \( i \)'s payoff under feedback Nash equilibrium and denote by \( S^N \) the accumulated pollution under feedback Nash equilibrium. The feedback Nash equilibrium is \( E_i^N(S) = (E_i^N(S),...,E_n^N(S)) \), where \( E_i^N(S), i=1,...,n \) refers to player \( i \)'s feedback Nash equilibrium strategy. To get the feedback Nash equilibrium, assuming the value function \( V_i(S,t) \) is differentiable, then the HJB equations is

\begin{equation}
\rho V_i(S,t) = \max \{ \ln E_i - \phi_i S + V_i(S,t) \left( \sum_{i} E_i - \delta S \right) \}. \tag{9}
\end{equation}

By a similar solving procedure for \( E_i^N, V(I,S,t) \), we get that

\begin{equation}
E_i^N = \frac{\rho + \delta}{\phi_i} \forall i \in I, \tag{10}
\end{equation}

\begin{equation}
V_i(S,t) = -\frac{\phi_i}{\rho + \delta} S + \frac{1}{\rho} \left( \ln \left( \frac{\rho + \delta}{\phi_i} \right) - \phi_i \sum_{j \neq k} \phi_j \right) \forall i \in I. \tag{11}
\end{equation}

By (10) player \( i \)'s feedback Nash equilibrium strategy is a constant too. The difference between cooperative and noncooperative emission rates is that, in the former, each player takes into account the sum of the cost coefficient, whereas in the latter, each player determines her emissions taking into account only her own cost coefficient. So we conclude that each player's cooperative emission level lower than that under noncooperation, i.e. \( E_i^N < E_i^N, \forall i \in I \).

To get accumulated stock of pollution \( S^N(t) \) at time \( t \), we put \( E_i^N \) into (1) and choose the initial condition \( S^N(t) = S^i(t), \tau \in [0,t] \). Then we obtain

\begin{equation}
S^N(t) = S^i(t) e^{-(\rho + \delta) t} + \frac{\rho + \delta}{\sum_{i} \phi_i} (1 - e^{-(\rho + \delta) t}). \tag{12}
\end{equation}

**Step 3** Computation of optimal payoff for intermediate coalitions

Assuming player in coalition \( K \subseteq I \) chooses strategy to maximize the total payoff of the coalition \( K \) and the left-out players choose feedback Nash equilibrium strategy \( E_j^N(S), j \in I \setminus K \). Denote by \( E_i^N(S), i \in K \) the player \( i \)'s strategy in coalition \( K \) and denote by \( S^N \) the accumulated pollution under this scenario. The value function is obtained by solving the following dynamic programming problem,

\begin{equation}
\max_{S^N} \sum_{i \in K} W_i = \sum_{i \in K} \int_{0}^{t} (\ln E_i - \phi_i S) e^{-(\rho + \delta) \tau} d \tau
\end{equation}

s.t. \( S^N = \sum_{j \in K} E_j - \delta S, S(t) = S^i(t), E_j = E_j^N, j \in I \setminus K \).

Assuming value function \( V_i(S,t) \) is differentiability with respect to \( S \), and then the HJB equations is

\begin{equation}
\rho V(K,S,t) = \max \{ \sum_{i \in K} (\ln E_i - \phi_i S) + V(K,S,t) \left( \sum_{i \in K} E_i - \delta S \right) \}.
\end{equation}

By a similar solving procedure for \( E_i^N, V(I,S,t) \), we get that

\begin{equation}
E_i^N = \frac{\rho + \delta}{\phi_i} \forall i \in K, \tag{13}
\end{equation}

\begin{equation}
V(K,S,t) = -\frac{\phi_i}{\rho + \delta} S + \frac{1}{\rho} \left( \ln \left( \frac{\rho + \delta}{\phi_i} \right) - \phi_i \sum_{j \neq k} \phi_j \right) \forall i \in K. \tag{14}
\end{equation}
By (11) we know that each player in coalition $K$ have the same emission level $\sum_{j \in K} \rho_j$, thus the emission level of coalition $K$ is $\sum_{j \in K} k(\rho + \delta)$. However, the sum of the noncooperative emission levels of members in coalition $K$ equal to $\sum_{j \in K} \rho_j^\delta$. Thus we can conclude that any coalition emission level is lower than the sum of noncooperative emission levels of the coalition members.

To get accumulated stock of pollution $S^K(t)$ at time $t$, we put $E_s^i, E_s^j, i \in K, j \in I \setminus K$ into (1) and choose the initial condition $(S^K(\tau) = 0, \tau \in [0, t]$. Then we obtain

$$S^K(t) = S'(\tau) e^{\rho t} + \sum_{j \in K} \frac{1}{\delta} (1 - e^{\rho (t - \tau)}).$$

(13)

**Step 4** Definition of the characteristic function

The characteristic function of the cooperative differential game $(S(t), \Gamma)$ is defined as follows:

$$\phi(K) = \phi_1(K, S(t)) + \sum_{i \in K} \frac{1}{\delta} (1 - e^{\rho (t - \tau)}).$$

(14)

Step 5 Computation of the Shapley value.

The component $i$ of the Shapley value $\phi(t, S(t)) = (\phi_1(t, S(t)), \ldots, \phi_n(t, S(t)))$ in subgame $(S(t), \Gamma)$ is given by

$$\phi_i(t, S(t)) = \sum_{(K, n, k \in I)} \frac{(n-k)(k-1)!}{n!} [v(K, S(t)) - v(K \setminus \{i\}, S(t))].$$

(15)

Step 6 Definition of the time-consistent IDP.

At instant of time $t, t \in [0, \infty)$, allocate to player $i$ the following amount:

$$\beta_i(t) = \phi_i(t, S(t)) - \frac{d}{dt} \phi_i(t, S(t)).$$

(16)

The following theorem shows that $\beta(t) = (\beta_1(t), \ldots, \beta_n(t))$ given by (15) is a time-consistent IDP.

**Theorem 1** The vector $\beta(t) = (\beta_1(t), \ldots, \beta_n(t))$ where $\beta_i(t), i = 1, \ldots, n$ is given by (15) is a time-consistent IDP.

**Proof.** We first show that the vector $\beta(t) = (\beta_1(t), \ldots, \beta_n(t))$ given by (15) is an IDP, i.e.

$$\int_0^\infty e^{-\rho t} \beta_i(t) dt = \phi_i(t, S(t)).$$

Multiply (15) by the discount factor $e^{-\rho t}$ and integrate

$$\int_0^\infty e^{-\rho t} \beta_i(t) dt = \phi_i(t, S(t)) = \phi_i(t, S(t)).$$

Next we show that $\beta(t) = (\beta_1(t), \ldots, \beta_n(t))$ is a time-consistent IDP. We have

$$\int_0^\infty e^{-\rho t} \beta_i(t) dt = \phi_i(t, S(t)) = \phi_i(t, S(t)).$$

By definition 2, the vector $\beta(t) = (\beta_1(t), \ldots, \beta_n(t))$ given by (15) is a time-consistent IDP.

**IV. AN EXAMPLE OF THE MODEL**

Given the game consisting of three players, i.e. $I = \{1, 2, 3\}$, according to the characteristic function value constructed by step 4 of the algorithm, we obtain that

$$\phi_1(t, S(t)) = \frac{1}{\rho + 3\rho} \left[ \phi_1(t, S(t)) - \phi_2(t, S(t)) - \phi_3(t, S(t)) \right] + \frac{3}{\rho} \phi_2(t, S(t)) - \phi_3(t, S(t)) - 6, \quad i = 1, 2, 3$$

(17)

According to (15), we have deduced

$$\beta_1(t) = -\phi_1 e^{-\rho t} - \frac{3}{\rho + \delta} \sum_{j \in I} \phi_j,$$

$$+ \frac{1}{6} \left[ \phi_2(t, S(t)) + \phi_3(t, S(t)) \right] + \frac{3}{\rho} \phi_2(t, S(t)) - \phi_3(t, S(t)) - 6, \quad i = 1, 2, 3.$$
By theorem 1 \( \beta(t) = (\beta_1(t), \beta_2(t), \beta_3(t)) \) determined by (17) is a time-consistent IDP.

Note that the characteristic function defined in this paper may not satisfy the super-additivity, therefore, in the following, theorem 2 gives the sufficient condition of characteristic function in this example satisfying the property of super-additivity.

**Theorem 2** For three players \( i, j, h \in \{1, 2, 3\} \) who are different from each other, if

\[
\frac{\ln[(\phi_i + \phi_j)^2]}{\phi_i \phi_j} \leq 2 \quad \text{and} \quad \frac{\ln[(\phi_i + \phi_j + \phi_k)^3]}{(\phi_i + \phi_j) \phi_k} \leq 4
\]

then the characteristic function defined by (14) is super-additive, i.e.

\[
v(K, S, t) + v(L, S, t) \leq v(K \cup L, S, t).
\]

\( \forall K, L \subseteq I, K \cap L = \emptyset, I = \{1, 2, 3\} \)

**Proof.** By the definition of the characteristic function and \( S^K, S^L, S^{K \cup L} \) determined by (13), we have

\[
v(K, S, t) + v(L, S, t) - v(K \cup L, S, t)
\]

\[
= \frac{1}{\rho} \left[ \ln \left( \frac{\sum_{i \in K} \phi_i}{\sum_{j \in L} \phi_j} \right) \right] - \sum_{i \in K} \frac{1}{\phi_i} - \sum_{j \in L} \frac{1}{\phi_j} - \frac{1}{\phi_k} \left( \phi_k \right)
\]

\[
- \frac{1}{\delta} \left( \sum_{i \in K} \frac{1}{\phi_i} - \sum_{j \in L} \frac{1}{\phi_j} - \sum_{k \in K \cup L} \frac{1}{\phi_k} \right),
\]

where \( k = |K|, l = |L| \). Since \( \delta > 0, t \geq \tau, \phi_i > 0, \forall i \in I \), then

\[
- \frac{1}{\delta} \left( \sum_{i \in K} \frac{1}{\phi_i} + \sum_{j \in L} \frac{1}{\phi_j} + \sum_{k \in K \cup L} \frac{1}{\phi_k} \right) \leq 0.
\]

Next we prove that

\[
\frac{1}{\rho} \left[ \ln \left( \frac{\sum_{i \in K} \phi_i}{\sum_{j \in L} \phi_j} \right) \right] - \sum_{i \in K} \frac{1}{\phi_i} - \sum_{j \in L} \frac{1}{\phi_j} + \sum_{k \in K \cup L} \frac{1}{\phi_k} \leq 0.
\]

Case 1 \( k = l = 1 \). Assuming that \( K = \{\alpha\}, L = \{\beta\}, \forall \alpha, \beta \in \{1, 2, 3\}, \alpha \neq \beta \), then

\[
\frac{1}{\rho} \left[ \ln \left( \frac{\sum_{i \in K} \phi_i}{\sum_{j \in L} \phi_j} \right) \right] - \sum_{i \in K} \frac{1}{\phi_i} - \sum_{j \in L} \frac{1}{\phi_j} + \sum_{k \in K \cup L} \frac{1}{\phi_k} \leq 0
\]

where \( k = |K| = 1, l = |L| = 1 \), \( k = l = 1 \)

\[
\frac{1}{\rho} \left[ \ln \left( \frac{\sum_{i \in K} \phi_i}{\sum_{j \in L} \phi_j} \right) \right] - \sum_{i \in K} \frac{1}{\phi_i} - \sum_{j \in L} \frac{1}{\phi_j} + \sum_{k \in K \cup L} \frac{1}{\phi_k} \leq 0
\]

where \( k = l = 1 \)

Integrating Case 1 and Case 2 we have completed the proof of theorem 2.
V. CONCLUSION

We establish an imputation distribution procedure to ensure each country wouldn’t deviate from the initial agreement. At the same time, we prove that emission levels of any coalition are lower than the sum of noncooperative emission level of the coalition members. Thus we obtained that pollutant accumulation under cooperation is lower than that under noncooperation. In this paper, we choose a simple model of ecological economy; a topic for future research would be to use a more realistic representation. In addition, other types of differential games, such as linear quadratic differential game, whether the corresponding algorithm can be obtained is a further research project.

ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China (Grant Nos. 71171120, 71571108), Projects of International (Regional) Cooperation and Exchanges of National Natural Science Foundation of China (Grant No. 71411130215), Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20133706110002), Natural Science Foundation of Shandong Province of China (Grant No. ZR2015GZ007).

REFERENCES

[1] D. Yeung, “A differential game of industrial pollution management,” Ann. Oper. Res. 37: 297–311, 1992.
[2] E. Dockner and N. V. Long, “International pollution control: cooperative versus noncooperative strategies,” Environ. Econ. Manag. 25: 13–29, 1993.
[3] M. Stimming, “Capital accumulation subject to pollution control: open-loop versus feedback investment strategies,” Ann. Oper. Res. 88: 309–336, 1999.
[4] G. Martín-Herrán and G. Zaccour, “Credibility of incentive equilibrium strategies in linear-state differential games,” J. Optim. Theory Appl. 126(2): 367–389, 2005.
[5] L. A. Petrosyan and G. Zaccour, “Time-consistent Shapley value allocation of pollution cost reduction,” J. Econ. Dyn. Control. 27(3): 381–398, 2003.
[6] S. Jørgensen, G. Martín-Herrán and G. Zaccour, “Agreement and time consistency in linear-state differential games,” J. Optim. Theory Appl. 119(1): 49–63, 2003.
[7] N. Kozlovskaya, L. A. Petrosyan and N. Zenkevich, “Coalitional solution of a game-theoretic emission reduction model,” Int. Game Theory Rev. 12(3): 275-286, 2010.
[8] D. Yeung, “Dynamically consistent cooperative solution in a differential game of transboundary industrial pollution,” J. Optim. Theory Appl. 134: 143–160, 2007.
[9] L. A. Petrosyan, “Stable solutions of differential games with many participants,” Viestnik of Leningrad University, 19: 46–52, 1977.
[10] L. A. Petrosyan and N. N. Danilov, “Stability of solutions in non-zero sum differential games with transferable payoffs,” Viestnik of Leningrad University, 1: 52–59, 1979.
[11] L. Wang, H. Gao, L. A. Petrosyan, H. Qiao and A. Sedakov, “Strategically supported cooperation in dynamic games with coalition structures,” Science China-Mathematics, 59(5): 1015-1028, 2016.