Lecture Notes on Non-Cooperative Game Theory

Tamer Başar *

July 26, 2010

These lecture notes have been prepared as a supplement to the series of 20 lectures to be delivered by the author for the Game Theory Module of the Graduate Program in Network Mathematics at the Hamilton Institute and CTVR in Trinity College, Dublin, Ireland, August 3-6, 2010. They are intended to accompany the actual lectures to be delivered, and should not be viewed as a comprehensive, exhaustive coverage of game theory.

Throughout the lecture notes, the following conventions and acronyms will be adopted and used:

ZSG: Zero-sum game

NZSG: Nonzero-sum game   NZSDG: Nonzero-sum dynamic (differential) game

NE: Nash equilibrium (or equilibria)

SP: Saddle point

SPE: Saddle point equilibrium (or equilibria)

SES: Stackelberg equilibrium solution (strategy)

MSNE: Mixed-strategy Nash equilibrium

MSSPE: Mixed-strategy saddle-point equilibrium

BSNE: Behavioral-strategy Nash equilibrium

BSSPE: Behavioral-strategy saddle-point equilibrium

CE: Correlated equilibrium

OL: Open loop   CL: Closed loop   CLFB: Closed-loop feedback

$\mathbf{P}i$: Player i, where $i \in \mathcal{N}$, with $\mathcal{N} := \{1, \ldots, N\}$ denoting the players set

*Swanlund Endowed Chair, Electrical and Computer Engineering and Center for Advanced Study Professor, University of Illinois at Urbana Champaign. *Mailing address:* Coordinated Science Laboratory, University of Illinois, 1308 West Main Street, Urbana, IL 61801-2307, USA. Tel: (+1) 217 333-3607; Fax: (+1) 217 244-1653; Email: basar1@illinois.edu.

Contents

1 Lecture 1: A general introduction to game theory, its origins, and classifications

1.1 What is game theory?

Game theory deals with strategic interactions among multiple decision makers, called *players* (and in some context *agents*), with each player's preference ordering among multiple alternatives captured in an objective function for that player, which she either tries to maximize (in which case the objective function is a *utility* function or *benefit* function) or minimize (in which case we refer to the objective function as a *cost* function or a *loss* function). For a non-trivial game, the objective function of a player depends on the choices (*actions*, or equivalently *decision variable*) of at least one other player, and generally of all the players, and hence a player cannot simply optimize her own objective function independent of the choices of the other players. This thus brings in a coupling between the actions of the players, and binds them together in decision making even in a non-cooperative environment. If the players were able to enter into a cooperative agreement so that the selection of actions or decisions is done collectively and with full trust, so that all players would benefit to the extent possible, and no inefficiency would arise, then we would be in the realm of *cooperative game theory*, with issues of bargaining, coalition formation, excess utility distribution, etc. of relevance there; cooperative game theory will not be covered in this overview; see for example *Owen (1995)*[2], *Vorob'ev (1977)*[4], or *Fudenberg and Tirole (1991)*[3]. See also the recent survey article [5], which emphasizes applications of cooperative game theory to communication systems.

If no cooperation is allowed among the players, then we are in the realm of *non-cooperative game theory*, where first one has to introduce a satisfactory solution concept. Leaving aside for the moment the issue of how the players can reach a satisfactory solution point, let us address the issue of if the players are at such a solution point, what are the minimum features one would expect to see there. To first order, such a solution point should have the property that if all players but one stay put, then the player who has the option of moving away from the solution point should not have any incentive to do so because she cannot improve her payoff. Note that we cannot allow two or more players to move collectively from the solution point, because such a collective move requires cooperation, which is not allowed in a non-cooperative game. Such a solution point where none of the players can improve her payoff by a unilateral move is known as a *non-cooperative* *equilibrium* or *Nash equilibrium*, named after John Nash, who introduced it and proved that it exists in finite games (that is games where each player has only a finite number of alternatives), some sixty years ago; see *Nash (1950, 1951)* [16, 17]. We will discuss this result in detail later in these Notes, following some terminology, a classification of non-cooperative games according to various attributes, and a mathematical formulation.

We say that a non-cooperative game is *nonzero-sum* if the sum of the players' objective functions cannot be made zero after appropriate positive scaling and/or translation that do not depend on the players' decision variables. We say that a two-player game is *zero-sum* if the sum of the objective functions of the two players is *zero* or can be made zero by appropriate positive scaling and/or translation that do not depend on the decision variables of the players. If the two players' objective functions add up to a constant (without scaling or translation), then the game is sometimes called *constant sum*, but according to our convention such games are also zero sum. A game is a *finite game* if each player has only a finite number of alternatives, that is the players pick their actions out of finite sets (action sets); otherwise the game is an *infinite game*; finite games are also known as *matrix games.* An infinite game is said to be a *continuous-kernel game* if the action sets of the players are continua, and the players' objective functions are continuous with respect to action variables of all players. A game is said to be *deterministic* if the players' actions uniquely determine the outcome, as captured in the objective functions, whereas if the objective function of at least one player depends on an additional variable (state of nature) with a known probability distribution, then we have a *stochastic game.* A game is a *complete information* game if the description of the game (that is, the players, the objective functions, and the underlying probability distributions (if stochastic)) is common information to all players; otherwise we have an *incomplete information* game. We say that a game is *static* if players have access to only the *a priori* information (shared by all), and none of the players has access to information on the actions of any of the other players; otherwise what we have is a *dynamic game.* A game is a *single-act game* if every player acts only once; otherwise the game is *multi-act.* Note that it is possible for a single-act game to be dynamic and for a multi-act game to be static. A dynamic game is said to be a *differential game* if the evolution of the decision process (controlled by the players over time) takes place in continuous time, and generally involves a differential equation; is it takes place over a discrete-time horizon, a dynamic game is sometimes called a *discrete-time game.*

1.2 Past and the present

Game Theory has enjoyed over 65 years of scientific development, with the publication of the *Theory of Games and Economic Behavior* by John von Neumann and Oskar Morgenstern [14] generally acknowledged to *kick start* the field. It has experienced incessant growth in both the number of theoretical results and the scope and variety of applications. As a recognition of the vitality of the field a total of 8 Nobel Prizes were given in Economic Sciences for work primarily in game theory, with the first such recognition given in 1994 to John Harsanyi, John Nash, and Reinhard Selten "for their pioneering analysis of equilibria in the theory of non-cooperative games." The second set of Nobel Prizes in game theory went to Robert Aumann and Thomas Schelling in 2005, "for having enhanced our understanding of conflict and cooperation through game-theory analysis." And the most recent one was in 2007, recognizing Leonid Hurwicz, Eric Maskin, and Roger Myerson, "for having laid the foundations of mechanism design theory." I should add to this list of highest-level awards in game theory, also the Crafoord Prize in 1999 (which is the highest prize in Biological Sciences), which went to John Maynard Smith (along with Ernst Mayr and G. Williams) "for developing the concept of evolutionary biology," where Smith's recognized contributions had a strong game-theoretic underpinning, through his work on evolutionary games and evolutionary stable equilibrium [8, 9, 10].

Even though von Neumann and Morgenstern's 1944 book is taken as the starting point of the scientific approach to game theory, game-theoretic notions and some isolated key results date back to earlier years. Sixteen years earlier, in 1928, John von Neumann himself had resolved completely an open fundamental problem in zero-sum games, that *every finite two-player zero-sum game admits a saddle point in mixed strategies*, which is known as the *Minimax Theorem* [15] — a result which Emile Borel had conjectured to be false eight years before. Some early traces of game-theoretic thinking can be seen in the 1982 work (*Considérations sur la théorie mathématique du jeu*) of André-Marie Ampère (1775-1836), who was influenced by the 1777 writings (Essai d'Arithmétique Morale) of Georges Louis Buffon (1707-1788).

Which event or writing has really started game-theoretic thinking or approach to decision making (in law, politics, economics, etc.) may be a topic of debate, but what is indisputable is that the second half of the twentieth century was a golden era of game theory, and the twenty-first century has started with a *big bang*, and is destined to be a *platinum* era with the proliferation of text-books, monographs, and journals covering the theory and applications (to an ever-growing breadth) of static and dynamic games. Some selected textbooks that cover non-cooperative as well as cooperative games, with various shades of mathematical rigor and sophistication, are (an incomplete list):

- G. Owen, Game Theory, 3rd edition, Academic Press, 1995
- D. Fudenberg, J. Tirole, Game Theory, MIT Press, 1991
- T. Başar and G.J. Olsder, Dynamic Noncooperative Game Theory, 2nd edition, SIAM Classics, 1999 (original: Academic Press, 1982)
- R.B. Myerson, Game Theory: Analysis of Conflict, Harvard, 1991
- N.H. Vorobev, Game Theory, Springer Verlag, 1977
- R. Gibbons, Game Theory for Applied Economists, Princeton University Press, 1992
- K. Binmore, Fun and Games, D.C. Heath and Co, 1992
- J.D. Williams, The Compleat Strategyst. McGraw-Hill, 1954
- W. Poundstone, Prisoners Dilemma, Doubleday, 1992

Two international societies exist,

- International Society of Dynamic Games (1990 -)
- Game Theory Society (1999 -)

as well as several regional ones, and the rate at which conferences and symposia on game theory or related to game theory are being organized is growing from year to year. Several journals in other fields (economics, biology, political science, sociology, communications, various disciplines of engineering, etc.) are publishing papers that use tools of game theory in increasing numbers, and there are at least four journals that are primarily on game theory:

- Games and Economic Behavior
- International J. Game Theory
- International Game Theory Review
- J. Dynamic Games and Applications (*just launched*)

Hence, in all respects game theory is on an upward slope in terms of its vitality, the wealth of topics that fall in its scope, the richness of the conceptual framework it offers, the range of applications, and the challenges it presents to an inquisitive mind.

1.3 On these Lecture Notes

This set of notes will expose the reader to some of the fundamental results in **non-cooperative game theory**, and highlight some current applications. It will cover both static and dynamic games, and zero-sum as well as nonzero-sum games. Details of proofs will not be given in most cases, and they can be found in standard texts (listed earlier), particularly in [1], as well as in some relevant papers. In the discussion of selected applications, some relevant journal or conference paper references will be given, and in fact reprints or pre-prints of some of these papers will be attached. As the title of the *Notes* indicates, the coverage will be restricted to non-cooperative games.

For readers who are familiar with optimization theory in its static (linear and nonlinear programming [11]) or dynamic (optimal control or calculus of variations [12]) forms, they can be viewed as special cases of a theory for nonzero-sum static and dynamic games when the number of players is restricted to *one*. It is therefore natural for tools of single-criterion optimization to play a role in the derivation and characterization of non-cooperative equilibrium solutions of zero-sum and nonzero-sum static and dynamic games, as we will see throughout these lectures.

2 Lecture 2: Non-cooperative games and equilibria

2.1 Main elements and equilibrium solution concepts

For a precise formulation of a non-cooperative game, we have have to specify (i) the number of players, (ii) the possible actions available to each player, and any constraints that may be imposed on them, (iii) the objective function of each player which she attempts to optimize (minimize or maximize, as the case may be), (iv) any time ordering of the execution of the actions if the players are allowed to act more than once, (v) any information acquisition that takes place and how the information available to a player at each point in time depends on the past actions of other players, and (vi) whether there is a player (*nature*) whose action is the outcome of a probabilistic event with a fixed (known) distribution. Here we will first consider formulation of games where only items (i)-(iii) above are relevant, that is players act only once, the game is static so that players do not acquire information on other players' actions, and there is no nature player. Subsequently, in later lectures, we will consider more general formulations, particularly dynamic games, that will incorporate all the ingredients listed above.

Accordingly, we consider an N-player game, with $\mathcal{N} := \{1, \ldots, N\}$ denoting the Players set. The decision or action variable of Player i is denoted by $x_i \in X_i$, where X_i is the action set of Player i. The action set could be a finite set (so that the player has only a finite number of possible actions), or an infinite but finite-dimensional set (such as the unit interval, $[0, 1]$), or an infinite-dimensional set (such as the space of all continuous functions on the interval $[0, 1]$). We let x denote the N-tuple of action variables of all players, $x := (x_1, \ldots, x_N)$. Allowing for possibly coupled constraints, we let $\Omega \subset X$ be the constraint set for the game, where X is the N-product of $X_1, \ldots, X_N$; hence for an N-tuple of action variables to be feasible, we need $x \in \Omega$ (for example, with $N = 2$, we could have a coupled constraint set described by: $0 \le x_1, x_2 \le 1,\ x_1 + x_2 \le 1$).

If we consider the players to be minimizers, the objective function (loss function or cost function) of Player i will be denoted by $L_i(x_i, x_{-i})$, where x_{-i} stands for the action variables of all players except the i'th one. If the players are maximizers, then the objective function (utility function) of Player i will be denoted by $V_i(x_i, x_{-i})$. In these *Notes*, we will use L_i and V_i interchangeably, depending on whether a player is respectively a minimizer or a maximizer. Note that a game where all players are minimizers, with cost functions L_i's, can be seen as one where all players are maximizers, with utility functions $V_i \equiv -L_i$'s.

Now, an N-tuple of action variables $x^* \in \Omega$ constitutes a *Nash equilibrium* (or, *non-cooperative equilibrium*) (NE) if, for all $i \in \mathcal{N}$,

$$L_i(x_i^*, x_{-i}^*) \leq L_i(x_i, x_{-i}^*)\,, \quad \forall\, x_i \in X_i\,, \text{ such that } (x_i, x_{-i}^*) \in \Omega\,, \tag{1}$$

or, if the players are maximizers,

$$V_i(x_i^*, x_{-i}^*) \geq L_i(x_i, x_{-i}^*)\,, \quad \forall\, x_i \in X_i\,, \text{ such that } (x_i, x_{-i}^*) \in \Omega\,. \tag{2}$$

If $N = 2$, and $L_1 \equiv -L_2 =: L$, then we have a two-player zero-sum game (ZSG), with Player 1 minimizing L and Player 2 maximizing the same quantity. In this case, the Nash equilibrium becomes the *saddle-point equilibrium* (SPE), which is formally defined as follows, where we leave out the coupling constraint set Ω (or simply assume it to be equal to the product set $X := X_1 \times X_2$): A pair of actions $(x_1^*, x_2^*) \in X$ is in *saddle-point equilibrium* (SPE) for a game with cost function L, if

$$L(x_1^*, x_2) \leq L(x_1^*, x_2^*) \leq L(x_1, x_2^*)\,, \quad \forall (x_1, x_2) \in X\,. \tag{3}$$

This also implies that the order in which minimization and maximization are carried out is inconsequential, that is

$$\min_{x_1 \in X_1} \max_{x_2 \in X_2} L(x_1, x_2) = \max_{x_2 \in X_2} \min_{x_1 \in X_1} L(x_1, x_2) = L(x_1^*, x_2^*) =: L^*$$

where the first expression on the left is known as the *upper value* of the game, the second expression is the *lower value* of the game, and L^* is known as the *value* of the game.[1] Note that we generally have

$$\min_{x_1 \in X_1} \max_{x_2 \in X_2} L(x_1, x_2) \geq \max_{x_2 \in X_2} \min_{x_1 \in X_1} L(x_1, x_2)\,,$$

or more precisely

$$\inf_{x_1 \in X_1} \sup_{x_2 \in X_2} L(x_1, x_2) \geq \sup_{x_2 \in X_2} \inf_{x_1 \in X_1} L(x_1, x_2)\,,$$

which follows directly from the obvious inequality

$$\sup_{x_2 \in X_2} L(x_1, x_2) \geq \inf_{x_1 \in X_1} L(x_1, x_2)\,,$$

[1] Upper and lower values are defined in more general terms using infimum (inf) and supremum (sup) replacing minimum and maximum, respectively, to account for the facts that minima and maxima may not exist. When the action sets are finite, however, the latter always exist.

since the LHS expression is only a function of x_1 and the RHS expression only a function of x_2.

Next, note that the value of a game, whenever it exists (which certainly does if there exists a saddle point), is *unique*. Hence, if there exist another saddle-point solution, say $(\hat{x}_1, \hat{x}_2)$, then $L(\hat{x}_1, \hat{x}_2) = L^*$. Moreover, these multiple saddle points are *orderly interchangeable*, that is the pairs $(x_1^*, \hat{x}_2)$ and $(\hat{x}_1, x_2^*)$ are also in saddle-point equilibrium. This property that saddle-point equilibria enjoy do not extend to multiple Nash equilibria (for nonzero-sum games): multiple Nash equilibria are generally not interchangeable, and further they do not lead to the same values for the Players' cost functions, the implication being that when players switch from one equilibrium to another, some players may benefit from that switch (in terms of reduction in cost) while others may see an increase in their costs. Further, if the players pick their actions randomly from the set of multiple Nash equilibria of the game, then the resulting N-tuple of actions may not be in Nash equilibrium.

Now coming back to the zero-sum game, if there is no value, which essentially means that the upper and lower values are not equal, in which case the former is strictly higher than the latter:

$$\min_{x_1 \in X_1} \max_{x_2 \in X_2} L(x_1, x_2) > \max_{x_2 \in X_2} \min_{x_1 \in X_1} L(x_1, x_2)\,,$$

then a saddle point does not exist. We then say in this case that *the zero-sum game does not have a saddle point in pure strategies*. This opens the door for looking for a *mixed-strategy* equilibrium. A *mixed strategy* is for each player a probability distribution over his action set, which we denote by p_i for Player i. This argument also extends to the general N-player game, which may not have a Nash equilibrium in pure strategies (actions, in this case). In search of a mixed-stragy equilibrium, L_i is replaced by its expected value taken with respect to the mixed strategy choices of the players, which we denote for Player i by $J_i(p_1, \ldots, p_N)$. Nash equilibrium over mixed strategies is then introduced as before, with just J_i's replacing L_i's, and p_i's replacing x_i's, and $p_i \in \mathcal{P}_i$, where $\mathcal{P}_i$ is the set of all probability distributions on X_i (we do not bring Ω into the picture here, assuming that the constraint sets are rectangular). If X_i is finite, then p_i will be a probability vector, taking values in the probability simplex determined by X_i. In either case, the N-tuple $(p_1^*, \ldots, p_N^*)$ is in (mixed-strategy) Nash equilibrium (MSNE) if

$$J_i(p_i^*, p_{-i}^*) \leq J_i(p_i, p_{-i}^*)\,, \quad \forall\, p_i \in \mathcal{P}_i\,. \tag{4}$$

This readily leads, in the case of zero-sum games, as a special case, to the following definition of a saddle point in mixed strategies: A pair (p_1^*, p_2^*) constitutes a *saddle point in mixed strategies* (or a *mixed-strategy saddle-point equilibrium*) (MSSPE), if

$$J(p_1^*, p_2) \leq J(p_1^*, p_2^*) \leq J(p_1, p_2^*)\,, \quad \forall\,(p_1, p_2) \in \mathcal{P}\,.$$

where $J(p_1, p_2) = E_{p_1,p_2}[L(x_1, x_2)]$, and $\mathcal{P} := \mathcal{P}_1 \times \mathcal{P}_2$. Here $J^* = J(p_1^*, p_2^*)$ is the value of the zero-sum game in mixed strategies.

2.2 Security strategies

If there is no Nash equilibrium in pure strategies, and the players do not necessarily want to adopt mixed strategies, an alternative approach is for each player to pick that pure strategy which will safeguard her losses under worst scenarios. This will entail each player essentially playing a zero-sum game, minimizing her cost function against collective maximization of all other players. A strategy (or an action, in this case) that provides a loss ceiling for a player is known as a *security strategy* for that player. Assuming again rectangular action product sets, security strategy $x_i^s \in X_i$ for Player i is defined through the relationship

$$\sup_{x_{-i} \in X_{-i}} L_i(x_i^s, x_{-i}) = \inf_{x_i \in X_i} \sup_{x_{-i} \in X_{-i}} L_i(x_i, x_{-i}) =: \bar{L}_i$$

where the "sup" could be replaced with "max" if the action sets are finite. Note that, the RHS value, $\bar{L}_i$, is the upper value of the zero-sum game played by Player i. Also note that even if the security strategies of the players, say $x^s := \{x_i^s,\ i \in \mathcal{N}\}$, are unique, then this N-tuple would not necessarily constitute an equilibrium in any sense. In the actual play, the player will actually end up doing better than just safeguarding their losses, since $L^i(x^s) \leq \bar{L}^i$ for all $i \in \mathcal{N}$.

The notion of a security strategy could naturally be extended to also mixed strategies. Using the earlier notation, $p_i^s \in \mathcal{P}_i$ would be a *mixed security strategy* for Player i if

$$\sup_{p_{-i} \in \mathcal{P}_{-i}} J_i(p_i^s, p_{-i}) = \inf_{p_i \in \mathcal{P}_i} \sup_{p_{-i} \in \mathcal{P}_{-i}} J_i(p_i, p_{-i}) =: \bar{J}_i$$

Remark. If the original game is a two-player zero-sum game, and the upper and lower values are equal, then security strategies for the players will have to be in SPE. If the upper and lower values are not equal in pure strategies, but are in mixed strategies, then mixed security strategies for the players will have to be in MSPE.

2.3 Strategic equivalence

Strategic equivalence is a useful property (observation) that facilitates study of non-cooperative equilibria of nonzero-sum games (NZSGs). Let us now make the simple observation that given an N-player NZSG of the type introduced in this lecture, if two operations are applied to the loss function of a player, *positive scaling* and *translation* which does not depend on the action variable of that player, this being so for every player, then the set of NE of the resulting NZSG is identical to the set of NE of the original game. In view of this property, we say that the two games are *strategically equivalent.* In mathematical terms, if $\tilde{L}_i$'s are the cost functions of the players in the transformed game, then we have, for some constants $\alpha_i > 0,\ i \in \mathcal{N}$, and some functions $\beta_i(x_{-i}),\ i \in \mathcal{N}$,

$$\tilde{L}_i(x_i, x_{-i}) = \alpha_i \, L_i(x_i, x_{-i}) + \beta_i(x_{-i}) \, , \ i \in \mathcal{N} \, .$$

Now note that, if for a given NZSG, there exist α_i's and β_i's of the types above, such that $\tilde{L}_i$ is independent of i, that is the transformed NZSG features the same cost function, say $\tilde{L}$, for all players, then we have a single objective game, or equivalently a *team* problem. Any NE of this transformed game (which is a team) is a *person-by-person* optimal solution of the team problem. That is, if $x_i^*,\ i \in \mathcal{N}$ is one such solution, we have

$$\tilde{L}(x_i,^* x_{-i}^*) = \min_{x_i \in X_i} \tilde{L}(x_i, x_{-i}^*) \, , \ \forall \, i \in \mathcal{N}$$

which is not as strong as the globally minimizing solution for $\tilde{L}$:

$$\tilde{L}(x_i,^* x_{-i}^*) = \min_{x_i \in X_i} \min_{x_{-i} \in X_{-i}} \tilde{L}(x_i, x_{-i}^*) \, .$$

Clearly, the latter implies the former, but not *vice versa*. Consider, for example, the two-player game where each player has two possible actions, for which $\tilde{L}$ admits the matrix representation

$$\tilde{L} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \tag{5}$$

where Player 1 is the row player, and Player 2 the column player (and both are minimizers). The south-east entry (row 2, column 2) is clearly a person-by-person optimal solution (NE), but is not the globally minimum one, which is the north-west entry (row 1, column 1) (which is of course also a person-by-person optimal solution). Of course, if the players were to cooperate, they would unquestionably pick the latter, but since this is a non-cooperative game, they are not allowed to correlate their choices. With the entries as above, however, the chances of them ending up at the global minimum are very high, because neither one would end up worse than the inferior NE if they stick to the first row and first column (even if one player inadvertently deviates). But this is not the whole story, because it would be misleading to make the *mutual benefit* argument by working on the transformed game. Consider now the following two-player, two-action NZSG, where again Player 1 is the row player, and Player 2 the column player:

$$L_1 = \begin{pmatrix} 99 & 1 \\ 100 & 1 \end{pmatrix}, \quad L_2 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \tag{6}$$

This game has two pure-strategy NE, (row 1, column 1) and (row 2, column 2), the same as the game $\tilde{L}$. In fact it is easy to see that the two games are strategically equivalent (subtract 99 from the first column of L_1). But now, Player 1 would prefer the south-west entry (that is what was inferior in the transformed game), which shows that that are perils in jumping to conclusions based on a transformed game.

When this all comes handy, however, is when the transformed game as a team problem can be shown to have a unique person-by-person optimal solution, which is also the globally optimal team solution. We will later see games with structures where this would happen. Then, there would be no ambiguity in the selection of the unique NE.

For a given NZSG, if there exists a strategically equivalent team problem, then we say that the original game is *team like*. There could also be situations when a game is strategically equivalent to a zero-sum team problem, that is there exists a proper subset of $\mathcal{N}$, say $\mathcal{N}_1$, such that for $i \in \mathcal{N}_1$, $\tilde{L}_i$ is independent of i, say $\tilde{L}$, **and** for $j \notin \mathcal{N}_1$, $\tilde{L}_j \equiv -\tilde{L}$. This means that there exists a strategically equivalent game where players in $\mathcal{N}_1$ form a team, playing against another team comprised of all players outside $\mathcal{N}_1$. In particular, if $N = 2$, we have every NE of the original game equal to the SPE of the transformed strategically equivalent ZSG.

3 Lecture 3: Finite games, and existence and computation of NE

3.1 Zero-sum finite games and the Minimax Theorem

Let us first consider two-player zero-sum finite games, or equivalently matrix games. For any such game we have to specify the cardinality of action sets X_1 and X_2 (card (X_1) and card (X_2)), and the objective function $L(x_1, x_2)$ defined on the product of these finite sets. As per our earlier convention, Player 1 is the minimizer and Player 2 the maximizer. Let card $(X_1) = m$ and card $(X_2) = n$, that is the minimizer has m choices and the maximizer has n choices, and let the elements of X_1 and X_2 be ordered according to some (could be arbitrary) convention. We can equivalently associate an $m \times n$ matrix A with this game, whose entries are the values of $L(x_1, x_2)$, following the same ordering as that of the elements of the action sets, that is ij'th entry of A is the value of $L(x_1, x_2)$ when x_1 is the i'th element of X_1 and x_2 is the j'th element of X_2. Player 1's choices are then the rows of the matrix A and Player 2's are its columns.

It is easy to come of with example matrix games where a saddle point does not exist in pure strategies, with perhaps the simplest one being the game known as *Matching Pennies*, where

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \tag{7}$$

and each entry is cost to Player 1 (minimizer), and payoff to Player 2 (maximizer). Here there is no row-column combination at which the players would not have an incentive to deviate and improve their returns.

The next question is whether there exists a saddle point in mixed strategies. Assume that Player 1 now picks row 1 and row 2 with equal probability $\frac{1}{2}$. Then, regardless of whether Player 2 picks column 1 or column 2, she will face the same expected cost of 0. Hence, in response to this equal probability choice of Player 1, Player 2 is indifferent between the two actions available to her; she could pick column 1, or column 2, or any probability mix between the two. Likewise, if Player 2 picks column 1 and column 2 with equal probability $\frac{1}{2}$, this time Player 1 faces an expected cost of 0 regardless of her choice. In view of this, the mixed strategy pair $\left(p_1^* = (\frac{1}{2}, \frac{1}{2}), p_2^* = (\frac{1}{2}, \frac{1}{2})\right)$ is a MSSPE, and in fact is the *unique* one. The SP value in mixed strategies is 0.

To formalize the above, let A be an $m \times n$ matrix representing the finite ZSG, and as before let p_1 and p_2 be the probability vectors for Players 1 and 2, respectively (both column vectors, and note that in this case p_1 is of dimension m and p_2 is of dimension n, and components of each are nonnegative and add up to 1). We can then rewrite the expected cost function as

$$J(p_1, p_2) = p_1' A p_2 \,.$$

By the *minimax theorem*, due to *John von Neumann (1928)*, J indeed admits a saddle point, which means that the matrix game A has a saddle point in mixed strategies, that is there exists a pair (p_1^*, p_2^*) such that for all other probability vectors p_1 and p_2, of dimensions m and n, respectively, the following pair of saddle-point inequalities hold:

$$p_1^{*\prime} A p_2 \leq p_1^{*\prime} A p_2^* \leq p_1' A p_2^* \tag{8}$$

The quantity $p_1^{*\prime} A p_2^*$ is the *value* of the game in mixed strategies. This result is now captured in the following *Minimax Theorem*.

Theorem 1 *Every finite two-person zero-sum game has a saddle point in mixed strategies.*

Proof (outline): The proof uses the alternating hypotheses lemma in matrix theory, which says that given the matrix A as above, either there exists $y \in \mathbf{R}^m, y \geq 0$, such that $y' A \leq 0$, or there exists $z \in \mathbf{R}^n, z \geq 0$, such that $Az \geq 0$. Details can be found in ([1], p. 26). $\diamond$

3.2 Neutralization and domination

A mixed strategy that assigns positive probability to every action of a player is known as an *inner mixed strategy*. A MSSPE where both strategies are inner mixed is known as an *inner MSSPE*, or a completely mixed MSSPE. Note that if (p_1^*, p_2^*) is an inner MSSPE, then $p_1' A p_2^*$ is independent of p_1 on the m-dimensional probability simplex, and $p_2' A' p_1^*$ is independent of p_2 on the n-dimensional probability simplex. The implication is that in an inner MSSPE all the players do is to *neutralize* each other, and the solution would be the same if their roles were reversed (that is, Player 1 the maximizer, and Player 2 the minimizer). This suggests an obvious computational scheme for solving for the MSSPE, which involves solving linear algebraic equations for p_1 and p_2, of course provided that MSSPE is inner.

Now, if MSSPE is not inner but is proper mixed, that is it is not a pure-strategy SPE, then a similar neutralization will hold in a lower dimension. For example, if (p_1^*, p_2^*) is a MSSPE where some components of p_2^* are zero, then p_1^* will neutralize only the actions of Player 2 corresponding to the remaining components of p_2^* (which are positive), with the expected payoff for Player 2 (which is minus the cost) corresponding to the non-neutralized actions being no smaller than the neutralized ones. In this case, whether a player is a minimizer or a maximizer does make a difference. The following game, which is an expanded version of *Matching Pennies*, where Player 2 has a third possible action illustrates this point:

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & -1 \end{pmatrix} \tag{9}$$

Here, the MSSPE is $\left(p_1^* = (\frac{1}{2}, \frac{1}{2}), p_2^* = (\frac{1}{2}, \frac{1}{2}, 0)\right)$, where Player 1 neutralizes only the first two actions of Player 2, with the expected cost of the third action being $-\frac{1}{2}$, lower than 0, and hence Player 2, being the maximizer, would not put any positive weight on it. Note that in this game it does make a difference whether a player is a minimizer or a maximizer, because if we reverse the roles (now Player 1 is the maximizer, and Player 2 the minimizer), the SP value in mixed strategies is no longer 0, but is $-\frac{1}{3}$, with the MSSPE being $\left(p_1^* = (\frac{2}{3}, \frac{1}{3}), p_2^* = (0, \frac{1}{3}, \frac{2}{3})\right)$. Player 2 ends up not putting positive probability to the first column, which is *dominated* by the third column. *Domination* can actually be used to eliminate columns and/or rows which will not affect the MSSPE, and this will lead to reduction in the size of the game (and hence make computation of MSSPE more manageable). MSSPE of a reduced ZS matrix game (reduced through domination) is also an MSSPE of the original ZSG (with appropriate lifting to the higher dimension, by assigning eliminated columns or rows zero probability), but in this process some mixed SP strategies may also be eliminated. This, however, is not a major issue because the MSSPE value is unique. If only strictly dominated rows and columns are eliminated,[2] then all mixed SP strategies are preserved (see, [1, 4]).

3.3 Off-line computation of MSSPE

We have seen in the section above that inner MSSPE can be computed using the idea of neutralization and solving linear algebraic equations. The same method can in principle be applied to MSSPE that are not inner, but then one has to carry out an enumeration by setting some components of the probability vectors to zero, and looking for neutralization in a reduced dimension—a process which converges because MSSPE exists by the minimax theorem. In this process, domination can

[2] We say that *a row strictly dominates another row* if the difference between the two vectors (first one minus the second one) has all negative entries. Likewise, *a column strictly dominates another column* if the difference has all positive entries.

be used (as discussed above) to eliminate some rows or columns, which would sometimes lead to a (reduced) game with an inner MSSPE.

Yet another approach to computation of MSSPE is a graphical one, which however is practical only when one of the players has only two possible actions ([1], pp. 29-31). And yet another off-line computational method is to use the powerful tool of linear programming (LP). One can actually show that there is a complete equivalence between a matrix game and an LP. The following proposition captures this result, a proof of which can be found in [1].

Proposition 1 *Given a ZS matrix game described by the $m \times n$ matrix A, let B be another matrix game (strategically equivalent to A), obtained from A by adding an appropriate positive constant to make all its entries positive. Let $V_m(B)$ denote the SP value of B in mixed strategies. Introduce the two LPs:*[3]

Primal LP: $\max y'\mathbf{1}_m$ *such that* $B'y \leq \mathbf{1}_n\,,\; y \geq 0$

Dual LP: $\min z'\mathbf{1}_n$ *such that* $Bz \geq \mathbf{1}_m\,,\; z \geq 0,$

with their optimal values (if they exist) denoted by V_p and V_d, respectively. Then:

(i) Both LPs admit solutions, and $V_p = V_d = 1/V_m(B)$.

(ii) If (y^*, z^*) solves matrix game B, $y^*/V_m(B)$ solves the primal LP, and $z^*/V_m(B)$ solved the dual LP.

(iii) If $\tilde{y}^*$ solves the primal LP, and $\tilde{z}^*$ solves the dual LP, the pair $(\tilde{y}^*/V_p, \tilde{z}^*/V_d)$ constitutes a MSSPE for the matrix game B, and hence for A, and $V_m(B) = 1/V_p$.

3.4 Nonzero-sum finite games and Nash's Theorem

We now move on to N-player NZS finite games, and study the Nash equilibrium (NE), introduced in Lecture 2. As in the case of ZSGs, it is easy to come up with examples of games which do not admit NE in pure strategies. The question then is whether there is a counterpart of the minimax theorem in this case, which guarantees existence of NE in mixed strategies. This is indeed the case—a result established by *John Nash (1951)*[17], and captured in the following theorem.

Theorem 2 *Every finite N-player nonzero-sum game has a Nash equilibrium in mixed strategies.*

Proof (outline): We provide here an outline of the proof for the two-player case. Extension to the N-player case follows similar lines. In the two-player case, the game is known as a *bi-matrix game*,

[3] The notation $\mathbf{1}_m$ below stands for the m-dimensional column vector whose entries are all 1's.

characterized by two $m \times n$ matrices, $A = \{a_{ij}\}$ and $B = \{b_{ij}\}$, where A's entries are the cost to Player 1, and B's entries are the cost to Player 2 (both players are minimizers). Player 1 has m actions) and Player 2 has n actions. Let $p_1 = y$ be a mixed strategy for Player 1, and $p_2 = z$ be a mixed strategy for Player 2, which belong to the simplices Y and Z, respectively. Then, expected costs for the players can be written as

$$J_1 = y'Az = \sum\sum y_i a_{ij} z_j \qquad \text{and} \qquad J_2 = y'Bz = \sum\sum y_i b_{ij} z_j$$

Introduce the functions $\psi_k^1,\ k = 1, \ldots, m$, and $\psi_\ell^2,\ \ell = 1, \ldots, n$ by

$$\psi_k^1(y,z) := J_1(y,z) - \sum_j a_{kj} z_j\,, \qquad \psi_\ell^2(y,z) := J_2(y,z) - \sum_i b_{i\ell} y_i\,,$$

and let $c_k^1(y,z) := \max\{\psi_k^1(y,z), 0\}$, $c_\ell^2(y,z) := \max\{\psi_\ell^2(y,z), 0\}$, and note that both $c_k^1(\cdot,\cdot)$ and $c_\ell^2(\cdot,\cdot)$ are continuous on $Y \times Z$, since ψ_k^1 and ψ_ℓ^2 are. Now, introduce the transformations

$$\bar{y}_k = \frac{y_k + c_k^1(y,z)}{1 + \sum_i c_i^1(y,z)}\,, \quad \bar{z}_\ell = \frac{z_\ell + c_\ell^2(y,z)}{1 + \sum_j c_j^2(y,z)}\,,$$

and note that $\bar{y} \in Y$ and $\bar{z} \in Z$. The above induces a continuous map, T, from $Y \times Z$ into itself, that is

$$\begin{pmatrix} \bar{y} \\ \bar{z} \end{pmatrix} = T(y,z) \tag{10}$$

The proof proceeds by showing that if (y^*, z^*) is a MSNE of the bi-matrix game, then it is a fixed point of T, and conversely any fixed point of T (which exists, using Brouwer's fixed point theorem; see [1]),[4] is a MSNE of the bi-matrix game. The first part follows by noting that $\psi_k^1(y^*,z^*) \le 0, \forall k$, and $\psi_\ell^2(y^*,z^*) \le 0, \forall \ell$, and hence that $c_k^1(y^*,z^*) = c_\ell^2(y^*,z^*) = 0$, and therefore $((y^*)'\ (z^*)')' = T(y^*,z^*)$. And in the other direction it is a proof by contradiction. $\diamond$

Note that clearly the minimax theorem follows from this one since zero-sum games are special cases of nonzero-sum games. The main difference between the two, however, is that in zero-sum games the *value* is unique (even though there may be multiple saddle-point solutions), whereas in genuine nonzero-sum games the expected cost N-tuple to the players under multiple Nash equilibria need not be the same. In zero-sum games, multiple equilibria have the ordered interchangeability property, whereas in nonzero-sum games they do not, as we have discussed in Lecture 2. As an

[4] Brouwer's theorem says that a continuous mapping, f, of a closed, bounded, convex subset (S) of a finite-dimensional space into itself (that is, S) has a fixed point, that is a $p \in S$ such that $f(p) = p$.

illustrative example, consider the *Battle of the Sexes* game, where a husband and wife are faced with the decision of choosing between going to a concert and going to a soccer game, where the former is preferred by her and the latter by him. They also prefer being together at one of these events over going their separate ways. Two bi-matrix games that capture this scenario are (and they are strategically equivalent):

$$\text{Game 1}: \begin{pmatrix} (2,1) & (0,0) \\ (0,0) & (1,2) \end{pmatrix} \qquad \text{Game 2}: \begin{pmatrix} (1,0) & (-1,-1) \\ (-1,-1) & (0,1) \end{pmatrix} \tag{11}$$

Here wife is the row player, and husband is the column player, and they are both utility maximizers. The first listed choice for both of them is *Concert* (C), and the second listed choice is *Soccer* (S). The ordered payoff $(2,1)$, for example, means that her payoff is 2 and his is 1.

Since these two games are strategically equivalent, they admit the same set of equilibria (but not the values). There are two pure-strategy NE, $(C.C)$ and (S,S), with corresponding payoff pairs $(2,1)$ and $(1,2)$ for Game 1, but also a MSNE, $((2/3,1/3),\ (1/3,2/3))$, with a corresponding expected payoff pair of $(2/3,2/3)$ for Game 1. And these multiple NE are not interchangeable.

The notions of inner mixed equilibria, neutralization, and domination introduced earlier in the context of SPE and MSSPE equally apply here, and particularly the inner MSNE also has the neutralization property and can be solved using algebraic equations. These equations, however, will not be linear unless $N = 2$, that is the NZSG is a bi-matrix game. In two-player NZSGs, a counterpart of the LP equivalence exists, but this time it is a bi-linear program, as captured in the following proposition; for a proof, see [1], pp. 96-97.

Proposition 2 *For a bi-matrix game (A,B), where players are minimizers, a pair (y^*,z^*) constitutes a MSNE if, and only if, there exists a pair of real numbers (p^*,q^*) such that the quadruple (y^*,z^*,p^*,q^*) solves the bi-linear program:*

$$\min_{y,z,p,q} \ [y'AZ + y'Bz + p + q]$$

such that

$$Az \geq -p\mathbf{1}_m\,, \quad B'y \geq -q\mathbf{1}_n\,, \quad y \geq 0\,,\ z \geq 0\,,\ \ y'\mathbf{1}_m = 1\,,\ z'\mathbf{1}_n = 1\,.$$

3.5 On-line computation of MSSPE and MSNE: Fictitious play

In the discussion of the computation of MSNE as well as MSSPE we have so far focused on off-line methods, where the assumption was that the players have access to the entire game parameters (including other players' payoff or cost matrices). This however may not always be possible, which then begs the question as to whether it would be possible for the players to end up at a MSSPE or MSNE by following a process where each observes others' actions in a repetition of the game, and builds probabilistic beliefs (empirical probabilities) on other players' moves. Such a process is known as a *fictitious play* (FP). We say that the process converges in beliefs to equilibrium (MSSPE or MSNE, as the case may be) if the sequence of beliefs converges to an equilibrium. We further say that a game has the *fictitious play property* (FPP) if every fictitious play process converges in beliefs to equilibrium.

For some historical background, the fictitious play process was first suggested by Brown (1949)[28] as a mechanism to compute MSNE of a finite NZSG. Robinson(1950)[29] then proved that every two-player ZSG has the FPP. Miyasawa (1961)[30] proved (using a particular tie-breaking rule) that every two-player 2×2 bi-matrix game has the FPP. Shapley (1964)[32] constructed an example of a 3×3 two-player bi-matrix game which does not have the FPP. Thirty-two years later, Monderer and Shapley (1996)[31] proved the FPP for games with identical interests. A stochastic extension of FP, where each player's payoff function includes an additional entropy term to randomize her actions, were discussed recently by Shamma and Arslan(2004, 2005)[33, 34].

Let us now be more precise regarding the FPP, and restrict the discussion to bi-matrix games, represented by (A, B), where the players are maximizers. Let us assume that at times $k = 0, 1, 2, \ldots$, the players pick some actions, according to some rule, and they observe each other's past actions, and based on the history build empirical probabilities using the frequency of occurrences. Let $v_2(i, k)$ be a zero-one variable, which takes the value one if Player 2 chooses her action i at time k, and otherwise it is zero. If $\tilde{z}(k)$ denotes the empirical probability vector Player 1 constructs at time k based on actions of Player 2 up to (and including) time k, then we have the natural relationship

$$\tilde{z}_i(k+1) = \frac{1}{k+1} \sum_{\ell=0}^{k} v_2(i, \ell)\,, \quad i = 1, \ldots, n,$$

which can be written as a first-order difference equation

$$\tilde{z}_i(k+1) = \frac{k}{k+1} \tilde{z}_i(k) + \frac{1}{k+1} v_2(i, k)\,, \quad i = 1, \ldots, n\,. \tag{12}$$

Likewise, we can have a recursion for Player 2's empirical probability on Player 1:

$$\tilde{y}_j(k+1) = \frac{k}{k+1} \tilde{y}_j(k) + \frac{1}{k+1} v_1(j, k)\,, \; j = 1, \ldots, m\,. \tag{13}$$

Now, $v_2(i,k)$ and $v_1(j,k)$ are generated by each player maximizing her expected payoff given the empirical probability associated with the other player, namely $v_1(j,k) = 1$ if the j'th entry of the vector $A\tilde{z}(k)$ is maximum. Of course in general the maximizing integer will not be unique, in which case some *random* selection will have to be made from among the maximizing integers. An alternative is to add a softening entropy term to the maximand, carry out the maximization over probabilities, and then pick the action according to the maximizing probability distribution. Specifically, Player 1 maximizes at time k the function

$$U_1(y) = y'A\tilde{z}(k) + \tau_1 H(y)\,, \quad H(y) := -\sum_{j=1}^{m} y_j \log(y_j)$$

where $\tau_1 > 0$ is some parameter, and H is the entropy function. Since entropy is a strictly concave function of y, $U_1(y)$ is strictly concave, being maximized over the closed simplex Y; hence it has a unique maximum. Denote this maximizing solution by $\beta_1(\tilde{z}(k))$, which is a probability vector for each $\tilde{z}(k)$. The expression for β_1 actually turns out to be:

$$\beta_1(\tilde{z}) = \sigma(A\tilde{z}/\tau_1)$$

where σ is the *logit* or *soft max* function, mapping a Euclidean space(in this case of dimension m) into the appropriate probability simplex (in this case Y), whose i'th component is given by

$$(\sigma(x))_i = e^{x_i} / \sum_{j=1}^{m} e^{x_j}\,.$$

Now choose the action at time k, say $a_1(k)$, in the support set of $\beta_1(\tilde{z}(k))$ and randomly, with the property that $E[a_1(k)] = \beta_1(\tilde{z}(k))$; we write this relationship as

$$a_1(k) = \mathrm{rand}(\beta_1(\tilde{z}(k)))$$

where "rand" stands for the *randomizer function.* Similarly for Player 2,

$$a_2(k) = \mathrm{rand}(\beta_2(\tilde{y}(k)))\,, \quad \beta_2(\tilde{y}(k)) = \arg\max_{z\in Z}\,[\tilde{y}(k)'Bz + \tau_2 H(z),$$

where β_2 admits the closed-form expression $\beta_2(\tilde{y}) = \sigma(B'\tilde{z}/\tau_2)$. Now, in view of this, taking averages in (12) and (13), we arrive at the vector recursions

$$\tilde{z}(k+1) = \frac{k}{k+1}\tilde{z}(k) + \frac{1}{k+1}\beta_2(\tilde{y}(k))\,, \tag{14}$$

$$\tilde{y}(k+1) = \frac{k}{k+1}\tilde{y}(k) + \frac{1}{k+1}\beta_1(\tilde{y}(k))\,. \tag{15}$$

which are the updates of the empirical frequencies, based on the players' observations of their actions over the course of repetitions of the game. If $\tau_1 = \tau_2 = 0$, then we have the standard (classical) FP, but when they are taken to be positive we have the stochastic FP, which rewards randomization and hence leads to unique mixed-strategy responses. If these parameters are taken to be small but positive, then they can be viewed as capturing random perturbations in the entries of the payoff matrices.

To obtain convergence proofs, it is useful to work with continuous, differential equation versions of the recursions (14) and (15) (known as continuous-time FP), which are obtained by letting the time that elapses between two successive updates vanish: [5]

$$\dot{\tilde{y}}(t) = -\tilde{y}(t) + \beta_1(\tilde{z}(t)) \;\text{ and }\; \dot{\tilde{z}}(t) = -\tilde{z}(t) + \beta_2(\tilde{y}(t))\,,$$

If these converge (that is if the coupled systems are asymptotically stable), then the limit point, say (y^*, z^*) is a fixed point of

$$y = \beta_1(z)\,, \quad z = \beta_2(y)$$

which is also a NE of the game with utility functions

$$U_1(y,z) = y'Az + \tau_1 H(y)\,, \;\; U_2(y,z) = y'Bz + \tau_2 H(y)\,,$$

and for sufficiently small τ_1 and τ_2, they constitute an approximation to the MSNE of the bi-matrix game (A, B). It has been shown by Shamma and Arslan (2004)[33] using Lyapunov stability techniques that under some non-singularity conditions, if we have either a zero-sum game ($A = -B$), or an identical interest game (A=B), or a bi-matrix game with either $m = 2$ or $n = 2$, convergence takes place, that is

$$\lim_{t\to\infty}(y(t) - \beta_1(z(t))) = 0 \;\text{ and }\; \lim_{t\to\infty}(z(t) - \beta_2(y(t))) = 0\,,$$

It should be noted, however, that this does not necessarily imply that the original discrete-time iteration converges almost surely. Further recent work on this topic can be found in [34], [35], [36], as well as [37] which discusses applications to security games.

[5] Here we have reversed the order of appearance.

4 Lecture 4: Games in extensive form

If players act in a game more than once, and at least one player has information (complete or partial) on past actions of other players, then we are in the realm of *dynamic games* (as mentioned earlier), for which a complete description (in finite games) involves a tree structure where each node is identified with a player along with the time when she acts, and branches emanating from a node show the possible moves of that particular player. A player, at any point in time, could generally be at more than one node—which is a situation that arises when the player does not have complete information on the past moves of other players, and hence may not know with certainty which particular node she is at at any particular time. This uncertainty leads to a clustering of nodes into what is called *information sets* for that player. A precise definition of extensive form of a dynamic game now follows.

Definition 1 Extensive form *of an N-person nonzero-sum finite game without chance moves is a tree structure with*

(i) a specific vertex indicating the starting point of the game,

(ii) N cost functions*, each one assigning a real number to each terminal vertex of the tree, where the ith cost function determines the loss to be incurred to $\mathbf{P}i$,*

(iii) a partition of the nodes of the tree into N player sets*,*

(iv) a subpartition of each player set into information sets *$\{\eta^i_j\}$, such that the same number branches emanate from every node belonging to the same information set and no node follows another node in the same information set.* ◇

What players decide on within the framework of the extensive form is not their actions, but their *strategies*, that is what action they should take at each information set. They then take specific actions (or actions are executed on their behalf), dictated by the strategies chosen as well as the progression of the game (decision) process along the tree. A precise definition now follows.

Definition 2 *Let N^i denote the class of all information sets of $\mathbf{P}i$, with a typical element designated as η^i. Let $U^i_{\eta^i}$ denote the set of alternatives of $\mathbf{P}i$ at the nodes belonging to the information* *set η^i. Define $U^i = \cup U^i_{\eta^i}$, where the union is over $\eta^i \in N^i$. Then, a* strategy *γ^i for $\mathbf{P}i$ is a mapping from N^i into U^i, assigning one element in U^i for each set in N^i, and with the further property that $\gamma^i(\eta^i) \in \mathbf{U}^i_{\eta^i}$ for each $\eta^i \in \mathcal{N}^i$. The set of all strategies of $\mathbf{P}i$ is called his* strategy set *(*space*), and it is denoted by Γ^i.* $\diamond$

Let $J^i(\gamma^1, \ldots, \gamma^N)$ denote the loss incurred to $\mathbf{P}i$ when the strategies $\gamma^1 \in \Gamma^1, \ldots, \gamma^N \in \Gamma^N$ are adopted by the players. This construction leads to what is known as the *normal form* of the dynamic game, which in a sense is no different from the matrix forms we have seen in the earlier lectures. In particular, for a finite game with a finite duration (that is players act only a finite number of times), the number of elements in each Γ^i is finite, and hence the game can be viewed as a matrix game, of the type considered earlier. In this normal form, the concept of Nash equilibrium (NE) is introduced in exactly the same way as in static games, with now the actions variables replaced by strategies. Hence we have:[6]

Definition 3 *An N-tuple of strategies $\gamma^* := \{\gamma^{1*}, \gamma^{2*}, \ldots, \gamma^{N*}\}$ with $\gamma^{i*} \in \Gamma^i$, $i \in \mathcal{N}$ constitutes a* noncooperative (Nash) equilibrium solution *for an N-person nonzero-sum finite game in extensive form, if the following N inequalities are satisfied for all $\gamma^i \in \Gamma^i$, $i \in \mathcal{N}$:*[7]

$$J^{1*} := J^i(\gamma^{i*}, \gamma^{-i*}) \leq J^i(\gamma^i, \gamma^{-i*})$$

The N-tuple of quantities $\{J^{1*}, \ldots, J^{N*}\}$ is known as **a** *Nash equilibrium outcome of the nonzero-sum finite game in extensive form.* $\diamond$

I emphasize the word **a** in the last sentence of the preceding definition, since Nash equilibrium solution could possibly be non-unique, with the corresponding set of Nash values being different. This then leads to a partial ordering in the set of all Nash equilibrium solutions.

As in the case of static (matrix) games, pure-strategy NE may not exist in dynamic games also. This leads to the introduction of mixed strategies, which are defined (quite analogously to the earlier definition) as probability distributions on Γ^i's, that is for each player as a probability

[6]Even though the discussion in this lecture uses the framework of N-player non-cooperative games with NE as the solution concept, it applies as a special case to two-player zero-sum games, by taking $J^1 = -J^2$ and noting that in this case NE becomes SPE.

[7]Using the earlier convention, the notation γ^{-i} stands for the collection of all players' strategies, except the i'th one.

distribution on the set of all her pure strategies; denote such a collection for $\mathbf{P}i$ by $\bar{\Gamma}^i$. A MSNE is then defined in exactly the same way as before. Again, since in normal form a finite dynamic game with a finite duration (I will call such games *finite-duration multi-act finite games*) can be viewed as a matrix game, there will always exist a MSNE by Nash's theorem:

Proposition 3 *Every N-person nonzero-sum finite-duration multi-act finite game in extensive form admits a Nash equilibrium solution in mixed strategies (MSNE).*

A MSNE may not be desirable in a multi-act game, because it allows for a player to correlate her choices across different information sets. A *behavioral strategy*, on the other hand, allows a player to assign independent probabilities to the set of actions at each information set (that is independent across different information sets); it is an appropriate mapping whose domain of definition is the class of all the information sets of the player. By denoting the behavioral strategy set of $\mathbf{P}i$ by $\hat{\Gamma}^i$, and the average loss incurred to $\mathbf{P}i$ as a result of adoption of the behavioral strategy N-tuple $\{\hat{\gamma}^1 \in \hat{\Gamma}^1, \ldots, \hat{\gamma}^N \in \hat{\Gamma}^N\}$ by $\hat{J}(\hat{\gamma}^1, \ldots, \hat{\gamma}^N)$, the definition of a Nash equilibrium solution in behavioral strategies (BSNE) may be obtained directly from Definition 3 by replacing γ^i, Γ^i and J^i with $\hat{\gamma}^i$, $\hat{\Gamma}^i$ and $\hat{J}^i$, respectively. A question of interest now is whether a BSNE is necessarily also a MSNE. The following proposition settles that.

Proposition 4 *Every BSNE of an N-person nonzero-sum multi-act game also constitutes a Nash equilibrium in the larger class of mixed strategies (that is a MSNE).*

Proof. Let $\hat{\gamma}^* := \{\hat{\gamma}^{1*} \in \hat{\Gamma}^1, \ldots, \hat{\gamma}^{N*} \in \hat{\Gamma}^N\}$ denote an N-tuple of behavioral strategies stipulated to be in Nash equilibrium, and let $\bar{\Gamma}^i$ denote the mixed-strategy set of $\mathbf{P}i$, $i \in \mathbf{N}$. Since $\hat{\Gamma}^i \subset \bar{\Gamma}^i$, we clearly have $\hat{\gamma}^{i*} \in \bar{\Gamma}^i$, for every $i \in \mathbf{N}$. Assume, to the contrary, that $\hat{\gamma}^*$ is not a MSNE; then this implies that there exists at least one i (say, $i = N$, without any loss of generality) for which the corresponding inequality of [8]

$$\hat{J}^i(\hat{\gamma}^i*, \hat{\gamma}^{-i}*) \leq \hat{J}^i(\hat{\gamma}^i, \hat{\gamma}^{-i}*)\,, \quad \forall \hat{\gamma}^i \in \hat{\Gamma}^i\,, \quad \forall i \in \mathcal{N}$$

is not satisfied for all $\bar{\gamma}^i \in \bar{\Gamma}^i$. In particular, there exists a $\bar{\gamma}^N \in \bar{\Gamma}^N$ such that

$$\hat{J}^{N*} > \hat{J}^N(\hat{\gamma}^{1*}; \ldots; \hat{\gamma}^{N-1*}; \bar{\gamma}^N) \triangleq F^*. \tag{i}$$

[8]Here, by an abuse of notation, we take $\hat{J}^i$ to denote the average loss to $\mathbf{P}i$ under also mixed-strategy N-tuples.

Now, abiding by our standard convention, let Γ^N denote the pure-strategy set of $\mathbf{P}N$, and consider the quantity

$$F(\gamma^N) \triangleq \hat{J}^N(\hat{\gamma}^{1*};\ldots;\hat{\gamma}^{N-1*};\gamma^N)$$

defined for each $\gamma^N \in \Gamma^N$. (This is well defined since $\Gamma^N \subset \hat{\Gamma}^N$.) The infimum of this quantity over Γ^N is definitely achieved (say, by $\gamma^{N*} \in \Gamma^N$), since Γ^N is a finite set. Furthermore, since $\bar{\Gamma}^N$ is comprised of all probability distributions on Γ^N,

$$\inf_{\bar{\Gamma}^N} F(\gamma^N) = \inf_{\Gamma^N} F(\gamma^N) = F(\gamma^{N*}).$$

We therefore have

$$F^* = F(\gamma^{N*}), \qquad (ii)$$

and also the inequality

$$\hat{J}^{N*} > F(\gamma^{N*})$$

in view of (i). But this is impossible since $\hat{J}^{N*} = \inf_{\hat{\Gamma}^N} F(\gamma^N)$ and $\Gamma^N \subset \hat{\Gamma}^N$, thus completing the proof of the proposition. $\diamond$

Even though MSNE exists in all finite-duration multi-act finite games, there is no guarantee that BSNE will exist. One can in fact construct games where a BSNE will not exist, but it is also possible to impose structures on a game such that BSNE will exist; for details see [1].

Given multi-act games which are identical in all respects except in the construction of the information sets, one can introduce a partial ordering among them depending on the *relative richness* of their strategy sets (induced by the information sets). One such ordering is introduced below, followed by a specific result that it leads to.

Definition 4 *Let I and II be two N-person multi-act nonzero-sum games with fixed orders of play, and with the property that at the time of her act each player has perfect information concerning the current level of play, that is, no information set contains nodes of the tree belonging to different levels of play. Further let Γ^i_{I} and Γ^i_{II} denote the strategy sets of $\mathbf{P}i$ in I and II, respectively. Then, I is* informationally inferior *to II if $\Gamma^i_{\mathrm{I}} \subseteq \Gamma^i_{\mathrm{II}}$ for all $i \in \mathcal{N}$, with strict inclusion for at least one i.* $\diamond$

Proposition 5 *Let* I *and* II *be two N-person multi-act nonzero-sum games as introduced in Definition 4, so that* I *is informationally inferior to* II. *Then,* *(i) any NE for* I *is also a NE for* II,

(ii) if $\{\gamma^1, \ldots, \gamma^N\}$ *is a NE for* II *so that* $\gamma^i \in \Gamma_I^i$ *for all* $i \in \mathcal{N}$*, then it is also a NE for* I.

Proof. We will prove this for the case when each player acts only once; extension to the more general multi-act game follows similar lines. (i) If $\gamma^* \in \Gamma_I^N\}$ constitutes a NE for I, then inequalities

$$J^i * := J^i(\gamma^{i*}, \gamma^{-i*}) \leq J^i(\gamma^i, \gamma^{-i*})$$

are satisfied for all $\gamma^i \in \Gamma_I^i$, $i \in \mathcal{N}$. But, since $\Gamma_I^i \subseteq \Gamma_{II}^i$, $i \in \mathcal{N}$, we clearly also have $\gamma^i \in \Gamma_{II}^i$, $i \in \mathcal{N}$. Now assume, to the contrary, that $\{\gamma^{1*}, \ldots, \gamma^{N*}\}$ is not a NE solution of II. Then, this implies that there exists at least one i (say, $i = N$, without any loss of generality) for which the corresponding inequality above is not satisfied for all $\gamma^i \in \Gamma_{II}^i$. In particular, there exists $\tilde{\gamma}^N \in \Gamma_{II}^N$ such that

$$J^{N*} > J^N(\gamma^{-N*}, \tilde{\gamma}^N). \qquad (i)$$

Now, the N-tuple of strategies $\{\gamma^{1*}, \ldots, \gamma^{N-1*}, \tilde{\gamma}^N\}$ leads to a unique path of action, and consequently to a unique outcome, in the single-act game II. Let us denote the information set of $\mathbf{P}N$, which is actually traversed by this path, by $\tilde{\eta}_{II}^N$, and the specific element (node) of $\tilde{\eta}_{II}^N$, intercepted by $\tilde{n}^N$. Let us further denote the information set of $\mathbf{P}N$ in the game I, which includes the node $\tilde{n}^N$, by $\tilde{\eta}_I^N$. Then, there exists at least one element in Γ_I^N (say, $\bar{\gamma}^N$) with the property $\bar{\gamma}^N(\tilde{\eta}_I^N) = \tilde{\gamma}^N(\tilde{\eta}_{II}^N)$. If this strategy replaces $\tilde{\gamma}^N$ on the RHS of inequality (i), the value of J^N clearly does not change, and hence we equivalently have

$$J^{N*} > J^N(\gamma^{1*}, \ldots, \gamma^{N-1*}, \bar{\gamma}^N).$$

But this inequality contradicts the initial hypothesis that the N-tuple $\{\gamma^{1*}, \ldots, \gamma^{N*}\}$ was in Nash equilibrium for the game I. This then completes the proof of part (i).

(ii) Part (ii) of the proposition can be proven analogously. $\diamond$

An important conclusion to be drawn from the result above is that dynamic games will generally admit a plethora of NE, because for a given game the NE of all inferior games will also constitute NE of the original game, and these are generally not even partially orderable—which arises due to informational richness. We call such occurrence of multiple NE *informational non-uniqueness*, which is a topic that will be revisited later.

5 Lecture 5: Refinements on Nash equilibrium

As we have seen in previous lectures, finite NZSGs will generally have multiple NE, in both pure and mixed strategies, and these equilibria are generally not interchangeable, with each one leading to a different set of equilibrium cost values or payoff values to the players, and they are not strictly ordered. We will also see in later lectures, when we discuss dynamic games that the presence of of multiple NE is more a rule rather than an exception, with the multiplicity arising in that case because of the *informational richness* of the underlying decision problem (in addition to the structure of the players' cost matrices). As a means of shrinking the set of Nash equilibria in a rational way, refinement schemes have been introduced in the literature; we discuss in this lecture some of those relevant to finite games.

To motivate the discussion, let us start with a two-player matrix game (A, B) where the players are minimizers and have identical cost matrices (which is what we called a *team problem* earlier).

$$A = B = \begin{array}{c} \\ U \\ D \\ \\ \end{array}\begin{array}{c} \mathbf{P}2 \\ \begin{array}{|c|c|} \hline 0 & 1 \\ \hline 1 & 1 \\ \hline \end{array} \\ L \quad R \end{array} \; \mathbf{P}1 \tag{16}$$

The game admits two pure-strategy Nash equilibria: (U, L) and (D, R). Note, however, that if we perturb the entries of the two matrices slightly, and independently:

$$A + \Delta A = \begin{array}{c} \mathbf{P}2 \\ \begin{array}{|c|c|} \hline \epsilon_{11}^1 & 1+\epsilon_{12}^1 \\ \hline 1+\epsilon_{21}^1 & 1+\epsilon_{22}^1 \\ \hline \end{array} \end{array} \; \mathbf{P}1; \quad B + \Delta B = \begin{array}{c} \mathbf{P}2 \\ \begin{array}{|c|c|} \hline \epsilon_{11}^2 & 1+\epsilon_{12}^2 \\ \hline 1+\epsilon_{21}^2 & 1+\epsilon_{22}^2 \\ \hline \end{array} \end{array} \; \mathbf{P}1;$$

where ϵ_{ij}^k, i, j, $k = 1, 2$, are infinitesimally small (positive or negative) numbers, then (U, L) will still retain its equilibrium property (as long as $|\epsilon_{ij}^k| < 1/2$), but (D, R) will not. More precisely, there will exist infinitely many perturbed versions of the original game for which (D, R) will not constitute a Nash equilibrium. Hence, in addition to admissibility,[9] (U, L) can be singled out in this case as the Nash solution that is *robust* to infinitesimal perturbations in the entries of the cost matrices.

Can such perturbations be induced naturally by some behavioral assumptions imposed on the players? The answer is yes, as we discuss next. Consider the scenario where a player who intends to play a particular pure strategy (out of a set of n possible alternatives) errs and plays with some

[9] A NE is said to be *admissible* if there is no other NE which yields better outcome for all players.

small probability one of the other $n-1$ alternatives. In the matrix game (16), for example, if both players err with equal (independent) probability $\epsilon > 0$, the resulting matrix game is (A_ϵ, B_ϵ), where

$$A_\epsilon = B_\epsilon = \begin{array}{c|c|c|c} & \multicolumn{2}{c}{\mathbf{P}2} & \\ U & \epsilon(2-\epsilon) & 1-\epsilon+\epsilon^2 & \\ D & 1-\epsilon+\epsilon^2 & 1-\epsilon^2 & \mathbf{P}1 \\ & L & R & \end{array}$$

Note that for all $\epsilon \in (0, 1/2)$ this matrix game admits the unique Nash equilibrium (U, L), with a cost pair of $(\epsilon(2-\epsilon), \epsilon(2-\epsilon))$, which converges to $(0,0)$ as $\epsilon \downarrow 0$, thus recovering one of the Nash cost pairs of the original game. A Nash equilibrium solution that can be recovered this way is known as a *perfect equilibrium*, which was first introduced in precise terms by *Selten (1975)*[38], in the context of N-player games in extensive form.[10] Given a game of perfect recall,[11] denoted $\mathcal{G}$, the idea is to generate a sequence of games, $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_k, \ldots$, a limiting equilibrium solution of which (in behavioral strategies, and as $k \to \infty$)[12] is an equilibrium solution of $\mathcal{G}$. If $\mathcal{G}_k$ is obtained from $\mathcal{G}$ by forcing the players at each information set to choose every possible alternative with positive probability (albeit small, for those alternatives that are not optimal), then the equilibrium solution(s) of $\mathcal{G}$ that are recovered as a result of the limiting procedure above is (are) called *perfect equilibrium* (*equilibria*).[13] *Selten (1975)* has shown that every finite game in extensive form with perfect recall (and as a special case in normal form) admits at least one perfect equilibrium, thus making this refinement scheme a legitimate one.

The procedure discussed above, which amounts to "completely" perturbing a game with multiple equilibria, is one way of obtaining perfect equilibria; yet another one, as introduced by *Myerson (1978)*[40], is to restrict the players to use completely mixed strategies (with some lower positive bound on the probabilities) at each information set. Again referring back to the matrix game (A, B) of (16), let the players' mixed strategies be restricted to the class

$$\hat{\gamma}^1 = \begin{cases} U & \text{w.p.} \quad y \\ D & \text{w.p.} \quad 1-y \end{cases} ; \quad \hat{\gamma}^2 = \begin{cases} L & \text{w.p.} \quad z \\ R & \text{w.p.} \quad 1-z \end{cases}$$

[10] Selten's construction and approach also apply to static games of the types discussed heretofore, where slight perturbations are made in the entries of the matrices, instead of at information sets.

[11] A game is one with *perfect recall* if all players recall their past moves—a concept that applies to games in extensive form.

[12] As introduced in the previous lecture, *behavioral strategy* is a mixed strategy for each information set of a player (in a dynamic game in extensive form). When the context is static games, it is identical with mixed strategy.

[13] This is also called "trembling hand equilibrium", as the process of erring at each information set is reminiscent of a "trembling hand" making unintended choices with small probability. Here, as $k \to \infty$, this probability of unintended plays converges to zero.

where $\epsilon \leq y \leq 1-\epsilon$, $\epsilon \leq z \leq 1-\epsilon$, for some (sufficiently small) positive ϵ. Over this class of strategies, the average cost functions of the players will be

$$\hat{J}^1 = \hat{J}^2 \quad = \quad -yz + 1,$$

which admits (assuming that $0 \leq \epsilon < \frac{1}{2}$) a unique Nash equilibrium:

$$p_\epsilon^{1^*} = \hat{\gamma}_\epsilon^{2^*} = \left\{ \begin{array}{lll} L & \text{w.p.} & 1-\epsilon \\ R & \text{w.p.} & \epsilon \end{array} \right. \quad ; \quad \hat{J}_\epsilon^{1^*} = \hat{J}_\epsilon^{2^*} = 1 - (1-\epsilon)^2.$$

Such a solution is called an *ϵ-perfect equilibrium (Myerson, 1978)*, which in the limit as $\epsilon \downarrow 0$ clearly yields the perfect Nash equilibrium obtained earlier. Myerson in fact proves, for N-person games in normal form, that every perfect equilibrium can be obtained as the limit of an appropriate ϵ-perfect equilibrium, with the converse statement also being true. More precisely, letting y^i denote a mixed strategy for Player i, and Y^i the simplex of probabilities, we have:

Proposition 6 *For an N-person finite game in normal form, a MSNE $\{y^{i^*} \in Y^i, i \in \mathcal{N}\}$ is a perfect equilibrium if, and only if, there exist some sequences $\{\epsilon_k\}_{k=1}^\infty$, $\{y_{\epsilon_k}^i \in \overset{\circ}{Y}{}^i, i \in \mathcal{N}\}_{k=1}^\infty$ such that*

i) $\epsilon_k > 0$ *and* $\lim_{k\to\infty} \epsilon_k = 0$

ii) $\{y_{\epsilon_k}^i, i \in \mathcal{N}\}$ *is an ϵ_k-perfect equilibrium*

iii) $\lim_{k\to\infty} y_{\epsilon_k}^i = y^{i^*}$, $i \in \mathcal{N}$.

Furthermore, a perfect equilibrium necessarily exists, and every perfect equilibrium is a NE. $\diamond$

Even though perfect equilibrium provides a refinement of Nash equilibrium with some appealing properties, it also carries some undesirable features as the following example of an identical cost matrix game (due to *Myerson (1978)*) exhibits:

$$A = B = \begin{array}{c} \\ U \\ M \\ D \\ \\ \end{array} \begin{array}{c} \mathbf{P}2 \\ \begin{array}{|c|c|c|} \hline 0 & 1 & 10 \\ \hline 1 & 1 & 8 \\ \hline 10 & 8 & 8 \\ \hline \end{array} \\ L \quad M \quad R \end{array} \ \mathbf{P}1. \tag{17}$$

Note that this is a matrix game derived from (16) by adding a completely dominated row and a completely dominated column. It now has three Nash equilibria: (U,L), (M,M), (D,R), the first two of which are perfect equilibria, while the last one is not. Hence, inclusion of completely dominated rows and columns could create additional perfect equilibria not present in the original game—a feature that is clearly not desirable. To remove this shortcoming of perfect equilibrium, *Myerson (1978)* introduced what is called *proper equilibria*, which corresponds to a particular construction of the sequence of strategies used in Proposition 6. *Proper equilibrium* is defined as in Proposition 6, with only the ϵ_k*-perfect equilibrium* in ii) replaced by the notion of ϵ_k*-proper equilibrium* to be introduced next. Toward this end, let $\bar{J}^i(j; y_\epsilon)$ denote the average cost to $\mathbf{P}i$ when he uses his j'th strategy (such as j'th column or row of the matrix) in the game and all the other players use their mixed strategies y_ϵ^k, $k \in \mathcal{N}$, $k \neq i$. Furthermore, let $y_\epsilon^{i,j}$ be the probability attached to his j'th strategy under the mixed strategy y_ϵ^i. Then, the N-tuple $\{y_\epsilon^i, i \in \mathcal{N}\}$ is said to be in ϵ*-proper equilibrium* if the strict inequality

$$\bar{J}^i(j; y_\epsilon) \quad > \quad \bar{J}^i(k; y_\epsilon)$$

implies that $y_\epsilon^{i,j} \leq \epsilon\, y_\epsilon^{i,k}$, this being so for every j, $k \in \mathbf{M}_i$,[14] and every $i \in \mathcal{N}$. In other words, an ϵ-proper equilibrium is one in which every player is giving his better responses much more probability weight than this worse responses (by a factor $1/\epsilon$), regardless of whether those "better" responses are "best" or not. *Myerson (1978)*[40] proves that such an equilibrium necessarily exists, that is:

Proposition 7 *Every finite N-player game in normal form admits at least one proper equilibrium. Furthermore, every proper equilibrium is a perfect equilibrium (but not vice versa).* $\diamond$

Remark 1 Note that in the matrix game (17) there is only one proper equilibrium, which is (U, L), the perfect equilibrium of (16). $\diamond$

Another undesirable feature of a perfect equilibrium is that it is very much dependent on whether the game is in extensive or normal form (whereas the Nash equilibrium property is form-independent). As it has been first observed by *Selten (1975)*[38], and further elaborated on by *van Damme (1984)*[41], a perfect equilibrium of the extensive form of a game need not be perfect in the normal form, and conversely a perfect equilibrium of the normal form need not be perfect in the extensive form. To remove this undesirable feature, *van Damme (1984)* (see also [42]) introduced

[14] $\mathbf{M}_i$ is the set of all pure strategies of Player 1, with corresponding labeling of positive integers.

the concept of *quasi-perfect* equilibria for games in extensive form, and has shown that a proper equilibrium of a normal form game induces a quasi-perfect equilibrium in every extensive form game having this normal form. Quasi-perfect equilibrium is defined as a behavioral strategy combination which prescribes at every information set a choice that is optimal against mistakes ("trembling hands") of the other players; its difference from perfect equilibrium is that here in the construction of perturbed matrices each player ascribes "trembling hand" behavior to all other players (with positive probability), but not to himself.

Other types of refinement have also been proposed in the literature, such as *sequential equilibria (Kreps and Wilson, 1982)*[43], and *strategic equilibria (Kohlberg and Mertens, 1986)*[44], which we do not further discuss here. None of these, however, are uniformly powerful, in the sense of shrinking the set of Nash equilibria to the smallest possible set. We will revisit the topic of "refinement on Nash equilibria" later in the context of infinite dynamic games and with emphasis placed on the issue of time consistency. In the context of infinite dynamic games, *Başar (1976)*[39] introduced stochastic perturbations in the system dynamics ("trembling dynamics") to eliminate multiplicity of Nash equilibria.

6 Lecture 6: Correlated equilibria

One important message that should have come through in the lectures so far is that NE is generally utterly *inefficient*, meaning that if the players had somehow correlated their choices of their actions, or better had collaborated in their selections, they all would be able to do better (than any of the NE) in terms of the outcome. In mathematical terms, NE is generally *not Pareto-efficient*, that is using the notation of Lectures 2 and 3, if $x^* \in X$ is a NE, it would be possible to find another N-tuple $\tilde{x} \in X$ such that $L_i(\tilde{x}) \leq L_i(x^*)$ for all $i \in \mathcal{N}$, with strict inequality for at least one i.

The question now is how to improve the costs (or payoffs) to players while still preserving the non-cooperative nature of the decision process. One way of doing this is through incentive strategies, or mechanism designs, which will be discussed in a future lecture. Another one is correlating the choices of the players through some signaling mechanisms, which leads to the notion of *correlated equilibrium* (CE) which is what I cover in this lecture.

First, let me start with an example scenario. Consider the situation faced by two drivers when they meet at an intersection (simultaneously). If both proceed, then that will lead to collision, and hence result in extreme cost to both drivers. If both yield, then they lose some time, which entails some cost, whereas if one yields and the other one proceeds, then the one that yields incurs some cost and the one that proceeds receives positive payoff. This can be modeled as a two-player 2×2 bi-matrix game, of the type below (where the first row and first column correspond to *Cross* (C), and the second row and the second column correspond to *Yield* (Y), and both players are minimizers):

$$\text{Intersection Game}: \quad \begin{pmatrix} (10,10) & (-5,0) \\ (0,-5) & (1,1) \end{pmatrix} \tag{18}$$

The game admits two pure-strategy NE, (C, Y) and (Y, C), and one MSNE, $\left((\frac{3}{8}, \frac{5}{8}), (\frac{3}{8}, \frac{5}{8})\right)$. The costs to the players (drivers) under the two pure-strategy NE are $(-5, 0)$ and $(0, -5)$, respectively, and under the MSNE (expected cost) $(\frac{5}{8}, \frac{5}{8})$. Note that both pure-strategy NE are uniformly better than the MSNE for both players, and therefore so is any convex combination of the two: $(-5\lambda, -5(1-\lambda))$, $\lambda \in [0, 1]$. Any pair in this convex combination can be achieved through correlated randomization, but the question is how such outcomes (or even better ones) can be attained through non-cooperative play. How can a randomization device be installed without any enforcement?

Of course, an obvious answer is to install a *traffic light*, which would function as a randomization device which with a certain probability would tell the players whether to cross or yield. Note that such a *signal* would help the players to correlate their actions. For example, if the traffic light tells with probability 0.55 cross to Player 1 (green light), and yield to Player 2 (red light); with probability 0.4 the other way around; and with the remaining probability (0.05) both players to yield, then the resulting expected cost pair is $(-2.7, -1.95)$. Note that these expected costs add up to -4.65, which is somewhat worse than any convex combination of the two pure-strategy NE (where the sum is -5), but it is a *safe* outcome and can only be achieved through correlation. Another noteworthy point is that actually the players do not have to obey the traffic light, but once it is there it is to their advantage to use it as a signal to correlate their moves; in that sense what this yields is an *equilibrium*, which is called a *correlated equilibrium* (CE). We now proceed with a precise definition of CE for bi-matrix games.

Consider a bi-matrix game (A, B), where the matrices are $m \times n$. Consider a randomization device which with probability p_{ij} signals $\mathbf{P}1$ to use row i and $\mathbf{P}2$ to use column j. This generates an $m \times n$ probability matrix

$$P = \{p_{ij}\},\ p_{ij} \geq 0,\ \sum_i \sum_j p_{ij} = 1$$

which we call a *correlated mixed strategy* (CMS). Such a strategy is in equilibrium if whenever the signal dictates $\mathbf{P}1$ to use row i, his expected cost cannot be lower by using some other action, *i.e.*,

$$\sum_{j=1}^{n} \left[a_{ij} p_{ij} / \sum_{\ell} p_{i\ell} \right] \leq \sum_{j=1}^{n} \left[a_{kj} p_{ij} / \sum_{\ell} p_{i\ell} \right] \quad \forall k \neq i$$

which can equivalently be written as

$$\sum_{j=1}^{n} (a_{ij} - a_{kj})\, p_{ij} \leq 0 \ \ \forall k \neq i\,. \tag{19}$$

Likewise for $\mathbf{P}2$, if j is the signal,

$$\sum_{i=1}^{m} (b_{ij} - b_{i\ell})\, p_{ij} \leq 0 \ \ \forall \ell \neq j\,. \tag{20}$$

Definition 5 *A* correlated equilibrium *(CE) for the bi-matrix game (A, B) is a correlated mixed strategy P that satisfies (19) for all $i = 1, \ldots, m$, and (20) for all $j = 1, \ldots, n$.*

Remark. If x is a mixed strategy for **P**1 and y is a mixed strategy for **P**2, then $P = xy'$ is a correlated mixed strategy for the bi-matrix game. Note that in this case $p_{ij} = x_i y_j$. But this is only a one-direction relationship, because not all correlated mixed strategies can be written this way. Hence, the set of all correlated mixed strategies for the bi-matrix game is larger than the set of all mixed strategy pairs. Furthermore, if (x^*, y^*) is a MSNE, then $P^* = x^* y^{*\prime}$ is a CE, which then implies that CE always exists. $\diamond$

Now, going back to the intersection game considered earlier, (19) and (20) lead to the set of inequalities:

$$\begin{aligned} 10p_{11} - 6p_{12} &\leq 0 \\ -10p_{21} + 6p_{22} &\leq 0 \\ 10p_{11} - 6p_{21} &\leq 0 \\ -10p_{12} + 6p_{22} &\leq 0 \end{aligned}$$

These have to be solved subject to the condition that elements of P are nonnegative and add up to 1. There are several solutions; below are a selected few.

(i) $p_{11} = 0,\ p_{12} = 0.55,\ p_{21} = 0.4,\ p_{22} = 0.05$ (this is what we had earlier)

(i) $p_{11} = 0,\ p_{12} = 0.5,\ p_{21} = 0.5,\ p_{22} = 0$ (this one is efficient)

(iii) $p_{11} = 0,\ p_{12} = 0.7,\ p_{21} = 0.3,\ p_{22} = 0$ (this one is also efficient)

(iv) $p_{11} = 0.1,\ p_{12} = 0.4,\ p_{21} = 0.3,\ p_{22} = 0.2$ (dangerous and inefficient)

(v) $p_{11} = \frac{9}{64},\ p_{12} = \frac{15}{64},\ p_{21} = \frac{15}{64},\ p_{22} = \frac{25}{64}$ (dangerous and highly inefficient, is the MSNE)

(vi) $p_{11} = 0,\ p_{12} = 1,\ p_{21} = 0,\ p_{22} = 0$ (corresponds to pure-strategy NE)

(vii) $p_{11} = 0,\ p_{12} = 0,\ p_{21} = 1,\ p_{22} = 0$ (the other pure-strategy NE)

Note that in this game, it is not possible to do uniformly better than any convex combination of the two pure-strategy NE. There are other examples, however, where one can. The following bi-matrix game which has that property is due to *Aumann*:

$$\begin{pmatrix} (0,4) & (5,5) \\ (1,1) & (4,0) \end{pmatrix} \tag{21}$$

We will label the two actions of Player 1 (row player) as U (up) and D (down), and those of Player 2 as L (left) and R (right); both are minimizers.

The game admits again two pure-strategy NE, (U, L) and (D, R), and one MSNE, $\left((\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2})\right)$. The costs to the players under the two pure-strategy NE are $(0, 4)$ and $(4, 0)$, respectively, and under the MSNE (expected cost) $(\frac{5}{2}, \frac{5}{2})$.

Now let the randomization device have three equally likely states: E, F, G. Player 1 observes E (perfectly), and Player 2 observes G (again perfectly). One can show that the following correlated mixed strategy pair for the players is a CE:

Player 1 plays U when she observes E, and D otherwise

Player 2 plays R when she observes G, and L otherwise.

Let us now compute the expected costs when the players use the correlated mixed strategy above. If E occurs, then the action pair is (U,L), leading to cost pair of $(0, 4)$; if G occurs, the action pair is (D,R), with resulting cost pair $(4, 0)$; and finally if F occurs, the action pair is (D, L), with corresponding cost $(1, 1)$. Since the three events each occur with equal probability $\frac{1}{3}$, the expected cost is $\left(\frac{5}{3}, \frac{5}{3}\right)$, which is better (for both players) than any convex combination of the two pure-strategy NE.

7 Lecture 7: NE of infinite/continuous-kernel games

7.1 Formulation, existence and uniqueness

We now go back to the general class of N-player games introduced through (1), with X_i being a finite-dimensional space (for example, m_i-dimensional Euclidean space, $\mathbf{R}^{m_i}$), for $i \in \mathcal{N}$; L_i a continuos function on the product space X, which of course is also finite-dimensional (for example, if $X_i = \mathbf{R}^{m_i}$, X can be viewed as $\mathbf{R}^m$, where $m := \sum_{i \in \mathcal{N}} m_i$); and the constraint set Ω a subset of X. This class of games is known as *continuous-kernel games with coupled constraints*, and of course if the constraints are not coupled, for example with each player having a separate constraint set $\Omega_i \subset X_i$, this would also be covered as a special case. Now, further assume that Ω is closed, bounded, and convex, and for each $i \in \mathcal{N}$, $L_i(x_i, x_{-i})$ is convex in $x_i \in X_i$ for every $x_{-i} \in \times_{j \neq i} X_j$. Then the basic result for such games is that they admit Nash equilibria in pure strategies (but the equilibria need not be unique), as stated in the theorem below, due to *Rosen (1965)*[18].

Theorem 3 *For the N-player nonzero-sum continuous-kernel game formulated above, with the constraint set Ω a closed, bounded, and convex subset of $\mathbf{R}^m$, and with $L_i(x_i, x_{-i})$ convex in x_i for each x_{-i}, and each $i \in \mathcal{N}$, there exists a Nash equilibrium in pure strategies.*

Proof. First introduce the function

$$L(x; v) := \sum_{i=1}^{N} L_i(x_{-i}, v_i),$$

and note that x is in NE if

$$L(x; x) \leq L(x; v), \quad \forall v \in X \ . \tag{22}$$

Further note that under the hypotheses of the theorem, the function $L(x; v)$ is jointly continuous in x and v and is convex in v for every fixed x, with $(x, v) \in X \times X$. Introduce the reaction set for the game :

$$Tx = \{v \in X : L(x; v) \leq L(x; w) \quad \forall w \in X\}$$

which, by the continuity and convexity properties of L cited above, is an upper semicontinuous (usc) mapping[15] that maps each point x in the convex and compact set X into a closed convex

[15] Let X and Y be two normed spaces (such as R^m and R^n), and f be a set-valued mapping of X into 2^Y (the set of all subsets of Y). Then, f is *upper semicontinuous* at a point $x_o \in X$, if for any sequence $\{x_{(i)}\}$ converging to x_o, and any sequence $\{y_{(i)}\} \in f(x_{(i)})\}$ converging to y_o, we have $y_o \in f(x_o)$. f is usc if it is usc at each point of X.

subset of X. Then, by the Kakutani fixed point theorem,[16] there exists a point $x^* \in X$ such that $x^* \in Tx^*$, or equivalently that it minimizes $L(x^*; v)$ over $v \in X$. Such a point indeed constitutes a NE, because if it does not then this would imply that for some $i \in \mathcal{N}$ there would be a $\bar{x}_i \in X_i(x^*_{-i})$ such that

$$L_i(x^*_{-i}, \bar{x}_i) < L_i(x^*) \ ,$$

which would in turn imply (by adding $L_j(x^*)$ to both sides and summing over $j \in \mathcal{N}, j \neq i$) the strict inequality

$$L(x^*; \bar{x}) < L(x^*; x^*), \quad \bar{x} \triangleq (x^*_{-i}, \bar{x}_i) \ ,$$

contradicting the initial hypothesis that x^* minimizes $L(x^*; v)$ over $v \in X$. $\diamond$

Remark. If the constraint sets are decoupled, and $L_i(x_i, x_{-i})$ is strictly convex in $x_i \in \Omega_i$, then there is an alternative proof for Theorem 3 which uses Brouwer's fixed-point theorem. Under the given hypotheses, it follows from Weirstrass theorem and strict convexity that the minimization problem

$$\min_{x_i \in \Omega_i} L_i(x_i, x_{-i})$$

admits a unique solution for each x_{-i}, this being so for each $i \in \mathcal{N}$, that is, there exists a unique map $T_i : \Omega_i \to \Omega_{-i}$,[17] such that the solution to the minimization problem is

$$x_i = T_i(x_{-i}) \,, \quad i \in \mathcal{N} \tag{23}$$

Furthermore, T_i is continuous on Ω_{-i}. Clearly, every pure-strategy NE has to provide a solution to (23), and *vice versa*. Stacking these maps, there exists a corresponding continuous map $T : \Omega \to \Omega$, whose components are the T_i's, and (23) is equivalent to $x = T(x)$, which is a fixed-point equation. Since T is a continuous mapping of Ω into itself, and Ω is a closed and bounded subset of a finite-dimensional space (and thus compact), by Brouwer's fixed-point theorem T has a fixed point, and hence a NE exists. $\diamond$

For the special class of 2-person ZSGs structured the same way as the NZSG of Theorem 3, a similar result clearly holds (as a special case), implying the existence of a SPE (in pure strategies).

[16] This fixed point theorem says that if S is a compact subset of R^n, and f is an usc function which assigns to each $x \in S$ a closed and convex subset of S, then there exists $x \in S$ such that $x \in f(x)$.

[17] T_i is known as the reaction function (or response function) of Player i to other players' actions.

Note that in this case the single objective function ($L \equiv L_1 \equiv -L_2$) to be minimized by Player 1 and maximized by Player 2, is convex in x_1 and concave in x_2, in view of which such zero-sum games are known as *convex-concave games*. Even though convex-concave games could admit multiple saddle-point solutions, they are ordered interchangeable, and the values of the games are unique (which is not the case for multiple Nash equilibria in genuine nonzero-sum games, as we have also seen earlier). Now, if the convexity-concavity is replaced by strict convexity-concavity (for ZSGs), then the result can be sharpened as below, which however has no a counterpart for Nash equilibria in genuine nonzero-sum games.

Theorem 4 *For a two-person zero-sum game on closed, bounded and convex finite-dimensional action sets $\Omega_1 \times \Omega_2$, defined by the continuous kernel $L(x_1, x_2)$, let $L(x_1, x_2)$ be strictly convex in x_1 for each $x_2 \in \Omega_2$ and strictly concave in x_2 for each $x_1 \in \Omega_1$. Then, the game admits a unique pure-strategy SPE.*

Proof. Existence of SPE is a direct consequence of Theorem 3. Furthermore, by strict convexity and strict concavity, there can be no SPE outside the class of pure strategies. Hence only uniqueness within the class of pure strategies remains to be proven, which, however, follows readily from the interchangeability property of multiple SPs, in view of strict convexity/concavity. $\diamond$

If the structural assumptions of Theorem 3 do not hold, then a pure-strategy Nash equilibrium may not exist, but there may exist one in mixed strategies. Mixed strategy (MS) for a player (say, Player i) is a probability distribution on that player's action set, which we take to be a closed and bounded subset, Ω_i, of $X_i = \mathbf{R}^{m_i}$, and denote a MS of Player i by p_i, and the set of all probability distribution on Ω_i by $\mathcal{P}_i$. NE, then, is defined by the N-tuple of inequalities (4), using the expected values of L_i's given mixed strategies of all the players, which we denote by J_i as in Lecture 2. The following theorem now states the basic result on existence of MSNE in continuous-kernel games.

Theorem 5 *For the N-player continuous-kernel NZSG formulated above, with the constrained action set Ω_i for Player i a closed and bounded subset of $\mathbf{R}^{m_i}$, and with $L_i(x_i, x_{-i})$ continuous on $\Omega = \Omega_1 \times \cdots \times \Omega_N$, for each $i \in \mathcal{N}$, there exists a MSNE, $(p_1^*, \ldots, p_N^*)$, satisfying (4).*

Proof. A proof of this theorem can be found in *Owen (1974)*[19]. The underlying idea is to make the kernels L_i discrete so as to obtain an N-person matrix game that suitably approximates the original game in the sense that a MSNE of the latter (which exists by Nash's theorem) is arbitrarily close to a mixed equilibrium solution of the former. Compactness of the action spaces ensures that a limit to the sequence of solutions obtained for approximating finite matrix games exists. $\diamond$

As a special case of Theorem 5 we now have:

Corollary 7.1 *Every continuous-kernel* 2*-player ZSG with compact action spaces has a MSSPE.* $\diamond$

I conclude this section with an example of a zero-sum game whose cost functional is continuous but not convex-concave, and which has a MSSPE.

Example. Consider the two-person ZSG on the square $[0,2] \times [0,2]$ characterized by the kernel $L = (x_1 - x_2)^2$. It can readily be verified that the *upper value* is 1 and the *lower value* is 0, and hence a pure-strategy SPEdoes not exist. In the extended class of mixed strategies, however, a candidate SP solution is

$$x_1^* = 1 \quad \text{w.p. } 1; \quad x_2^* = \begin{cases} 0 & \text{w.p. } \frac{1}{2} \\ 2 & \text{w.p. } \frac{1}{2}. \end{cases}$$

It can readily be verified that this pair of strategies indeed provides a MSSPE. $\diamond$

Remark. As in the case of finite (matrix) games, the existence of a pure-strategy NE does not preclude existence of also a genuine MSNE,[18] and all such (multiple) NE are generally non-interchangeable, unless the game is a ZSG or is strategically equivalent to one. $\diamond$

7.2 Stability and computation

We have seen in the previous section that when the cost functions of the players are strictly convex in a continuous-kernel NZSG, then the NE is completely characterized by the solution of a fixed-point equation, namely (23). Since solutions of fixed-point equations can be obtained recursively (under some condition), this brings up the possibility of computing the NE recursively, using the iteration

$$x_i(k+1) = T_i(x_{-i}(k)), \quad k = 0, 1, \ldots, \quad i \in \mathcal{N} \tag{24}$$

where k stands for times of updates by the players. Note that this admits an *on-line computation* interpretation for the underlying game, where each player needs to know only the most recent

[18] The qualifier *genuine* is used here to stress the point that mixed strategies in this statement are not pure strategies (even though pure strategies are indeed special types of mixed strategies, with all probability weight concentrated on one point).

actions of the other players (and not their cost functions) and her own reaction function T_i (for which only the individual cost function of the player is needed). Hence, this recursion entails a distributed computation with little information on the parameters of the game, and lumping all players' actions together, and writing (24) as

$$x(k+1) = T(x(k)), \quad k = 0, 1, \ldots,$$

we note that the sequence generated converges for all possible initial choices, $x(0) = x_0$, if T is a contraction from X into itself.[19] As an immediate byproduct, we also have *uniqueness* of the NE.

The recursion above is not the only way one can generate a sequence converging to its fixed point. But before I discuss other possibilities, let me make a digression and talk about a classification of NE based on such recursions, provided by the notion of "stability" of the solution(s) of the fixed point equation. This discussion will then immediately lead to other possible recursions (for $N > 2$). For the sake of simplicity in the initial discussion, let us consider the two-player case (because in this case there is only one type of recursion for the fixed-point equation, as will be clear later). Given a NE (and assuming that the players are at the NE point), consider the following sequence of moves: (i) One of the players (say $\mathbf{P}1$) deviates from his corresponding equilibrium strategy, (ii) $\mathbf{P}2$ observes this and minimizes his cost function in view of the new strategy of $\mathbf{P}1$, (iii) $\mathbf{P}1$ now optimally reacts to that (by minimizing his cost function), (iv) $\mathbf{P}2$ optimally reacts to that optimum reaction, etc. Now, if this infinite sequence of moves converges back to the original NE solution, and this being so regardless of the nature of the initial deviation of $\mathbf{P}1$, we say that the NE is *stable*. If convergence is valid only under small initial deviations, then we say that the NE is *locally stable*. Otherwise, the NE is said to be *unstable*. A NZSG can of course admit more than one locally stable equilibrium solution, but a stable NE solution has to be unique.

The notion of *stability*, as introduced above for two-person games, brings in a refinement to the concept of NE, which finds natural extensions to the N-player case. Essentially, we have to require that the equilibrium be "restorable" under any rational re-adjustment scheme when there is a deviation from it by any player. For $N > 2$ this will depend on the specific scheme adopted, which brings us to the following formal definition of a stable Nash equilibrium.

[19]This follows from Banach's contraction mapping theorem. If T maps a normed space X into itself, it is a contraction if there exists $\alpha \in [0, 1)$ such that $\|T(x) - T(y)\| \leq \alpha \|x - y\|$, $\forall x, y \in X$.

Definition 6 *A NE $x_i^*, i \in \mathcal{N}$, is (globally) stable with respect to an adjustment scheme $\mathcal{S}$ if it can be obtained as the limit of the iteration:*

$$x_i^* = \lim_{k\to\infty} {x_i}^{(k)}, \tag{25}$$

$$ {x_i}^{(k+1)} = \arg\min_{x_i\in\Omega_i} L_i(x_{-i}^{(\mathcal{S}_k)}, x_i), \quad {x_i}^{(0)} \in \Omega_i, \quad i \in \mathcal{N} \tag{26}$$

where the superscript $\mathcal{S}_k$ indicates that the precise choice of $x_{-i}^{(\mathcal{S}_k)}$ depends on the re-adjustment scheme selected. $\diamond$

One possibility for the scheme above is $x_{-i}^{(\mathcal{S}_k)} = x_{-i}^{(k)}$, which corresponds to the situation where the players update (re-adjust) their actions simultaneously, in response to the most recently determined actions of the other players. Yet another possibility is

$$x_{-i}^{(\mathcal{S}_k)} = \left({x_1}^{(k+1)}, \ldots, {x_{i-1}}^{(k+1)}, {x_{i+1}}^{(k)}, \ldots, {x_N}^{(k)}\right)$$

where the players update in an predetermined (in this case numerical) order. A third possibility is

$$x_{-i}^{(\mathcal{S}_k)} = \left({x_1}^{m^i_{1,k}}, \ldots, {x_{i-1}}^{(m^i_{i-1,k})}, {x_{i+1}}^{(m^i_{i+1,k})}, \ldots, {x_N}^{(m^i_{N,k})}\right)$$

where $m^i_{j,k}$ is an integer-valued random variable, satisfying the bounds:

$$\max(0, k-d) \quad \leq \quad m^i_{j,k} \leq k+1, \quad j \neq i, j \in \mathcal{N}, i \in \mathcal{N};$$

which corresponds to a situation where $\mathbf{P}i$ receives action update information from $\mathbf{P}j$ at random times, with the delay not exceeding d time units.

Clearly, if the iteration of Definition 6 converges under any one of the re-adjustment schemes above (or any other re-adjustment scheme where a player receives update information from every other player infinitely often), then the NE is *unique*. Every unique NE, however, is not necessarily *stable*, nor is a NE that is stable with respect to a particular re-adjustment scheme is necessarily stable with respect to some other scheme. Hence *stability* is generally given with some qualification (such as "stable with respect to scheme $\mathcal{S}$" or "with respect to a given class of schemes"), except when $N = 2$, in which case all schemes (with at most a finite delay in the transmission of update information) lead to the same condition of stability, as one then has the simplified recursions

$${x_i}^{(r_{k+1,i})} = \tilde{T}_i({x_i}^{(r_{k,i})}), \quad k = 0, 1, \ldots; \quad i = 1, 2$$

where $r_{1,i}$, $r_{2,i}$, $r_{3,i}$, $\ldots$ denote the time instants when $\mathbf{P}i$ receives new action update information from $\mathbf{P}j$, $j \neq i$, i, $j = 1, 2$.

8 Lecture 8: Hierarchical finite games and Stackelberg equilibria

The Nash equilibrium solution concept that we have heretofore studied in these *Notes* provides a reasonable noncooperative equilibrium solution for nonzero-sum games when the roles of the players are symmetric, that is to say, when no single player dominates the decision process. However, there are yet other types of noncooperative decision problems wherein one of the players has the ability to enforce his strategy on the other player(s), and for such decision problems one has to introduce a hierarchical equilibrium solution concept. Following the original work of *H. von Stackelberg (1934)*[20], the player who holds the powerful position in such a decision problem is called the *leader*, and the other players who react (rationally) to the leader's decision (strategy) are called the *followers*. There are, of course, cases of multiple levels of hierarchy in decision making, with many leaders and followers; but for purposes of brevity and clarity in exposition I will first confine the discussion here to hierarchical decision problems which incorporate two players (decision makers) — one leader and one follower.

8.1 Stackelberg equilibria in pure strategies

To set the stage to introduce the hierarchical (Stackelberg) equilibrium solution concept, let us first consider the bimatrix game (A, B) displayed (under our standard convention) as

$$A = \begin{array}{c} \\ L \\ M \\ R \\ \\ \end{array} \begin{array}{c} \mathbf{P}2 \\ \begin{array}{|c|c|c|} \hline 0^{S_1} & 2 & 3/2^{S_2} \\ \hline 1 & 1^N & 3 \\ \hline -1 & 2 & 2 \\ \hline \end{array} \\ \begin{array}{ccc} L & M & R \end{array} \end{array} \mathbf{P}1, \; B = \begin{array}{c} \\ L \\ M \\ R \\ \\ \end{array} \begin{array}{c} \mathbf{P}2 \\ \begin{array}{|c|c|c|} \hline -1^{S_1} & 1 & -2/3^{S_2} \\ \hline 2 & 0^N & 1 \\ \hline 0 & 1 & -1/2 \\ \hline \end{array} \\ \begin{array}{ccc} L & M & R \end{array} \end{array} \mathbf{P}1 \qquad (27)$$

This bimatrix game clearly admits a unique NE in pure strategies, which is $\{M, M\}$, with the corresponding outcome being $(1, 0)$. Let us now stipulate that the roles of the players are not symmetric and **P**1 can enforce his strategy on **P**2. Then, before he announces his strategy, **P**1 has to take into account possible responses of **P**2 (the follower), and in view of this, he has to decide on the strategy that is most favorable to him. For the decision problem whose possible cost pairs are given as entries of A and B, above, let us now work out the reasoning that **P**1 (the leader) will have to go through. If **P**1 chooses L, then **P**2 has a unique response (that minimizes his cost) which is L, thereby yielding a cost of 0 to **P**1. If the leader chooses M, **P**2's response is again unique (which is M), with the corresponding cost incurred to **P**1 being 1. Finally, if he picks R,

P2's unique response is also R, and the cost to **P**1 is 2. Since the lowest of these costs is the first one, it readily follows that L is the most reasonable choice for the leader (**P**1) in this hierarchical decision problem. We then say that L is the Stackelberg strategy of the leader (**P**1) in this game, and the pair $\{L, L\}$ is the Stackelberg solution with **P**1 as the leader. Furthermore, the cost pair $(0, -1)$ is the Stackelberg (equilibrium) outcome of the game with **P**1 as the leader. It should be noted that this outcome is actually more favorable for both players than the unique Nash outcome — this latter feature, however, is not a rule in such games. If, for example, **P**2 is the leader in the bimatrix game (27), then the unique Stackelberg solution is $\{L, R\}$ with the corresponding outcome being $(3/2, -2/3)$ which is clearly not favorable for **P**1 (the follower) when compared with his unique Nash cost. For **P**2 (the leader), however, the Stackelberg outcome is again better than his Nash outcome.

The Stackelberg equilibrium (SE) solution concept introduced above within the context of the bimatrix game (27) is applicable to all two-person finite games in normal form, but provided that they exhibit one feature which was inherent to the bimatrix game (27) and was used implicitly in the derivation: *the follower's response to every strategy of the leader should be unique.* If this requirement is not satisfied, then there is ambiguity in the possible responses of the follower and thereby in the possible attainable cost levels of the leader. As an explicit example to demonstrate such a decision situation, consider the bimatrix game

$$A = \begin{array}{c} \\ L \\ R \\ \\ \end{array} \begin{array}{c} \mathbf{P}2 \\ \begin{array}{|c|c|c|} \hline 0 & 1 & 3 \\ \hline 2 & 2 & -1 \\ \hline \end{array} \\ L \quad M \quad R \end{array} \ \mathbf{P}1, \qquad B = \begin{array}{c} \\ L \\ R \\ \\ \end{array} \begin{array}{c} \mathbf{P}2 \\ \begin{array}{|c|c|c|} \hline 0 & 0 & 1 \\ \hline -1 & 0 & -1 \\ \hline \end{array} \\ L \quad M \quad R \end{array} \ \mathbf{P}1 \tag{28}$$

and with **P**1 acting as the leader. Here, if **P**1 chooses (and announces) L, **P**2 has two optimal responses L and M, whereas if **P**1 picks R, **P**2 again has two optimal responses, L and R. Since this multiplicity of optimal responses for the follower results in a multiplicity of cost levels for the leader for each one of his strategies, the Stackelberg solution concept introduced earlier cannot directly be applied here. However, this ambiguity in the attainable cost levels of the leader can be resolved if we stipulate that the leader's attitude is toward securing his possible losses against the choices of the follower within the class of his optimal responses, rather than toward taking risks. Then, under such a mode of play, **P**1's secured cost level corresponding to his strategy L would be 1, and the one corresponding to R would be 2. Hence, we declare $\gamma^{1*} = L$ as the unique Stackelberg strategy of $\mathbf{P}1$ in the bimatrix game of (28), when he acts as the leader.[20] The corresponding Stackelberg cost for $\mathbf{P}1$ (the leader) is $J^{1*} = 1$. It should be noted that, in the actual play of the game, $\mathbf{P}1$ could actually end up with a lower cost level, depending on whether the follower chooses his optimal response $\gamma^2 = L$ or the optimal response $\gamma^2 = M$. Consequently, the outcome of the game could be either $(1, 0)$ or $(0, 0)$, and hence we cannot talk about a unique Stackelberg equilibrium outcome of the bimatrix game (28) with $\mathbf{P}1$ acting as the leader. Before concluding our discussion on this example, we finally note that the admissible Nash outcome of the bimatrix game (28) is $(-1, -1)$ which is more favorable for both players than the possible Stackelberg outcomes given above.

We now provide a precise definition for the Stackelberg solution concept introduced above within the context of two bimatrix games, so as to encompass all two-person finite games of the single-act and multi-act type which do not incorporate any chance moves. For such a game, let Γ^1 and Γ^2 again denote the pure-strategy spaces of $\mathbf{P}1$ and $\mathbf{P}2$, respectively, and $J^i(\gamma^1, \gamma^2)$ denote the cost incurred to $\mathbf{P}i$ corresponding to a strategy pair $\{\gamma^1 \in \Gamma^1, \gamma^2 \in \Gamma^2\}$. Then, we have

Definition 7 *In a two-person finite game, the set $R^2(\gamma^1) \subset \Gamma^2$ defined for each $\gamma^1 \in \Gamma^1$ by*

$$R^2(\gamma^1) = \{\xi \in \Gamma^2 : J^2(\gamma^1, \xi) \leq J^2(\gamma^1, \gamma^2), \quad \forall \gamma^2 \in \Gamma^2\} \tag{29}$$

is the optimal response (rational reaction) set *of $\mathbf{P}2$ to the strategy $\gamma^1 \in \Gamma^1$ of $\mathbf{P}1$.* $\diamond$

Definition 8 *In a two-person finite game with $\mathbf{P}1$ as the leader, a strategy $\gamma^1 \in \Gamma^1$ is called a* Stackelberg equilibrium strategy *for the leader, if*

$$\max_{\gamma^2 \in R^2(\gamma^{1*})} J^1(\gamma^{1*}, \gamma^2) = \min_{\gamma^1 \in \Gamma^1} \max_{\gamma^2 \in R^2(\gamma^1)} J^1(\gamma^1, \gamma^2) \triangleq J^{1*}. \tag{30}$$

The quantity J^{1*} is the Stackelberg cost *of the leader. If, instead, $\mathbf{P}2$ is the leader, the same definition applies with only the superscripts 1 and 2 interchanged.* $\diamond$

Theorem 6 *Every two-person finite game admits a Stackelberg strategy for the leader.*

Proof. Since Γ^1 and Γ^2 are finite sets, and $R^2(\gamma^1)$ is a subset of Γ^2 for each $\gamma^1 \in \Gamma^1$, the result readily follows from (30). $\diamond$

[20] Of course, the "strategy" here could also be viewed as an "action" if what we have is a static game, but since we are dealing with normal forms here (which could have an underlying extensive form description) we will use the term "strategy" throughout, to be denoted by γ^i for $\mathbf{P}i$, and the cost to $\mathbf{P}i$ will be denoted by J^i.

Remark 2 The Stackelberg strategy for the leader does not necessarily have to be unique. But nonuniqueness of the equilibrium strategy does not create any problem here (as it did in the case of Nash equilibria), since the Stackelberg cost for the leader is unique. $\diamond$

Remark 3 If $R^2(\gamma^1)$ is a singleton for each $\gamma^1 \in \Gamma^1$, then there exists a mapping $T^2 : \Gamma^1 \to \Gamma^2$ such that $\gamma^2 \in R^2(\gamma^1)$ implies $\gamma^2 = T^2\gamma^1$. This corresponds to the case when the optimal response of the follower (which is T^2) is unique for every strategy of the leader, and it leads to the following simplified version of (30) in Definition 8:

$$J^1(\gamma^{1*}, T^2\gamma^{1*}) = \min_{\gamma^1 \in \Gamma^1} J^1(\gamma^1, T^2\gamma^1) \triangleq J^{1*}. \tag{31}$$

Here J^{1*} is no longer only a secured equilibrium cost level for the leader (**P**1), but it is the cost level that is actually attained. $\diamond$

From the follower's point of view, the equilibrium strategy in a Stackelberg game is any optimal response to the announced Stackelberg strategy of the leader. More precisely,

Definition 9 *Let* $\gamma^{1*} \in \Gamma^1$ *be a Stackelberg strategy for the leader* (**P**1). *Then, any element* $\gamma^{2*} \in R^2(\gamma^{1*})$ *is an* optimal strategy *for the follower* (**P**2) *that is* in equilibrium *with* γ^{1*}. *The pair* $\{\gamma^{1*}, \gamma^{2*}\}$ *is a* Stackelberg solution *for the game with* **P**1 *as the leader, and the cost pair* $(J^1(\gamma^{1*}, \gamma^{2*}), J^2(\gamma^{1*}, \gamma^{2*}))$ *is the corresponding* Stackelberg equilibrium outcome. $\diamond$

Remark 4 In the preceding definition, the cost level $J^1(\gamma^{1*}, \gamma^{2*})$ could in fact be lower than the Stackelberg cost J^{1*} — a feature which has already been observed within the context of the bimatrix game (28). However, if $R^2(\gamma^{1*})$ is a singleton, then these two cost levels have to coincide. $\diamond$

For a given two-person finite game, let J^{1*} again denote the Stackelberg cost of the leader (**P**1), and J^1_N denote any Nash equilibrium cost for the same player. We have already seen within the context of the bimatrix game (28) that J^{1*} is not necessarily lower than J^1_N, in particular, when the optimal response of the follower is not unique. The following proposition now provides one sufficient condition under which the leader never does worse in a "Stackelberg game" than in a "Nash game".

Proposition 8 *For a given two-person finite game, let J^{1*} and J^1_N be as defined before. If $R^2(\gamma^1)$ is a singleton for each $\gamma^1 \in \Gamma^1$, then*

$$J^{1*} \leq J^1_N.$$

Proof. Under the hypothesis of the proposition, assume to the contrary that there exists a Nash equilibrium solution $\{\gamma^{1\circ} \in \Gamma_1, \gamma^{2\circ} \in \Gamma^2\}$ whose corresponding cost to **P**1 is lower than J^{1*}, *i.e.*

$$J^{1*} > J^1(\gamma^{1\circ}, \gamma^{2\circ}). \qquad (i)$$

Since $R^2(\gamma^1)$ is a singleton, let $T^2 : \Gamma^2 \to \Gamma^1$ be the unique mapping introduced in Remark 3. Then, clearly, $\gamma^{2\circ} = T^2\gamma^{1\circ}$, and if this is used in (i), together with the RHS of (31), we obtain

$$\min_{\gamma^1 \in \Gamma^1} J^1(\gamma^1, T^2\gamma^1) = J^{1*} > J^1(\gamma^{1\circ}, T^2\gamma^{1\circ}),$$

which is a contradiction. $\diamond$

Remark 5 One might be tempted to think that if a nonzero-sum game admits a unique Nash equilibrium solution and a unique Stackelberg strategy (γ^{1*}) for the leader, and further if $R^2(\gamma^{1*})$ is a singleton, then the inequality of Proposition 8 still should hold. This, however, is not true as the following bimatrix game demonstrates

$$A = \begin{array}{c} \\ L \\ R \\ \\ \end{array} \begin{array}{c} \mathbf{P}2 \\ \begin{array}{|c|c|} \hline 0^{S_1} & 1 \\ \hline -1^{N} & 2 \\ \hline \end{array} \\ L \quad\; R \end{array} \mathbf{P}1, \qquad B = \begin{array}{c} \\ L \\ R \\ \\ \end{array} \begin{array}{c} \mathbf{P}2 \\ \begin{array}{|c|c|} \hline 0^{S_1} & 2 \\ \hline 1^{N} & 1 \\ \hline \end{array} \\ L \quad\; R \end{array} \mathbf{P}1.$$

Here, there exists a unique Nash equilibrium solution, as indicated, and a unique Stackelberg strategy $\gamma^{1*} = L$ for the leader (**P**1). Furthermore, the follower's optimal response to $\gamma^{1*} = L$ is unique (which is $\gamma^2 = L$). However, $0 = J^{1*} > J^1_N = -1$. This counterexample indicates that the sufficient condition of Proposition 8 cannot be relaxed any further in any satisfactory way. $\diamond$

8.2 Stackelberg equilibria in mixed and behavioral strategies

The motivation behind introducing mixed strategies in the investigation of saddle-point equilibria and Nash equilibria was that such equilibria do not always exist in pure strategies, whereas within the enlarged class of mixed strategies one can ensure existence of noncooperative equilibria. In the case of the Stackelberg solution of two-person finite games, however, an equilibrium always exists (cf. Theorem 6), and thus, at the outset, there seems to be no need to introduce mixed strategies. Besides, since the leader dictates his strategy on the follower, in a Stackelberg game, it might at first seem to be unreasonable to imagine that the leader would ever employ a mixed-strategy. Such an argument, however, is not always valid, and there are cases when the leader can actually do better (in the average sense) with a proper mixed strategy, than the best he can do within the class of pure strategies. As an illustration of such a possibility, consider the bi-matrix game (A, B) displayed below:

$$A = \begin{array}{c} \\ L \\ R \\ \\ \end{array} \begin{array}{c} \mathbf{P}2 \\ \begin{array}{|c|c|} \hline 1 & 0 \\ \hline 0 & 1 \\ \hline \end{array} \\ L \quad R \end{array} \ \mathbf{P}1, \qquad B = \begin{array}{c} \\ L \\ R \\ \\ \end{array} \begin{array}{c} \mathbf{P}2 \\ \begin{array}{|c|c|} \hline 1/2 & 1 \\ \hline 1 & 1/2 \\ \hline \end{array} \\ L \quad\quad R \end{array} \ \mathbf{P}1. \tag{32}$$

If $\mathbf{P}1$ acts as the leader, then the game admits two pure-strategy Stackelberg equilibrium solutions, which are $\{L, L\}$, and $\{R, R\}$, the Stackelberg outcome in each case being $(1, 1/2)$. However, if the leader ($\mathbf{P}1$) adopts the mixed strategy which is to pick L and R with equal probability $1/2$, then the average cost incurred to $\mathbf{P}1$ will be equal to $1/2$, quite independent of the follower's (pure or mixed) strategy. This value $\bar{J}^1 = 1/2$ is clearly lower than the leader's Stackelberg cost in pure strategies, which can further be shown to be the unique Stackelberg cost of the leader in mixed strategies, since any deviation from $(1/2, 1/2)$ for the leader results in higher values for $\bar{J}^1$, by taking into account the optimal responses of the follower.

The preceding result then establishes the significance of mixed strategies in the investigation of Stackelberg equilibria of two-person nonzero-sum games, and demonstrates the possibility that a proper mixed-strategy Stackelberg solution could lead to a lower cost level for the leader than the Stackelberg cost level in pure strategies. To introduce the concept of mixed-strategy Stackelberg equilibrium in mathematical terms, we take the two-person nonzero-sum finite game to be in normal form (without any loss of generality) and associate with it a bi-matrix game (A, B). Abiding by the earlier notation and terminology, let Y and Z denote the mixed-strategy spaces of $\mathbf{P}1$ and $\mathbf{P}2$, respectively, with their typical elements denoted by y and z. Then, we have:

Definition 10 *For a bi-matrix game (A, B), the set*

$$\bar{R}^2(y) = \{z^\circ \in Z : y'Bz^\circ \leq y'Bz, \forall z \in Z\} \tag{33}$$

is the optimal response (rational reaction) set of $\mathbf{P}2$ *in mixed strategies to the mixed strategy* $y \in Y$ *of* $\mathbf{P}1$. $\diamond$

Definition 11 *In a bi-matrix game* (A, B) *with* $\mathbf{P}1$ *acting as the leader, a mixed strategy* $y^* \in Y$ *is called a* mixed Stackelberg equilibrium strategy *for the leader if*

$$\max_{z \in \bar{R}^2(y^*)} y^{*\prime} A z = \inf_{y \in Y} \max_{z \in \bar{R}^2(y)} y' A z \triangleq \bar{J}^{1*}. \tag{34}$$

The quantity $\bar{J}^{1*}$ *is the* Stackelberg cost *of the leader in mixed strategies.* $\diamond$

It should be noted that the "maximum" in (34) always exists since, for each $y \in Y$, $y'Az$ is continuous in z, and $R^2(y)$ is a closed and bounded subset of Z (which is a finite dimensional simplex). Hence, $\bar{J}^{1*}$ is a well-defined quantity. The "infimum" in (34), however, cannot always be replaced by a "minimum", unless the problem admits a mixed Stackelberg equilibrium strategy for the leader. The following example now demonstrates the possibility that a two-person finite game might not admit a mixed-strategy Stackelberg strategy even though $\bar{J}^{1*} < J^{1*}$.

Example. Consider the following modified version of the bi-matrix game of (32):

$$A = \begin{array}{c|c|c|} \multicolumn{3}{c}{\mathbf{P}2} \\ \cline{2-3} L & 1 & 0 \\ \cline{2-3} R & 0 & 1 \\ \cline{2-3} \multicolumn{1}{c}{} & \multicolumn{1}{c}{L} & \multicolumn{1}{c}{R} \end{array} \ \mathbf{P}1, \qquad B = \begin{array}{c|c|c|} \multicolumn{3}{c}{\mathbf{P}2} \\ \cline{2-3} L & 1/2 & 1 \\ \cline{2-3} R & 1 & 1/3 \\ \cline{2-3} \multicolumn{1}{c}{} & \multicolumn{1}{c}{L} & \multicolumn{1}{c}{R} \end{array} \ \mathbf{P}1.$$

With $\mathbf{P}1$ as the leader, this bi-matrix game also admits two pure-strategy Stackelberg equilibria, which are $\{L, L\}$ and $\{R, R\}$, the Stackelberg cost for the leader being $J^{1*} = 1$. Now, let the leader adopt the mixed strategy $y = (y_1, (1 - y_1))'$, under which $\bar{J}^2$ is

$$\bar{J}^2(y, z) = y'Bz = \left(-\frac{7}{6}y_1 + \frac{2}{3}\right) z_1 + \frac{2}{3}y_1 + \frac{1}{3},$$

where $z = (z_1, (1 - z_1))'$ denotes any mixed strategy of $\mathbf{P}2$. Then, the mixed-strategy optimal response set of $\mathbf{P}2$ can readily be determined as

$$\bar{R}^2(y) = \begin{cases} \{z = (1, 0)\} & \text{if } y_1 > 4/7 \\ \{z = (0, 1)\} & \text{if } y_1 < 4/7 \\ Z & \text{if } y_1 = 4/7. \end{cases}$$

Hence, for $y_1 > 4/7$, the follower chooses "column 1" with probability 1, and this leads to an average cost of $\bar{J}^1 = y_1$ for $\mathbf{P}1$. For $y_1 < 4/7$, on the other hand, $\mathbf{P}2$ chooses "column 2" with probability

1, which leads to an average cost level of $\bar{J}^1 = (1 - y_1)$ for **P**1. Then, clearly, the leader will prefer to stay in this latter region; in fact, if he employs the mixed strategy $y = (4/7 - \epsilon, 3/7 + \epsilon)'$ where $\epsilon > 0$ is sufficiently small, his realized average cost will be $\bar{J}^1 = 3/7 + \epsilon$, since then **P**2 will respond with the unique pure-strategy $\gamma^2 = R$. Since $\epsilon > 0$ can be taken as small as possible, we arrive at the conclusion that $\bar{J}^{1*} = \frac{3}{7} < 1 = J^{1*}$. In spite of this fact, the leader does not have a mixed Stackelberg strategy since, for the only candidate $y^\circ = (4/7, 3/7)$, $\bar{R}^2(y^\circ) = Z$, and therefore $\max_{z \in \bar{R}^2(y^\circ)} {y^\circ}' Az = 4/7$ which is higher than $\bar{J}^{1*}$. $\diamond$

The preceding example thus substantiates the possibility that a mixed Stackelberg strategy might not exist for the leader, but he can still do better than his pure Stackelberg cost J^{1*} by employing some sub-optimal mixed strategy (such as the one $y = (4/7 - \epsilon, 3/7 + \epsilon)'$ in the example, for sufficiently small $\epsilon > 0$). In fact, whenever $\bar{J}^{1*} < J^{1*}$, there will always exist such an approximating mixed strategy for the leader. If $\bar{J}^{1*} = J^{1*}$, however, it is, of course, reasonable to employ the pure Stackelberg strategy which always exists by Theorem 6. The following proposition now verifies that $\bar{J}^{1*} < J^{1*}$ and $\bar{J}^{1*} = J^{1*}$ are the only two possible relations we can have between $\bar{J}^{1*}$ and J^{1*}; in other words, the inequality $\bar{J}^{1*} > J^{1*}$ never holds.

Proposition 9 *For every two-person finite game, we have*

$$\bar{J}^{1*} \leq J^{1*}. \tag{35}$$

Proof. Let Y_0 denote the subset of Y consisting of all one-point distributions. Analogously, define Z_0 as comprised of one-point distributions in Z. Note that Y_0 is equivalent to Γ^1, and Z_0 is equivalent to Γ^2. Then, for each $y \in Y_0$

$$\min_{z \in Z} y' Bz = \min_{z \in Z_0} y' Bz$$

since any minimizing solution in Z can be replaced by an element of Z_0. This further implies that, for each $y \in Y_0$, elements of $\bar{R}^2(y)$ are probability distributions on $R^2(y)$, where the latter set is defined by (33) with Z replaced by Z_0. Now, since $Y_0 \subset Y$,

$$\bar{J}^{1*} = \min_{y \in Y} \max_{z \in \bar{R}^2(y)} y' Az \leq \min_{y \in Y_0} \max_{z \in \bar{R}^2(y)} y' Az,$$

and further, because of the cited relation between $\bar{R}^2(\cdot)$ and $R^2(\cdot)$, the latter quantity is equal to

$$\min_{y \in Y_0} \max_{z \in R^2(y)} y' Az$$

which, by definition, is J^{1*}, since $R^2(y)$ is equivalent to the pure-strategy optimal response set of the follower, as defined by (29). Hence, $\bar{J}^{1*} \leq J^{1*}$. $\diamond$

Computation of a mixed-strategy Stackelberg equilibrium (whenever it exists) is not as straightforward as in the case of pure-strategy equilibria, since the spaces Y and Z are not finite. The standard technique is first to determine the minimizing solution(s) of

$$\min_{z \in Z} y' B z$$

as functions of $y \in Y$. This will lead to a decomposition of Y into subsets (regions), on each of which a reaction set for the follower is defined. (Note that in the analysis of the previous example, Y has been decomposed into three regions.) Then, one has to minimize $y'Az$ over $y \in Y$, subject to the constraints imposed by these reaction sets, and under the stipulation that the same quantity is maximized on these reaction sets whenever they are not singletons. This brute-force approach also provides approximating strategies for the leader, whenever a mixed Stackelberg solution does not exist, together with the value of $\bar{J}^{1*}$.

If the two-person finite game under consideration is a dynamic game in extensive form, then it is more reasonable to restrict attention to behavioral strategies. Stackelberg equilibrium within the class of behavioural strategies can be introduced as in Definitions 10 and 11, by replacing the mixed strategy sets with the behavioral strategy sets. Hence, using the earlier terminology and notation, we have the following counterparts of Definitions. 10 and 11, in behavioral strategies:

Definition 12 *Given a two-person finite dynamic game with behavioral-strategy sets* $(\hat{\Gamma}^1, \hat{\Gamma}^2)$ *and average cost functions* $(\hat{J}^1, \hat{J}^2)$*, the set*

$$\widehat{R}^2(\hat{\gamma}^1) = \{\hat{\gamma}^{2\circ} \in \hat{\Gamma}^2 : \hat{J}^2(\hat{\gamma}^1, \hat{\gamma}^{2\circ}) \leq \hat{J}^2(\hat{\gamma}^1, \hat{\gamma}^2), \forall \hat{\gamma}^2 \in \hat{\Gamma}^2\}, \tag{36}$$

is the optimal response (rational reaction) set *of* $\mathbf{P}2$ *in* behavioral strategies *to the behavioral strategy* $\hat{\gamma}^1 \in \hat{\Gamma}^1$ *of* $\mathbf{P}1$*.* $\diamond$

Definition 13 *In a two-person finite dynamic game with* $\mathbf{P}1$ *acting as the leader, a behavioral strategy* $\hat{\gamma}^{1*} \in \hat{\Gamma}^1$ *is called a* behavioral Stackelberg equilibrium *strategy for the leader if*

$$\sup_{\hat{\gamma}^2 \in \bar{R}^2(\hat{\gamma}^{1*})} \hat{J}^1(\hat{\gamma}^{1*}, \hat{\gamma}^2) = \inf_{\hat{\gamma}^1 \in \hat{\Gamma}^1} \sup_{\hat{\gamma}^2 \in \widehat{R}^2(\hat{\gamma}^1)} \hat{J}^1(\hat{\gamma}^1, \hat{\gamma}^2) \triangleq \hat{J}^{1*}. \tag{37}$$

The quantity $\hat{J}^{1*}$ *is the* Stackelberg cost *of the leader in* behavioral strategies*.* $\diamond$

9 Lecture 9: Continuous-kernel games and Stackelberg equilibria

This lecture is on the Stackelberg solution of static NZSGs when the number of alternatives available to each player is not a finite set and the cost functions are described by continuous kernels. For the sake of simplicity and clarity in exposition, the focus will be on two-person static games. A variety of possible extensions of the Stackelberg solution concept to N-person static games with different levels of hierarchy can be found in [1].

Here, I use a slightly different notation than in Lecture 7, with $u^i \in U^i$ denoting the action variable of $\mathbf{P}i$ (instead of $x_i \in \Omega_i$), where his action set U^i is assumed to be a subset of an appropriate metric space (such as X_i). The cost function J^i of $\mathbf{P}i$ is defined as a continuous function on the product space $U^1 \times U^2$. Then we can give to following general definition of a Stackelberg equilibrium solution, (SES) which is the counterpart of Definition 8 for infinite games.

Definition 14 *In a two-person game, with* $\mathbf{P}1$ *as the leader, a strategy* $u^{1^*} \in U^1$ *is called a* Stackelberg equilibrium strategy *for the leader if*

$$J^{1^*} \triangleq \sup_{u^2 \in R^2(u^{1^*})} J^1(u^{1^*}, u^2) \leq \sup_{u^2 \in R^2(u^1)} J^1(u^1, u^2) \tag{38}$$

for all $u^1 \in U^1$*. Here,* $R^2(u^1)$ *is the rational reaction set of the follower as introduced in (29).* $\diamond$

Remark 6 If $R^2(u^1)$ is a singleton for each $u^1 \in U^1$, in other words, if it is described completely by a reaction curve $T_2 : U^1 \to U^2$, then inequality (38) in the above definition can be replaced by

$$J^{1^*} \triangleq J^1(u^{1^*}, T_2(u^{1^*})) \leq J^1(u^1, T_2(u^1)) \tag{39}$$

for all $u^1 \in U^1$. $\diamond$

If a SES exists for the leader, then the LHS of inequality (38) is known as the *leader's Stackelberg cost*, and is denoted by J^{1^*}. A more general definition for J^{1^*} is, in fact,

$$J^{1^*} = \inf_{u^1 \in U^1} \sup_{u^2 \in R^2(u^1)} J^1(u^1, u^2), \tag{40}$$

which also covers the case when a Stackelberg equilibrium strategy does not exist. It follows from this definition that the Stackelberg cost of the leader is a well-defined quantity, and that there will always exist a sequence of strategies for the leader which will insure him a cost value arbitrarily close to J^{1^*}. This observation brings us to the following definition of ϵ Stackelberg strategies.

Definition 15 *Let $\epsilon > 0$ be a given number. Then, a strategy $u_\epsilon^{1^*} \in U^1$ is called an ϵ* Stackelberg strategy *for the leader* (**P**1) *if*

$$\sup_{u^2 \in R^2(u_\epsilon^{1^*})} J^1(u_\epsilon^{1^*}, u^2) \leq J^{1^*} + \epsilon.$$

$\diamond$

The following two properties of ϵ Stackelberg strategies now readily follow.

Property 1. In a two-person game, let J^{1^*} be a finite number. Then, given an arbitrary $\epsilon > 0$, an ϵ Stackelberg strategy necessarily exists. $\diamond$

Property 2. Let $\{u_{\epsilon_i}^{1^*}\}$ be a given sequence of ϵ Stackelberg strategies in U^1, and with $\epsilon_i > \epsilon_j$ for $i < j$ and $\lim_{j\to\infty} \epsilon_j = 0$. Then, if there exists a convergent subsequence $\{u_{\epsilon_{ik}}^{1^*}\}$ in U^1 with its limit denoted as u^{1^*}, and further if $\sup_{u^2 \in R^2(u^1)} J^1(u^1, u^2)$ is a continuous function of u^1 in an open neighborhood of $u^{1^*} \in U^1$, u^{1^*} is a Stackelberg strategy for **P**1. $\diamond$

The equilibrium strategy of the follower, in a Stackelberg game, would be any strategy that constitutes an optimal response to the one adopted (and announced) by the leader. Mathematically speaking, if u^{1^*} (respectively, $u_\epsilon^{1^*}$) is adopted by the leader, then any $u^2 \in R^2(u^1)$, (respectively, $u^2 \in R^2(u_\epsilon^{1^*})$ will be referred to as an *optimal strategy* for the follower that is in *equilibrium* with the Stackelberg (respectively, ϵ Stackelberg) strategy of the leader. This pair is referred to as a *Stackelberg* (respectively, ϵ Stackelberg) solution of the two-person game with **P**1 as the leader (see Definition 9). The following theorem now provides a set of sufficient conditions for two-person NZSGs to admit a SES.

Theorem 7 *Let U^1 and U^2 be compact metric spaces, and J^i be continuous on $U^1 \times U^2$, $i = 1, 2$. Further let there exist a finite family of continuous mappings $l^{(i)} : U^1 \to U^2$, indexed by a parameter $i \in I \triangleq \{1, \ldots, M\}$, so that $R^2(u^1) = \{u^2 \in U^2 : u^2 = l^{(i)}(u^1), i \in I\}$. Then, the two-person nonzero-sum static game admits a Stackelberg equilibrium solution.*

Proof. It follows from the hypothesis of the theorem that J^{1^*}, as defined by (40), is finite. Hence, by Property 1, a sequence of Stackelberg strategies exists for the leader, and it admits a convergent subsequence whose limit lies in U^1, due to compactness of U^1. Now, since $R^2(\cdot)$ can be constructed from a finite family of continuous mappings (by hypothesis),

$$\sup_{u^2 \in R^2(u^1)} J^1(u^1, u^2) = \max_{i \in I} J^1(u^1, l^{(i)}(u^1)),$$

and the latter function is continuous on U^1. Then, the result follows from Property 2. $\diamond$

Remark 7 The assumption of Theorem 7 concerning the structure of $R^2(\cdot)$ imposes some severe restrictions on J^2; but such an assumption is inevitable as the following example demonstrates. Take $U^1 = U^2 = [0, 1]$, $J^1 = -u^1u^2$ and $J^2 = (u^1 - \frac{1}{2})u^2$. Here, $R^2(\cdot)$ is determined by a mapping $l(\cdot)$ which is continuous on the half-open intervals $[0, \frac{1}{2}), (\frac{1}{2}, 1]$, but is multivalued at $u^1 = \frac{1}{2}$. The Stackelberg cost of the leader is clearly $J^{1^*} = -\frac{1}{2}$, but a Stackelberg strategy does not exist because of the "infinitely multivalued" nature of l. $\diamond$

When $R^2(u^1)$ is a singleton for every $u^1 \in U^1$, the hypothesis of Theorem 7 can definitely be made less restrictive. One such set of conditions is provided in the following corollary to Theorem 7 under which there exists a unique l which is continuous.

Corollary 9.1 *Every two-person nonzero-sum continuous-kernel game on the square, for which $J^2(u^1, \cdot)$ is strictly convex for all $u^1 \in U^1$ and* **P1** *acts as the leader, admits a SES.* $\diamond$

It should be noted that the SES for a two-person game exists under a set of sufficiency conditions which are much weaker than those required for existence of Nash equilibria. It should further be noted, however, that the statement of Theorem 7 does not also rule out the existence of a mixed-strategy Stackelberg solution which might provide the leader with a lower average cost. We have already observed occurrence of such a phenomenon within the context of matrix games, in the previous lecture, and we now investigate to what extent such a result could remain valid in continuous-kernel games.

If mixed strategies are also allowed, then permissible strategies for **P**i will be probability measures μ^i on the space U^i. Let us denote the collection of all such probability measures for **P**i by M^i. Then, the quantity replacing J^i will be the average cost function

$$\bar{J}^i(\mu^1, \mu^2) = \int_{U^1} \int_{U^2} J^i(u^1, u^2)\, \mathrm{d}\mu^1(u^1)\, \mathrm{d}\mu^2(u^2), \tag{41}$$

and the reaction set R^2 will be replaced by

$$\bar{R}^2(\mu^1) \triangleq \{\widehat{\mu}^2 \in M^2 : \bar{J}^2(\mu^1, \widehat{\mu}^2) \leq \bar{J}^2(\mu^1, \mu^2), \forall \mu^2 \in M^2\}. \tag{42}$$

Hence, we have:

Definition 16 *In a two-person game with* $\mathbf{P}1$ *as the leader, a mixed strategy* $\mu^{1^*} \in M^1$ *is called a* mixed Stackelberg equilibrium strategy *for the leader if*

$$\bar{J}^{1^*} \triangleq \sup_{\mu^2 \in \bar{R}^2(\mu^{1^*})} \bar{J}^1(\mu^{1^*}, \mu^2) \leq \sup_{\mu^2 \in \bar{R}^2(\mu^1)} \bar{J}^1(\mu^1, \mu^2)$$

for all $\mu^1 \in M^1$*, where* $\bar{J}^{1^*}$ *is known as the* average Stackelberg cost *of the leader in* mixed strategies. $\diamond$

Proposition 10

$$\bar{J}^{1^*} \quad \leq \quad J^{1^*} \tag{43}$$

Proof. Since M^i also includes all one-point measures, we have (by an abuse of notation) $U^i \subset M^i$. Then, for each $u^1 \in U^1$, considered as an element of M^i,

$$\begin{aligned} \inf_{\mu^2 \in M^2} \bar{J}^2(u^1, \mu^2) &\equiv \inf_{\mu^2 \in M^2} \int_{U^2} J^2(u^1, u^2)\, \mathrm{d}\mu^2(u^2) \\ &= \inf_{u^2 \in U^2} J^2(u^1, u^2), \end{aligned}$$

where the last equality follows since any infimizing sequence in M^2 can be replaced by a subsequence of one-point measures. This implies that, for one point measures in $M^1, \bar{R}^2(\mu^1)$ coincides with the set of all probability measures defined on $R^2(u^1)$. Now, since $M^1 \supset U^1$,

$$\bar{J}^{1^*} = \inf_{M^1} \sup_{\bar{R}^2(\mu^1)} \bar{J}^1(\mu^1, \mu^2) \leq \inf_{U^1} \sup_{\bar{R}^2(\mu^1)} \bar{J}^1(\mu^1, \mu^2),$$

and because of the cited relation between $\bar{R}^2(\mu^1)$ and $R^2(u^1)$, the last expression can be written as

$$\inf_{U^1} \sup_{R^2(u^1)} J^1(u^1, u^2) = J^{1^*}.$$

$\diamond$

We now show, by a counter-example, that, even under the hypothesis of Thm 7, it is possible to have strict inequality in (43).

(Counter-) Example. Consider a two-person continuous-kernel game with $U^1 = U^2 = [0, 1]$, and with cost functions

$$J^1 = \epsilon(u^1)^2 + u^1\sqrt{u^2} - u^2; \quad J^2 \quad = \quad (u^2 - (u^1)^2)^2,$$

where $\epsilon > 0$ is a sufficiently small parameter. The unique Stackelberg solution of this game, in pure strategies, is $u^{1^*} = 0$, $u^{2*} = (u^1)^2$, and the Stackelberg cost for the leader is $J^{1^*} = 0$. We now show that the leader can actually do better by employing a mixed strategy.

First note that the follower's unique reaction to a mixed strategy of the leader is $u^2 = E[(u^1)^2]$ which, when substituted into $\bar{J}^1$, yields the expression

$$\bar{J}^1 = \epsilon E[(u^1)^2] + E[u^1]\sqrt{\{E[(u^1)^2]\}} - E[(u^1)^2].$$

Now, if the leader uses the uniform probability distribution on $[0,1]$, his average cost becomes

$$\bar{J}^1 = \frac{\epsilon - 1}{3} + \frac{1}{2}\sqrt{\frac{1}{3}}$$

which clearly indicates that, for ϵ sufficiently small, $\bar{J}^{1^*} < 0 = J^{1^*}$. $\diamond$

The preceding example has displayed the fact that even Stackelberg games with *strictly convex* cost functionals may fail to admit only pure-strategy solutions, and the mixed Stackelberg solution may in fact be .[21] However, if we further restrict the cost structure to be quadratic, then only pure-strategy Stackelberg equilibria will exist.

Proposition 11 *Consider the two-person nonzero-sum game with $U^1 = \mathbf{R}^{m_1}$, $U^2 = \mathbf{R}^{m_2}$, and*

$$J^i = \frac{1}{2}u^{i\prime}R^i_{ii}u^i + u^{1\prime}R^i_{ij}u^j + \frac{1}{2}u^{j}.'R^i_{jj}u^j + u^{i\prime}r^i_i + u^{j\prime}r^i_j;\ i,j = 1,2, i \neq j,$$

where $R^i_{ii} > 0$, R^i_{ii}, R^i_{ij}, R^i_{jj} are appropriate dimensional matrices, and r^i_i, r^i_j are appropriate dimensional vectors. This "quadratic" game can only admit a pure-strategy Stackelberg solution, with either **P**1 *or* **P**2 *as the leader.*

Proof. Without any loss of generality take **P**1 as the leader, and assume, to the contrary, that the game admits a mixed-strategy Stackelberg solution, and denote the leader's optimal mixed strategy by μ^{1^*}. Furthermore, denote the expectation operation under μ^{1^*} by $E[\cdot]$. If the leader announces this mixed strategy, then the follower's reaction is unique and is given by

$$u^2 = -{R^2_{22}}^{-1}R^2_{21}E[u^1] - {R^2_{22}}^{-1}r^2_2.$$

[21] In retrospect, this should not be surprising since for the special case of zero-sum games (without pure-strategy saddle points) we have already seen that the minimizer could further decrease his guaranteed expected cost by playing a mixed strategy; here however it holds even if $J^1 \not\equiv -J^2$.

By substituting this in $\bar{J}^{1^*} = E[J^1]$, we obtain

$$\bar{J}^{1^*} = \frac{1}{2}E[u^{1\prime} R^1_{11} u^1] + E[u^1]' K E[u^1] + E[u^1]' k + c,$$

where

$$\begin{aligned}
K &\triangleq \frac{1}{2} R^{2}_{21}{}' R^{2}_{22}{}^{-1} R^{2}_{11} R^{2}_{22}{-1} R^{2}_{21} - R^{1}_{12} R^{2}_{22}{}^{-1} R^{2}_{21} \\
k &\triangleq r^1_1 - R^{2}_{21}{}' R^{2}_{22}{}^{-1} r^1_2 - R^{2}_{22}{}^{-1} r^2_2 + R^{2}_{21}{}' R^{2}_{22}{}^{-1} R^{1}_{22} R^{2}_{22}{}^{-1} r^2_2 \\
c &\triangleq \frac{1}{2} r^{2\prime}_2 R^{2}_{22}{}^{-1} R^{1}_{22} R^{2}_{22}{}^{-1} r^2_2 - r^{2\prime}_2 R^{2}_{22}{}^{-1} r^1_2.
\end{aligned}$$

Now, applying the Cauchy-Schwarz inequality on the first term of $\bar{J}^{1^*}$, we further obtain the bound

$$\bar{J}^{1^*} \geq \frac{1}{2}E[u^1]' R^1_{11} E[u^1] + E[u^1]' K E[u^1] + E[u^1]' k + c \qquad (i)$$

which depends only on the mean value of u^1. Hence

$$J^{1^*} = \inf_{U^1} \left\{ \frac{1}{2} u^{1\prime} R^1_{11} u^1 + u^{1\prime} K u^1 + u^{1\prime} k + c \right\} \leq \bar{J}^{1^*}.$$

This implies that enlargement of the strategy space of the leader, so as also to include mixed strategics, does not yield him any better performance. In fact, since $E[u^{1\prime} u^1] > E[u^1]' E[u^1]$, whenever the probability distribution is not one-point, it follows that the inequality in (i) is actually strict for the case of a proper mixed strategy. This then implies that, outside the class of pure-strategies, there can be no Stackelberg solution. ◇

10 Lecture 10: Quadratic games: Deterministic and stochastic

This lecture presents explicit expressions for the Nash, saddle-point, and Stackelberg equilibrium solutions of static nonzero-sum games in which the cost functions of the players are quadratic in the decision variables — the so-called *quadratic games*. The action (strategy) spaces will be taken as appropriate dimensional Euclidean spaces, but the results are also equally valid (under the right interpretation) when the strategy spaces are taken as infinite-dimensional Hilbert spaces. In that case, the Euclidean inner products will have to be replaced by the inner product of the underlying Hilbert space, and the positive definiteness requirements on some of the matrices will have to be replaced by *strong positive definiteness* of the corresponding self-adjoint operators. This lecture will also include some discussion on iterative algorithms for the computation of Nash equilibria in the quadratic case.

10.1 Deterministic games

A general quadratic cost function for $\mathbf{P}i$, which is strictly convex in his action variable, can be written as

$$J^i = \frac{1}{2}\sum_{j=1}^{N}\sum_{k=1}^{N} u^{j\prime} R^i_{jk} u^k + \sum_{j=1}^{N} r^{i\prime}_j u^j + c^i, \tag{44}$$

where $u^j \in U^j = \mathbf{R}^{m_j}$ is the m_j-dimensional action variable of $\mathbf{P}j$, R^i_{jk} is an $(m_j \times m_k)$-dimensional matrix with $R^i_{ii} > 0$, r^i_j is an m_j-dimensional vector and c^i is a constant. Without loss of generality, we may assume that, for $j \neq k$, $R^i_{jk} = {R^i_{kj}}'$, since if this were not the case, the corresponding two quadratic terms could be written as

$$u^{j\prime} R^i_{jk} u^k + u^{k\prime} R^i_{kj} u^j = u^{j\prime}\left(\frac{R^i_{jk} + {R^i_{kj}}'}{2}\right) u^k + u^{k\prime}\left(\frac{R^i_{jk} + {R^i_{kj}}'}{2}\right) u^j \tag{45}$$

and redefining R^i_{jk} as $(R^i_{jk} + {R^i_{jk}}')/2$, a symmetric matrix could be obtained. By an analogous argument, we may take R^i_{jj} to be symmetric, without any loss of generality.

Quadratic cost functions are of particular interest in game theory, firstly because they constitute second-order approximation to other types of nonlinear cost functions, and secondly because games with quadratic cost or payoff functions are analytically tractable, admitting, in general, closed-form equilibrium solutions which provide insight into the properties and features of the equilibrium solution concept under consideration.

To determine the NE solution in strictly convex quadratic games, we differentiate J^i with respect to u^i $(i \in \mathcal{N})$, set the resulting expressions equal to zero, and solve the set of equations thus obtained. This set of equations, which also provides a sufficient condition because of strict convexity, is

$$R^i_{ii} u^i + \sum_{j \neq i} R^i_{ij} u^j + r^i_i = 0 \quad (i \in \mathcal{N}), \tag{46}$$

which can be written in compact form as

$$Ru = -r \tag{47}$$

where

$$R \triangleq \begin{bmatrix} R^1_{11} & R^1_{12} & \cdots & R^1_{1N} \\ R^2_{12} & R^2_{22} & \cdots & R^2_{2N} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ R^N_{1N} & R^N_{2N} & \cdots & R^N_{NN} \end{bmatrix} \tag{48}$$

$$u' \triangleq (u^1, u^2, \ldots, u^N). \tag{49}$$

$$r' \triangleq (r^1_1, r^2_2, \ldots, r^N_N). \tag{50}$$

This then leads to the following result.

Proposition 12 *The quadratic N-player nonzero-sum static game defined by the cost functions (44) and with $R^i_{ii} > 0$, admits a Nash equilibrium solution if, and only if, (47) admits a solution, say u^*; this Nash solution is unique if the matrix R defined by (48) is invertible, in which case it is given by*

$$u^* = -R^{-1} r. \tag{51}$$

$\diamond$

Remark 8 Since each player's cost function is strictly convex and continuous in his action variable, quadratic nonzero-sum games of the type discussed above cannot admit a Nash equilibrium solution in mixed strategies. Hence, in strictly convex quadratic games, the equilibrium analysis can be confined to the class of pure strategies. $\diamond$

We now investigate the stability properties of the unique Nash solution of quadratic games, where the notion of stability was introduced in Lecture 7. Taking $N = 2$, and directly specializing recursion (24) to the quadratic case (with the obvious change in notation, and in a sequential update mode), we arrive at the following iteration :

$$u^{1(k+1)} = C_1 u^{2(k)} + d_1, \quad u^{2(k+1)} = C_2 u^{1(k+1)} + d_2, \quad k = 0, 1, \ldots \tag{52}$$

with an arbitrary starting choice $u^{2(0)}$, where

$$C_i = -(R^i_{ii})^{-1} R^i_{ij}, \quad d_i = -(R^i_{ii})^{-1} r^i_i, \quad j \neq i, \ i, j = 1, 2.$$

This iteration corresponds to the sequential (Gauss-Seidel) update scheme where $\mathbf{P}1$ responds to the most recent past action of $\mathbf{P}2$, whereas $\mathbf{P}2$ responds to the current action of $\mathbf{P}1$. The alternative to this is the parallel (Jacobi) update scheme where (52) is replaced by[22]

$$u^{1(k+1)} = C_1 u^{2(k)} + d_1, \quad u^{2(k+1)} = C_2 u^{1(k)} + d_2, \quad k = 0, 1, \ldots \tag{53}$$

starting with arbitrary initial choices $(u^{1(0)}, u^{2(0)})$ Then, the question of stability of the Nash solution (51), with $N = 2$, reduces to the question of stability of the fixed point of either (52) or (53). Note that, apart from a relabeling of indices, stability of these two iterations is equivalent to the stability of the single iteration:

$$u^{1(k+1)} = C_1 C_2 u^{1(k)} + C_1 d_2 + d_1.$$

Since this is a linear difference equation, a necessary and sufficient condition for it to converge (to the actual Nash strategy of $\mathbf{P}1$) is that the eigenvalues of the matrix $C_1 C_2$, or equivalently those of $C_2 C_1$ should be in the unit circle, *i.e.*

$$\rho(C_1 C_2) \equiv \rho(C_2 C_1) < 1 \tag{54}$$

where $\rho(A)$ is the spectral radius of the matrix A.

Note that the condition of stability is considerably more stringent than the condition of existence of a unique Nash equilibrium, which is

$$\det(I - C_1 C_2) \neq 0. \tag{55}$$

[22] This one corresponds to (24).

The question we address now is whether, in the framework of Gauss-Seidel or Jacobi iterations, this gap between (54) and could be (55) shrunk or even totally eliminated, by allowing players to incorporate memory into the iterations. While doing this, it would be desirable for the players to need to know as little as possible regarding the reaction functions of each other (note that no such information is necessary in the Gauss-Seidel or Jacobi iterations given above).

To study this issue, consider the Gauss-Seidel iteration (52), but with a one-step memory for (only) **P**1. Then, the "relaxed" algorithm will be (using the simpler notation $u^{1^{(k)}} = u_k$, $u^{2^{(k)}} = v_k$):

$$\left.\begin{aligned} u_{k+1} &= C_1 v_k + d_1 + A(u_k - C_1 v_k - d_1) \\ v_{k+1} &= C_2 u_{k+1} + d_2 \end{aligned}\right\} \tag{56}$$

where A is a gain matrix, yet to be chosen. Substituting the second (for v_k) into the first, we obtain the single iteration

$$u_{k+1} = [C + A(I - C)]u_k + (I - A)[d_1 + C_1 d_2]$$

where

$$C \triangleq C_1 C_2.$$

By choosing

$$A = -C(I - C)^{-1} \tag{57}$$

where the required inverse exists because of (55), we obtain a finite-step convergence, assuming that the true value of C_2 is known to **P**1. If the true value of C_2 is not known, but a nominal value is given in a neighborhood of which the true value lies, the scheme (56) along with the choice (57) for the nominal value, still leads to convergence (but not in a finite number of steps) provided that the neighborhood is sufficiently small (see *Başar, 1987*) [22].

Now, if the original scheme is instead the parallel (Jacobi) scheme, then a one-step memory for **P**1 will not be sufficient to obtain a finite-step convergence result as above. In this case we replace (56) by

$$\left.\begin{aligned} u_{k+1} &= C_1 v_k + d_1 + B(u_{k-1} - C_1 v_k - d_2) \\ v_{k+1} &= C_2 u_k + d_2 \end{aligned}\right\} \tag{58}$$

where B is another gain matrix. Note that here **P**1 uses, in the computation of u_{k+1}, not u_k but rather u_{k-1}. Now, substituting for v_k from the second into the first equation of (58), we arrive at the iteration

$$u_{k+1} = [C + B(I - C)]u_{k-1} + (I - B)[d_1 + C_1 d_2],$$

which again shows finite-step convergence, with B chosen as

$$B = -C(I - C)^{-1}. \tag{59}$$

Again, there is a certain neighborhood of nominal C_2 or equivalently of the nominal C, where the iteration (58) is convergent.

In general, however, the precise scheme according to which **P**2 responds to **P**1's policy choices may not be common information, and hence one would like to develop relaxation-type algorithms for **P**1 which would converge to the true equilibrium solution regardless of what particular scheme **P**2 adopts (for example, Gauss-Seidel or Jacobi). Consider, for example, the scheme where **P**2's responses for different k are modeled by

$$v_{k+1} = C_2 u_{k+1-i_k} + d_2, \tag{60}$$

where $i_k \geq 0$ is an integer denoting the delay in the receipt of current policy information by **P**2 from **P**1. The choice $i_k = 0$ for all k, would correspond to the Gauss-Seidel iteration, and the choice $i_k = 1$ for all k, to the Jacobi iteration — assuming that u_{k+1} is still determined according to (52). An extreme case would be the totally asynchronous communication where $\{i_k\}_{k\geq 0}$ could be any sequence of positive integers. Under the assumptions that **P**1 communicates new policy choices to **P**2 *infinitely often*, and he uses the simple ("nonrelaxed") iteration

$$u_{k+1} = C_1 v_k + d_1, \tag{61}$$

it is known from the work of *Chazan and Miranker (1969)* [21] that such a scheme converges if, and only if,

$$\rho(|C|) < 1 \tag{62}$$

where $|C|$ is the matrix derived from C by multiplying all its negative entries by -1.

This condition can be improved upon, however, by incorporating relaxation terms in (61), such as

$$u_{k+1} = \alpha u_k + (1 - \alpha)C_1 v_k + (1 - \alpha)d_1 \tag{63}$$

where α is some scalar. The condition for convergence of any asynchronously implemented version of (60) and (63) in this case is

$$\rho(\bar{A}(\alpha)) \quad < \quad 1 \tag{64}$$

where

$$\bar{A}(\alpha) \quad := \quad \begin{pmatrix} |\alpha| I & |(1-\alpha)C_1| \\ |C_2| & 0 \end{pmatrix}. \tag{65}$$

Clearly, there is a value of $\alpha \neq 0$ for which (64) requires a less stringent condition (on C_1 and C_2) than (62). For example, if C_1 and C_2 are scalars, and $\alpha = \frac{1}{2}$, (64) dictates

$$C_1 C_2 < 4$$

while (62) requires that $C_1 C_2 < 1$.

From a game theoretic point of view, each of the iteration schemes discussed above corresponds to a game with a sufficiently large number of stages and with a particular mode of play among the players. Moreover, the objective of each player is to minimize a kind of an average long horizon cost, with costs at each stage contributing to this average cost. Viewing this problem overall as a multi-act NZSG, we observe that the behavior of each player at each stage of the game is rather "myopic", since at each stage the players minimize their cost functions only under past information, and quite in ignorance of the possibility of any future moves. If the possibility of future moves is also taken into account, then the rational behavior of each player at a particular stage could be quite different. Such myopic decision making could make sense, however, if the players have absolutely no idea as to how many stages the game comprises, in which case there is the possibility that at any stage a particular player could be the last one to act in the game. In such a situation, risk-aversing players would definitely adopt "myopic" behavior, minimizing their current cost functions under only the past information, whenever given the opportunity to act.

Two-person zero-sum games

Since ZSGs are special types of two-person NZSGs with $J_1 = -J_2$ (**P**1 minimizing and **P**2 maximizing), in which case the NE solution concept coincides with the concept of SPE, a special version of Proposition 12 will be valid for quadratic zero-sum games. To this end, we first note that the relation $J_1 = -J_2$ imposes in (44) the restrictions

$$R_{12}^{1\,\prime} = -R_{21}^2, R_{11}^2 = -R_{11}^1, R_{22}^1 = -R_{22}^2, r_1^2 = -r_1^1, r_2^1 = -r_2^2, c_1 = -c_2,$$

under which matrix R defined by (48) can be written as

$$R = \begin{pmatrix} R_{11}^1 & R_{12}^1 \\ -R_{12}^{1\,\prime} & R_{22}^2 \end{pmatrix}$$

which has to be nonsingular for existence of a saddle point. Since R can also be written as the sum of two matrices

$$R = \begin{pmatrix} R_{11}^1 & 0 \\ 0 & R_{22}^2 \end{pmatrix} + \begin{pmatrix} 0 & R_{12}^1 \\ -R_{12}^{1\,\prime} & 0 \end{pmatrix}$$

the first one being positive definite and the second one skew-symmetric, and since eigenvalues of the latter are always imaginary, it readily follows that R is a nonsingular matrix. Hence we arrive at the conclusion that quadratic strictly convex-concave zero-sum games admit unique saddle-point equilibrium in pure strategies.

Corollary 10.1 *The strictly convex-concave quadratic zero-sum game*

$$\begin{aligned} J &= \frac{1}{2}u^{1\prime}R_{11}^1u^1 + u^{1\prime}R_{12}^1u^2 - \frac{1}{2}u^{2\prime}R_{22}^2u^2 + u^{1\prime}r_1^1 + u^{2\prime}r_2^1 + c^1; \\ & R_{11}^1 > 0, R_{22}^2 > 0, \end{aligned}$$

admits a unique saddle-point equilibrium in pure strategies, which is given by

$$\begin{aligned} u^{1^*} &= -[R_{11}^1 + R_{12}^1(R_{22}^2)^{-1}R_{12}^{1\,\prime}]^{-1}[r_1^1 + R_{12}^1(R_{22}^2)^{-1}r_2^1], \\ u^{2*} &= [R_{22}^2 + R_{12}^{1\,\prime}(R_{11}^1)^{-1}R_{12}^1]^{-1}[r_2^1 + R_{12}^{1\,\prime}(R_{11}^1)^{-1}r_1^1]. \end{aligned}$$

◇

Remark 9 The positive-definiteness requirements on R_{11}^1 and R_{22}^2 in Corollary 10.1 are necessary and sufficient for the game kernel to be strictly convex-strictly concave, but this structure is clearly not necessary for the game to admit a saddle point. If the game is simply convex-concave (that is, if the matrices above are nonnegative definite, with a possibility of zero eigenvalues), then a SPE will still exist provided that the upper and lower values are bounded.[23] If the quadratic game is not convex-concave, however, then either the upper or the lower value (or both) will be unbounded, implying that a saddle point will not exist. ◇

[23]For a convex-concave quadratic game, the upper value will not be bounded if, and only if, there exists a $v \in \mathrm{R}^{m_2}$ such that $v'R_{22}^2v = 0$ while $v'r_2^1 \neq 0$. A similar result also applies to the lower value.

Team problems

Yet another special class of NZSGs are the team problems in which the players (or equivalently, members of the team) share a common objective. Within the general framework, this corresponds to the case $J_1 \equiv J^2 \equiv \cdots \equiv J^N \triangleq J$, and the objective is to minimize this cost function over all $u^i \in U^i$, $i = 1, \ldots, N$. The resulting solution N-tuple $(u^{1^*}, u^{2^*}, \ldots, u^{N^*})$ is known as the *team-optimal solution*. The NE solution, however, corresponds to a weaker solution concept in team problems (as we have already seen), the so-called *person-by-person (pbp) optimality*. In a two-member team problem, for example, a pbp optimal solution (u^{1^*}, u^{2*}) dictates satisfaction of the pair of inequalities

$$\begin{aligned} J(u^{1^*}, u^{2^*}) &\leq J(u^1, u^{2^*}), \quad \forall u^1 \in U^1, \\ J(u^{1^*}, u^{2^*}) &\leq J(u^{1^*}, u^2), \quad \forall u^2 \in U^2, \end{aligned}$$

whereas a team-optimal solution (u^{1^*}, u^{2^*}) requires satisfaction of a single inequality

$$J(u^{1^*}, u^{2^*}) \leq J(u^1, u^2), \quad \forall u^1 \in U^1, u^2 \in U^2.$$

A team-optimal solution always implies pbp optimality, but not vice versa. Of course, if J is quadratic and strictly convex on the product space $U^1 \times \cdots \times U^N$, then a unique pbp optimal solution exists, and it is also team-optimal.[24] However, for a cost function that is strictly convex only on individual spaces U^i, but not on the product space, this latter property may not be true. Consider, for example, the quadratic cost function

$$J = (u^1)^2 + (u^2)^2 + 10u^1u^2 + 2u^1 + 3u^2$$

which is strictly convex in u^1 and u^2, separately. The matrix corresponding to R defined by (48) is

$$\begin{pmatrix} 2 & 10 \\ 10 & 2 \end{pmatrix}$$

which is clearly nonsingular. Hence a unique pbp optimal solution will exist. However, a team-optimal solution does not exist since the said matrix (which is also the Hessian of J) has one positive and one negative eigenvalue. By cooperating along the direction of the eigenvector corresponding to the negative eigenvalue, the members of the team can make the value of J as small as possible. In particular, taking $u^2 = -\frac{2}{3}u^1$ and letting $u^1 \to +\infty$, drives J to $-\infty$.

[24]This result may fail to hold true for team problems with strictly convex but nondifferentiable kernels.

The Stackelberg solution

We now elaborate on the SESs of quadratic games of type (44) but with $N = 2$, and **P1** acting as the leader. We first note that since the quadratic cost function J^i is strictly convex in u^i, by Proposition 11 we can confine our investigation of an equilibrium solution to the class of pure strategies. Then, to every announced strategy u^1 of **P1**, the follower's unique response will be as given by (46) with $N = 2$, $i = 2$:

$$u^2 = -(R^2_{22})^{-1}[R^2_{21}u^1 + r^2_2]. \tag{66}$$

Now, to determine the Stackelberg strategy of the leader, we have to minimize J^1 over U^1 and subject to the constraint imposed by the reaction of the follower. Since the reaction curve gives u^2 uniquely in terms of u^1, this constraint can best be handled by substitution of (66) in J^1 and by minimization of the resulting functional (to be denoted by $\tilde{J}^1$) over U^1. To this end, we first determine $\tilde{J}^1$:

$$\begin{aligned}\tilde{J}^1(u^1) \;=\;& \frac{1}{2}u^{1\prime}R^1_{11}u^1 + \frac{1}{2}[R^2_{21}u^1 + r^2_2]'(R^2_{22})^{-1}R^1_{22}(R^2_{22})^{-1}[R^2_{21}u^1 + r^2_2] \\ & -u^{1\prime}R^1_{21}(R^2_{22})^{-1}[R^2_{21}u^1 + r^2_2] + u^{1\prime}r^1_1 \\ & -[R^2_{21}u^1 + r^2_2]'(R^2_{22})^{-1}r^1_2 + c^1.\end{aligned}$$

For the minimum of $\tilde{J}^1$ over U^1 to be unique, we have to impose a strict convexity condition on $\tilde{J}^1$. Because of the quadratic structure of $\tilde{J}^1$, this condition amounts to having the coefficient matrix of the quadratic term in u^1 positive definite, which is

$$\begin{aligned}R^1_{11} + {R^2_{21}}'(R^2_{22})^{-1}R^1_{22}(R^2_{22})^{-1}R^2_{21} - R^1_{21}(R^2_{22})^{-1}R^2_{21} \\ -{R^2_{21}}'(R^2_{22})^{-1}{R^1_{21}}' > 0.\end{aligned} \tag{67}$$

Under this condition, the unique minimizing solution can be obtained by setting the gradient of $\tilde{J}^1$ equal to zero, which yields

$$\begin{aligned}u^{1^*} \;=\;& -[R^1_{11} + R^2_{21}(R^2_{22})^{-1}R^1_{22}(R^2_{22})^{-1}R^2_{21} - R^1_{21}(R^2_{22})^{-1}R^2_{21} \\ & -{R^2_{21}}'(R^2_{22})^{-1}{R^1_{21}}']^{-1}[{R^2_{21}}'(R^2_{22})^{-1}R^1_{22}R^1_{22}(R^2_{22})^{-1}r^2_2 \\ & -R^1_{21}(R^2_{22})^{-1}r^2_2 + r^1_1 - {R^2_{21}}'(R^2_{22})^{-1}r^1_2].\end{aligned} \tag{68}$$

Proposition 13 *Under condition (67), the two-person version of the quadratic game (44) admits a unique Stackelberg strategy for the leader, which is given by (68). The follower's unique response is then given by (66).* ◇

Remark. A sufficient condition for condition (67) is strict convexity of J^1 on the product space $U^1 \times U^2$. $\diamond$

10.2 Stochastic games

We now discuss stochastic static games with quadratic cost functions, for only the case $N = 2$. Stochasticity will enter the game through the cost functions of the players, as weights on the terms linear the action variables. Accordingly, the quadratic cost functions will be given by (where we differentiate between players using subscripts instead of superscripts):

$$\begin{aligned} L_1(u_1, u_2; \xi_1, \xi_2) &= \frac{1}{2} u_1' R_{11} u_1 + u_1' R_{12} u_2 + u_1' \xi_1 \\ L_2(u_1, u_2; \xi_1, \xi_2) &= \frac{1}{2} u_2' R_{22} u_2 + u_2' R_{21} u_1 + u_2' \xi_2 \end{aligned}$$

where R_{ii} are positive definite, and ξ_i's are random vectors of appropriate dimensions. **P**1 and **P**2 do not have access to the values of these random vectors, but they measure another pair of random vectors, y_1 (for **P**1) and y_2 (for **P**2), which carry some information on ξ_i's. We assume that all four random variables have bounded first and second moments, and their joint distribution is common information to both players.

P1 uses y_1 in the construction of her policy and subsequently action, where we denote her policy variable (strategy) by γ_1, so that $u_1 = \gamma_1(y_1)$. Likewise we introduce γ_2 as the strategy for **P**2, so that $u_2 = \gamma_2(y_2)$. These policy variables have no restrictions imposed on them other than *measurability* and that U_i's should have bounded first and second moments. Let Γ_1 and Γ_2 be the corresponding spaces where γ_1 and γ_2 belong. Then for each $\gamma_i \in \Gamma_i,\ i = 1, 2$, using $u_1 = \gamma_1(y_1)$ and $u_2 = \gamma_2(y_2)$ in L_1 and L_2, and taking expectation over the statistics of the four random variables, we arrive at the normal form of the game (in terms of the strategies), captured by the expected costs:

$$\begin{aligned} J_1(\gamma_1, \gamma_2) &= E[L_1(\gamma_1(y_1), \gamma_2(y_2); \xi_1, \xi_2)] \\ J_2(\gamma_1, \gamma_2) &= E[L_2(\gamma_1(y_1), \gamma_2(y_2); \xi_1, \xi_2)] \end{aligned}$$

We are looking for a NE in $\Gamma_1 \times \Gamma_2$, where NE is defined in the usual way.

Using properties of conditional expectation, for fixed $\gamma_2 \in \Gamma_2$, there exists a unique $\gamma_1 \in \Gamma_1$ that minimizes $J_1(\gamma_1, \gamma_2)$ over Γ_1. This unique solution is given by

$$\gamma_1(y_1) = R_{11}^{-1} \left[R_{12} E \left[\gamma_2(y_2) | y_1 \right] + E\xi_1 | y_1 \right] \right] =: T_1(\gamma_2)(y_1)$$

which is the unique response by **P**1 to a strategy of **P**2. Likewise, **P**2's response to **P**1 is unique:

$$\gamma_2(y_2) = R_{22}^{-1}\left[R_{21}E\left[\gamma_1(y_1)|y_2\right] + E\xi_2|y_2\right] =: T_2(\gamma_1)(y_2)$$

Hence, in the policy space, we will be looking for fixed point of

$$\gamma_1 = T_1(\gamma_2)\,, \quad \gamma_2 = T_2(\gamma_1)$$

and substituting the second one into the first, we have

$$\gamma_1 = (T_1 \circ T_2)(\gamma_1)$$

where $T_1 \circ T_2$ is the composite map. This will admit a unique solution if $T_1 \circ T_2$ is a contraction (note that Γ_1 is a Banach space).

Now, writing out this fixed point equation

$$\begin{aligned}\gamma_1(y) &= R_{11}^{-1}R_{12}R_{22}^{-1}R_{21}E\left[E\left[\gamma_1(y_1)|y_2\right]|y_1\right] + R_{11}^{-1}R_{12}R_{22}^{-1}E[\xi_2|y_1] - R_{11}^{-1}E[\xi_1|y_1] \\ &=: \tilde{T}_1(\gamma_1)(y_1) + R_{11}^{-1}R_{12}R_{22}^{-1}E[\xi_2|y_1] - R_{11}^{-1}E[\xi_1|y_1]\end{aligned}$$

Hence, $T_1 \circ T_2$ is a contraction if, and only if, $\tilde{T}_1$ is, and since conditional expectation is a non-expansive mapping, it follows that the condition for existence of NE (and its stability) is exactly the one obtained in the previous section for the deterministic game, that is,

$$\rho(C_1C_2) = \rho(R_{11}^{-1}R_{12}R_{22}^{-1}R_{21}) < 1$$

In this case, the recursion

$$\gamma_1^{(k+1)} = (T_1 \circ T_2)(\gamma_1^{(k)})$$

will converge for all $\gamma_1^{(0)} \in \Gamma_1$. Note that if this sequence converges, so does the one generated by

$$\gamma_1^{(k+1)} = T_1(\gamma_2^{(k)})\,, \quad \gamma_2^{(k+1)} = T_2(\gamma_1^{(k)})$$

for all $\gamma_i^{(0)} \in \Gamma_i,\ i = 1, 2$. And *the limit is the unique NE.*

If the four random vectors are jointly Gaussian distributed, then the unique NE will be affine in y_1 (for **P**1) and y_2 (for **P**2), which follows from properties of Gaussian random variables, by taking $\gamma_i^{(0)} = 0$. Further results on this class of stochastic games with $N > 2$, and when the players do not agree on a common underlying statistics for the uncertainty, can be found in [45].

11 Lecture 11: Dynamic infinite games: Nash equilibria

In this lecture, I discuss the Nash equilibria of dynamic infinite games, which are also known as discrete-time games. After a precise formulation of such games, the informational non-uniqueness feature is discussed in detail through an example game before stating some general results. Also, derivation of the open-loop and closed-loop feedback NE is outlined.

11.1 General formulation

I start with a general definition.

Definition 17 *An N-person discrete-time deterministic infinite dynamic game*[25] *of prespecified fixed duration involves:*

(i) An index set $\mathcal{N} = \{1, \dots, N\}$ called the players' set.

(ii) An index set $\mathbf{K} = \{1, \dots, K\}$ denoting the stages *of the game, where K is the maximum possible number of moves a player is allowed to make in the game.*

(iii) An infinite set X with some topological structure, called the state set (space) *of the game, to which the state of the game (x_k) belongs for all $k \in \mathbf{K}$ and $k = K + 1$.*

(iv) An infinite set U_k^i with some topological structure, defined for each $k \in \mathbf{K}$ and $i \in \mathcal{N}$, which is called the action (control) *set of $\mathbf{P}i$ at stage k. Its elements are the permissible* actions u_k^i *of $\mathbf{P}i$ at stage k.*

(v) A function $f_k : X \times U_k^1 \times \cdots \times U_k^N \to X$, defined for each $k \in \mathbf{K}$, so that

$$x_{k+1} = f_k(x_k, u_k^1, \dots, u_k^N), k \in \mathbf{K} \qquad (i)$$

for some $x_1 \in \mathbf{X}$ which is called the initial state *of the game. The difference equation (i) is called the* state equation *of the dynamic game, describing the evolution of the underlying decision process.*

(vi) A set Y_k^i with some topological structure, defined for each $k \in \mathbf{K}$ and $i \in \mathcal{N}$, and called the observation set *of $\mathbf{P}i$ at stage k, to which the observation y_k^i of $\mathbf{P}i$ belongs at stage k.*

[25] Also known as an "N-person deterministic multi-stage game."

(vii) A function $h_k^i : X \to Y_k^i$*, defined for each* $k \in \mathbf{K}$ *and* $i \in \mathcal{N}$ *, so that*

$$y_k^i = h_k^i(x_k), \quad k \in \mathbf{K}, \quad i \in \mathcal{N},$$

which is the state-measurement (-observation) *equation of* $\mathbf{P}i$ *concerning the value of* x_k*.*

(viii) A finite set η_k^i*, defined for each* $k \in \mathbf{K}$ *and* $i \in \mathcal{N}$ *as a subcollection of* $\{y_1^1, \ldots, y_k^1; y_1^2, \ldots, y_k^2; \ldots; y_1^N, \ldots, y_k^N; u_1^1, \ldots, u_{k-1}^1; u_1^2, \ldots, u_{k-1}^2; \ldots; u_1^N, \ldots, u_{k-1}^1\}$*, which determines the information gained and recalled by* $\mathbf{P}i$ *at stage* k *of the game. Specification of* η_k^i *for all* $k \in \mathbf{K}$ *characterizes the* information structure (pattern) *of* $\mathbf{P}i$*, and the collection (over* $i \in \mathcal{N}$*) of these information structures is the* information structure *of the game.*

(ix) A set N_k^i*, defined for each* $k \in \mathbf{K}$ *and* $i \in \mathcal{N}$ *as an appropriate subset of* $\{(Y_1^1 \times \cdots \times Y_k^1) \times \cdots \times (Y_1^N \times \cdot \times Y_k^N) \times (U_1^1 \times \cdots \times U_{k-1}^1) \times \cdots \times (U_1^N \times \cdots \times U_{k-1}^N)\}$*, compatible with* η_k^i*.* N_k^i *is called the* information space *of* $\mathbf{P}i$ *at stage* k*, induced by his information* η_k^i*.*

(x) A prespecified class Γ_k^i *of mappings* $\gamma_k^i : N_k^i \to U_k^i$ *which are the permissible* strategies *of* $\mathbf{P}i$ *at stage* k*. The aggregate mapping* $\gamma^i = \{\gamma_1^i, \gamma_2^i, \ldots, \gamma_K^i\}$ *is a* strategy *for* $\mathbf{P}i$ *in the game, and the class* Γ^i *of all such mappings* γ^i *so that* $\gamma_k^i \in \Gamma_k^i$*,* $k \in \mathbf{K}$*, is the strategy set (space) of* $\mathbf{P}i$*.*

(xi) A functional $L^i : (X \times U_1^1 \times \cdots \times U_1^N) \times (X \times U_2^1 \times \cdots \times U_2^N) \times \cdots \times (X \times U_K^1 \times, \ldots, \times U_K^N) \to \mathbf{R}$ *defined for each* $i \in \mathcal{N}$*, and called the* cost functional *of* $\mathbf{P}i$ *in the game of fixed duration.* $\diamond$

The preceding definition of a deterministic discrete-time infinite dynamic game is clearly not the most general one that could be given, first because the duration of the game need not be fixed, but be a variant of the players' strategies, and second because a "quantitative" measure might not exist to reflect the preferences of the players among different alternatives. In other words, it is possible to relax and/or modify the restrictions imposed by items (ii) and (xi) in Definition 17, and still retain the essential features of a dynamic game. A relaxation of the requirement of (ii) would involve introduction of a *termination set* $\Lambda \subset X \times \{1, 2, \ldots\}$, in which case we say that *the game terminates, for a given* N*-tuple of strategies, at stage* k*, if* k *is the smallest integer for which* $(x_k, k) \in \Lambda$.[26] Such a more general formulation clearly also covers fixed duration game problems in

[26]It is, of course, implicit here that x_k is determined by the given N-tuple of strategies, and the strategies are defined as in Definition 17, but by taking K sufficiently large.

which case $\Lambda = X \times \{K\}$ where K denotes the number of stages involved. A modification of (xi), on the other hand, might for instance involve a "qualitative" measure (instead of the "quantitative" measure induced by the cost functional), thus giving rise to the so-called *qualitative games* (as opposed to "*quantitative games*" covered by Definition 17). Any qualitative game (also called *game of kind*) can, however, be formulated as a quantitative game (also known as *game of degree*) by assigning a fixed cost of zero to paths and strategies leading to preferred states, and positive cost to the remaining paths and strategies. For example, in a two-player game, if **P**1 wishes to reach a certain subset Λ of the state set X after K stages and **P**2 wishes to avoid it, we can choose

$$L^1 = \begin{cases} 0 & \text{if } x_{K+1} \in \Lambda \\ 1 & \text{otherwise} \end{cases}, \qquad L^1 = \begin{cases} 0 & \text{if } x_{K+1} \in \Lambda \\ -1 & \text{otherwise} \end{cases}$$

and thus consider it as a zero-sum quantitative game.

Now, returning to Definition 17, we note that it corresponds to an *extensive form description* of a dynamic game, since the evolution of the game, the information gains and exchanges of the players throughout the decision process, and the interactions of the players among themselves are explicitly displayed in such a formulation. It is, of course, also possible to give a normal form description of such a dynamic game, which in fact readily follows from Definition 17. More specifically, for each fixed initial state x_1 and for each fixed N-tuple permissible strategies $\{\gamma^i \in \Gamma^i; i \in \mathcal{N}\}$ the extensive form description leads to a *unique* set of vectors $\{u_k^i \equiv \gamma_k^i(\eta_k^i), x_{k+1}; i \in \mathcal{N}, k \in \mathbf{K}\}$ because of the causal nature of the information structure and because the state evolves according to a difference equation. Then, substitution of these quantities into $L^i (i \in \mathcal{N})$ clearly leads to a unique N-tuple of numbers reflecting the corresponding costs to the players. This further implies existence of a composite mapping $J^i : \Gamma^1 \times \cdots \times \Gamma^N \to \mathbf{R}$, for each $i \in \mathcal{N}$, which is also known as the *cost functional* of $\mathbf{P}i$ $(i \in \mathcal{N})$. Hence, the permissible strategy spaces of the players (*i.e.* $\Gamma^1, \ldots, \Gamma^N$) together with these cost functions $(J^1, \ldots, J^N)$ constitute the *normal form* description of the dynamic game for each fixed initial state vector x_1.

It should be noted that, under the normal form description, there is no essential difference between infinite discrete-time dynamic games and finite games (the complex structure of the former being disguised in the strategy spaces and the cost functionals), and this permits us to adopt all the noncooperative equilibrium solution concepts (SPE, NE, SES) introduced earlierdirectly in the present framework.

Before concluding our discussion on the ingredients of a discrete-time dynamic game as presented in Definition 17 we now finally classify possible information structures that will be encountered in the following chapters, and also introduce a specific class of cost functions — the so-called *stage-additive* cost functions.

Definition 18 *In an N-person discrete-time deterministic dynamic game of prespecified fixed duration, we say that $\mathbf{P}i$'s information structure is*

(i) open-loop *(OL) pattern if $\eta_k^i = \{x_1\}$, $k \in \mathbf{K}$,*

(ii) closed-loop perfect state information *(CLPS) pattern if $\eta_k^i = \{x_1, \ldots, x_k\}$, $k \in \mathbf{K}$,*

(iii) closed-loop imperfect state information *(CLIS) pattern if $\eta_k^i = \{y_1^i, \ldots, y_k^i\}$, $k \in \mathbf{K}$,*

(iv) memoryless perfect state information *(MPS) pattern if $\eta_k^i = \{x_1, x_k\}$, $k \in \mathbf{K}$,*

(v) feedback (perfect state) information *(FB) pattern if $\eta_k^i = \{x_k\}$, $k \in \mathbf{K}$,*

(vi) feedback imperfect state information *(FIS) pattern if $\eta_k^i = \{y_k^i\}$, $k \in \mathbf{K}$,*

(vii) one-step delayed CLPS *(1DCLPS) pattern if $\eta_k^i = \{x^i, \ldots, x_{k-1}\}$, $k \in \mathbf{K}, k \neq 1$,*

(viii) one-step delayed observation sharing *(1DOS) pattern if $\eta_k^i = \{y_1, \ldots, y_{k-1}, y_k^i\}$, $k \in \mathbf{K}$, where $y_j \triangleq \{y_j^1, y_j^2, \ldots, y_j^N\}$.* ⋄

Definition 19 *In an N-person discrete-time deterministic dynamic game of prespecified fixed duration (i.e. K stages), $\mathbf{P}i$'s cost functional is said to be* stage-additive *if there exist $g_k^i : X \times X \times U_k^1 \times \cdots \times U_k^N \to \mathbf{R}$, $(k \in \mathbf{K})$ so that*

$$L^i(u^1, \ldots, u^N) = \sum_{k=1}^{K} g_k^i(x_{k+1}, u_k^1, \ldots, u_k^N, x_k), \tag{69}$$

where

$$u^j = (u_1^{j\prime}, \ldots, u_K^{j\prime})'.$$

Furthermore, if $L^i(u^1, \ldots, u^N)$ depends only on x_{K+1} (the terminal state), then we call it a terminal cost functional. ⋄

Remark 10 It should be noted that every stage-additive cost functional can be converted into a terminal cost functional, by introducing an additional variable z_k $(k \in \mathbf{K})$ through the recursive relation

$$z_{k+1} = z_k + g_k(f_k(x_k, u_k^1, \dots, u_k^N), u_k^1, \dots, u_k^N, x_k), z_1 = 0,$$

and by adjoining z_k to the state vector x_k as the last component. Denoting the new state vector as $\tilde{x}_k \triangleq (x_k', z_k)'$, the stage-additive cost functional (69) can then be written as

$$L^i(u^1, \dots, u^N) = (0, \dots, 0, 1)\tilde{x}_{K+1}$$

which is a terminal cost functional. $\diamond$

11.2 Informational non-uniqueness

This section is devoted to an elaboration on the occurrence of "informationally nonunique" Nash equilibria in discrete-time dynamic games, and to a general discussion on the interplay between information patterns and existence-uniqueness properties of noncooperative equilibria in such games. First, we consider, in some detail, a scalar three-person dynamic game which admits uncountably many Nash equilibria, and which features several important properties of infinite dynamic games. Then, we discuss these properties in a general context.

Consider a scalar three-person two-stage linear-quadratic dynamic game in which each player acts only once. The state equation is given by

$$x_3 = x_2 + u^1 + u^2; \quad x_2 = x_1 + u^3, \tag{70}$$

and the cost functionals are defined as

$$L^1 = (x^3) + (u^1)^2; \quad L^2 = -L^3 = -(x_3)^2 + 2(u^2)^2 - (u^3)^2. \tag{71}$$

In this formulation, u^i is the scalar unconstrained control variable of $\mathbf{P}i$ $(i = 1, 2, 3)$, and x_1 is the initial state whose value is known to all players. $\mathbf{P}1$ and $\mathbf{P}2$, who act at stage 2, have also access to the value of x_2 (*i.e.* the underlying information pattern is CLPS (or, equivalently, MPS) for both $\mathbf{P}1$ and $\mathbf{P}2$), and their permissible strategies are taken as twice continuously differentiable mappings from $\mathbf{R} \times \mathbf{R}$ into $\mathbf{R}$. A permissible strategy for $\mathbf{P}3$, on the other hand, is any measurable mapping from $\mathbf{R}$ into $\mathbf{R}$. This, then, completes description of the strategy spaces Γ^1, Γ^2 and Γ^2 (for $\mathbf{P}1$, $\mathbf{P}2$ and $\mathbf{P}3$, respectively), where we suppress the subindices denoting the corresponding stages since each player acts only once.

Now, let $\{\gamma^1 \in \Gamma^1, \gamma^2 \in \Gamma^2, \gamma^3 \in \Gamma^3\}$ denote any noncooperative (Nash) equilibrium solution for this three-person nonzero-sum dynamic game. Since L^i is strictly convex in u^i $(i = 1, 2)$, a set of necessary and sufficient conditions for γ^1 and γ^2 to be in equilibrium (with $\gamma^3 \in \Gamma^3$ fixed) is obtained by differentiation of L^i with respect to u^i $(i = 1, 2)$, thus leading to

$$\begin{aligned}
\bar{x}_2(\gamma^3) + 2\gamma^1(\bar{x}_2, x_1) + \gamma^2(\bar{x}_2, x_1) &= 0 \\
-\bar{x}_2(\gamma^3) - \gamma^1(\bar{x}_2, x_1) + \gamma^2(\bar{x}_2, x_1) &= 0
\end{aligned}$$

where

$$\bar{x}_2 \triangleq \bar{x}_2(\gamma^3) = x_1 + \gamma^3(x_1).$$

Solving for $\gamma^1(\bar{x}_2, x_1)$ and $\gamma^2(\bar{x}_2, x_1)$ from the foregoing pair of equations, we obtain

$$\gamma^1(\bar{x}_2, x_1) = -\frac{2}{3}\bar{x}_2 \tag{72}$$

$$\gamma^2(\bar{x}_2, x_1) = \frac{1}{3}\bar{x}_2 \tag{73}$$

which are the side conditions on the equilibrium strategies γ^1 and γ^2, and which depend on the equilibrium strategy γ^3 of $\mathbf{P}3$. Besides these side conditions, the Nash equilibrium strategies of $\mathbf{P}1$ and $\mathbf{P}2$ have no other natural constraints imposed on them. To put it in other words, *every Nash equilibrium strategy for* $\mathbf{P}1$ *will be a closed-loop representation of the open-loop value (72), and every Nash equilibrium strategy for* $\mathbf{P}2$ *will be a closed-loop representation of (73).*

To complete the solution of the problem, we now proceed to stage one. Since $\{\gamma^1, \gamma^2, \gamma^3\}$ constitutes an equilibrium triple, with $\{u^1 = \gamma^1(x_2, x_1), u^2 = \gamma^2(x_2, x_1)\}$ substituted into L^3 the resulting cost functional of $\mathbf{P}3$ (denoted as $\tilde{L}^3$) should attain a minimum at $u^3 = \gamma^3(x_1)$; and since γ^1 and γ^2 are twice continuously differentiable in their arguments, this requirement can be expressed in terms of the relations

$$\frac{d}{du^3}\tilde{L}^3(\gamma^3(x_1)) = x_3(1 + \gamma^1_{x_2} + \gamma^2_{x_2}) - 2\gamma^2_{x_2}\gamma^2 + \gamma^3(x_1) = 0$$

$$\frac{d^2}{(du^3)^2}\tilde{L}^3(\gamma^3(x_1)) = (1+\gamma^1_{x_2}+\gamma^2_{x_2})^2 + x_3(\gamma^1_{x_2x_2}+\gamma^2_{x_2x_2}) - 2(\gamma^2_{x_2})^2$$

$$-2\gamma^2_{x_2x_2}\gamma^2 + 1 > 0^{27}$$

where we have suppressed the arguments of the strategies. Now, by utilizing the side conditions (72)-(73) in the above set of relations, we arrive at a simpler set of relations which are, respectively,

$$\frac{2}{3}[x_1+\gamma^3(x_1)][1+\gamma^1_{x_2}(\bar{x}_2,x_1)] + \gamma^3(x_1) = 0 \tag{74}$$

$$[1+\gamma^1_{x_2}(\bar{x}_2,x_1)+\gamma^2_{x_2}(\bar{x}_2,x_1)]^2 - \frac{2}{3}\bar{x}_2\gamma^1_{x_2x_2}(\bar{x}_2,x_1) - 2[\gamma^2_{x_2}(\bar{x}_2,x_1)]^2 + 1 > 0, \tag{75}$$

where $\bar{x}_2$ is again defined as

$$\bar{x}_2 = x_1 + \gamma^3(x_1). \tag{76}$$

These are the relations which should be satisfied by an equilibrium triple, in addition to (72) and (73). The following proposition summarizes the result.

Proposition 14 *Any triple $\{\gamma^{1*}\in\Gamma^1, \gamma^{2*}\in\Gamma^2, \gamma^{3*}\in\Gamma^3\}$ that satisfies (72)-(73) and (74)-(75), and also possesses the additional feature that (74) with $\gamma^1=\gamma^{1*}$ and $\gamma^2=\gamma^{*2}$ admits a unique solution $\gamma^3=\gamma^{3*}$, constitutes a Nash equilibrium solution for the nonzero-sum dynamic game described by (70)-(71).*

Proof This result follows from the derivation outlined prior to the statement of the proposition. Uniqueness of the solution of (74) for each pair $\{\gamma^{1*},\gamma^{2*}\}$ is imposed in order to insure that the resulting γ^{3*} is indeed a globally minimizing solution for $\tilde{L}^3$. $\diamond$

We now claim that there exists an uncountable number of triplets that satisfy the requirements of Proposition 14. To justify this claim, and to obtain a set of explicit solutions, we consider the class of γ^1 and γ^2 described as

$$\begin{aligned}\gamma^1(x_2,x_1) &= = -\frac{2}{3}x_2 + p[x_2-\bar{x}_2(\gamma^3)]\\ \gamma^2(x_2,x_1) &= = \frac{1}{3}x_2 + q[x_2-\bar{x}_2(\gamma^3)],\end{aligned}$$

[27] Here, we could of course also have nonstrict inequality (*i.e.* $\geq$) in which case we also have to look at higher-order derivatives of $\tilde{L}^3$. We avoid this by restricting our analysis at the outset only to those equilibrium triples $\{\gamma^1,\gamma^2,\gamma^3\}$ which lead to an $\tilde{L}^3$ that is locally strictly convex at the solution point.

where p and q are free parameters. These structural forms for γ^1 and γ^2 clearly satisfy the side conditions (72) and (73), respectively. With these choices, (74) can be solved uniquely (for each p, q) to give

$$\gamma^3(x_1) = -[(2+6p)/(11+6p)]x_1, \quad p \neq -11/6,$$

with the existence condition (75) reading

$$\left(\frac{2}{3}+p\right)^2 + 2pq - q^2 + \frac{7}{9} > 0. \tag{77}$$

The scalar $\bar{x}_2$ is then given by

$$\bar{x}_2 = [9/(11+6p)]x_1.$$

Hence,

Proposition 15 *The set of strategies*

$$\begin{aligned}
\gamma^{1*}(x_2, x_1) &= -\frac{2}{3}x_2 + p\{x_2 - [9/(11+6p)]x_1\} \\
\gamma^{2*}(x_2, x_1) &= \frac{1}{3}x_2 + q\{x_2 - [9/(11+6p)]x_1\} \\
\gamma^{3*}(x_1) &= -[(2+6p)/(11+6p)]x_1
\end{aligned}$$

constitutes a Nash equilibrium solution for the dynamic game described by (70)-(71), for all values of the parameters p and q satisfying (77) and with $p \neq -11/6$. The corresponding equilibrium costs of the players are

$$\begin{aligned}
J^{1*} &= 2[6/(11+6p)]^2(x_1)^2 \\
J^{2*} &= -J^{3*} = -[(22+24p+36p^2)/(11+6p)^2](x_1)^2.
\end{aligned}$$

$\diamond$

Several remarks and observations are in order here, concerning the Nash equilibrium solutions presented above.

(1) The nonzero-sum dynamic game of this section admits *uncountably many* Nash equilibrium solutions, each one leading to a different equilibrium cost triple.

(2) Within the class of linear strategies, Proposition 15 provides the complete solution to the problem, which is parametrized by p and q.

(3) The equilibrium strategy of **P**3, as well as the equilibrium cost values of all three players, depend only on p (not on q), whereas the existence condition (77) involves both p and q. There is indeed an explanation for this: the equilibrium strategies of **P**1 and **P**2 are in fact representations of the open-loop values (72)-(73) on appropriate trajectories. By choosing a specific representation of (72), **P**1 influences the cost functional of **P**3 and thereby the optimization problem faced by him. Hence, for each different representation of (72), **P**3 ends up, in general, with a different solution to his minimization problem, which directly contributes to nonuniqueness of Nash equilibria. For **P**2, on the other hand, even though he may act analogously—*i.e.* choose different representations of (73)—these different representations do not lead to different minimizing solutions for **P**3 (but instead affect only the existence of a minimizing solution) since $L^2 \equiv -L^3$, *i.e.* **P**2 and **P**3 have completely conflicting goals (see the next remark for further clarification). Consequently, $\gamma^{3*}(x_1)$ is independent of q, but the existence condition explicitly depends upon q. (This is true also for nonlinear representations, as it can be seen from (74) and (75).)

(4) If **P**1 has access to only x_2 (and not to x_1), then, necessarily, $p = 0$, and both **P**1 and **P**3 have unique equilibrium strategies which are $\{\gamma^{1*}(x_2) = -\frac{2}{3}x_2, \gamma^{3*}(x_1) = -(2/11)x_1\}$. (This is true also within the class of nonlinear strategies.) Furthermore, the equilibrium cost values are also unique (simply set $p = 0$ in $J^{i*}, i = 1, 2, 3$, in Proposition 15). However, the existence condition (77) still depends on q, since it now reduces to $q^2 < 11/9$. The reason for this is that **P**2 has still the freedom of employing different representations of (73), which affects existence of the equilibrium solution but not the actual equilibrium state trajectory, since **P**2 and **P**3 are basically playing a zero-sum game (in which case the equilibrium (*i.e.* saddle-point) solutions are interchangeable).

(5) By setting $p = q = 0$ in Proposition 15, we obtain the *unique feedback* Nash equilibrium solution of the dynamic game under consideration (which exists since $p = 0 \neq 11/6$, and (77) is satisfied).

(6) Among the uncountable number of Nash equilibrium solutions presented in Proposition 15,

there exists a subsequence of strategies which brings **P**1's Nash cost arbitrarily close to zero which is the lowest possible value L^1 can attain. Note, however, that the corresponding cost for **P**3 approaches $(x_1)^2$ which is unfavorable to him.

Before concluding this section, it is worthy to note that the linear equilibrium solutions presented in Proposition 15 are not the only ones that the dynamic game under consideration admits, since (74) will also admit nonlinear solutions. To obtain an explicit nonlinear equilibrium solution, we may start with a nonlinear representation of (72), for instance

$$\gamma^1(x_2, x_1) = -\frac{2}{3}x_2 + p[x_2 - \bar{x}_2(\gamma^3)]^2,$$

substitute it into (74), and solve for a corresponding $\gamma^3(x_1)$, checking at the same time satisfaction of the second order condition (75). Such a derivation (of nonlinear Nash solutions in a linear-quadratic game) can be found in [46].

We had already seen appearance of "informationally non-unique" Nash equilibria in finite multi-act nonzero-sum dynamic games, which was mainly due to the fact that an increase in information to one or more players leaves the Nash equilibrium obtained under the original information pattern unchanged, but it also creates new equilibria (cf. Proposition 5). In infinite games, the underlying reason for occurrence of informationally non-unique Nash equilibria is essentially the same (though much more intricate), and a counterpart of Proposition 5 can be verified. Toward this end we first introduce the notion of "informational inferior" in such dynamic games.

Definition 20 *Let I and II be two N-person K-stage infinite dynamic games which admit precisely the same extensive form description except the underlying information structure (and, of course, also the strategy spaces whose descriptions depend on the information structure. Let η^i_{I} (respectively, η^i_{II}) denote the information pattern of* **P***i in the game I (respectively, II), and let the inclusion relation $\eta^i_{\mathrm{I}} \subseteq \eta^i_{\mathrm{II}}$ imply that whatever* **P***i knows at each stage of game I he also knows at the corresponding stages of game II, but not necessarily vice versa. Then, I is* informationally inferior *to II if $\eta^i_{\mathrm{I}} \subseteq \eta^i_{\mathrm{II}}$ for all $i \in \mathcal{N}$, with strict inclusion for at least one i.* ◇

Proposition 16 *Let I and II be two N-person K-stage infinite dynamic games as introduced in Definition 20, so that I is informationally inferior to II. Furthermore, let the strategy spaces of the* *players in the two games be compatible with the given information patterns and the constraints (if any) imposed on the controls, so that $\eta_{\mathrm{I}}^i \subseteq \eta_{\mathrm{II}}^i$ implies $\Gamma_{\mathrm{I}}^i \subseteq \Gamma_{\mathrm{II}}^i$, $i \in \mathcal{N}$. Then,*

(i) any Nash equilibrium solution for I *is also a Nash equilibrium solution for* II*,*

(ii) if $\{\gamma^1, \ldots, \gamma^N\}$ is a Nash equilibrium solution for II such that $\gamma^i \in \Gamma_{\mathrm{I}}^i$ for all $i \in \mathcal{N}$, then it is also a Nash equilibrium solution for I.

Proof Let $\{\gamma^{i*}; i \in \mathcal{N}\}$ constitute a Nash equilibrium solution for I. Then, by definition,

$$J^1(\gamma^{1*}, \gamma^{2*}, \ldots, \gamma^{N*}) \leq J^1(\gamma^1, \gamma^{2*}, \ldots, \gamma^{N*}), \quad \forall \gamma^1 \in \Gamma_{\mathrm{I}}^1;$$

therefore, **P**1 minimizes $J^1(\cdot, \gamma^{2*}, \ldots, \gamma^{N*})$ over Γ_{I}^1, with the corresponding solution being $\gamma^{1*} \in \Gamma_{\mathrm{I}}^1$.

Now consider minimization of the same expression over Γ_{II}^1 ($\supseteq \Gamma_{\mathrm{I}}^1$) which reflects an increase in deterministic information concerning the values of state. But, since we have a deterministic optimization problem, the minimum value of $J^1(\gamma^1, \gamma^{2*}, \ldots, \gamma^{N*})$ does not change with an increase in information. Hence,

$$\min_{\gamma^1 \in \Gamma_{\mathrm{II}}^1} J^1(\gamma^1, \gamma^{2*}, \gamma^{N*}) = J^1(\gamma^{1*}, \gamma^{2*}, \ldots, \gamma^{N*});$$

and furthermore since $\gamma^{1*} \in \Gamma_{\mathrm{II}}^1$ we have the inequality

$$J^1(\gamma^{1*}, \gamma^{2*}, \ldots, \gamma^{N*}) \leq J^1(\gamma^1, \gamma^{2*}, \ldots, \gamma^{N*}), \quad \forall \gamma^1 \in \Gamma_{\mathrm{II}}^1.$$

Since **P**1 was an arbitrary player in this discussion, it follows in general that

$$\begin{aligned} J^i(\gamma^{1*}, \ldots, \gamma^{i*}, \ldots, \gamma^{N*}) \quad &\leq \quad J^i(\gamma^{1*}, \ldots, \gamma^{i-1*}, \gamma^i, \gamma^{i+1*}, \ldots, \gamma^{N*}) \\ &\qquad\qquad \forall \gamma^i \in \Gamma_{\mathrm{II}}^i \quad (i \in \mathcal{N}) \end{aligned}$$

which verifies (i) of Proposition 16. Proof of (ii) is along similar lines. $\diamond$

Since there corresponds at least one informationally inferior game (viz. a game with an open-loop information structure) to every multi-stage game with CLPS information, the foregoing result clearly provides one set of reasons for existence of "informationally non-unique" Nash equilibria in infinite dynamic games (as Proposition 5 did for finite dynamic games). However, this is not yet the whole story, as it does not explain occurrence of uncountably many equilibria in such games. What is really responsible for this is the existence of uncountably many *representations* of a strategy under dynamic information. To elucidate somewhat further, consider the scalar three-person dynamic game of this section. We have already seen that, for each fixed equilibrium strategy γ^3 of **P**3, the equilibrium strategies of **P**1 and **P**2 have unique open-loop values given by (72) and (73), respectively, but they are otherwise free. We also know that there exist infinitely many closed-loop representations of such open-loop policies [1]; and since each one has a different structure, this leads to infinitely many equilibrium strategies for **P**3, and consequently to a plethora of Nash equilibria.

11.3 Open-loop and closed-loop feedback NE

The message of the previous section was that in formulating dynamic games and seeking NE for them, one has to make sure that informational non-uniqueness is avoided, particularly the occurrence of infinitely many NE. One way of doing that would be to disallow the players to receive dynamic information (such as state information) correlated with the actions of the players in the past; the open-loop (OL) information structure does that, but in many applications it may be too restrictive. Another way would be to refine the concept of NE in the spirit of the *perfectness* concept of Lecture 5. This could be done by perturbing the state dynamics by a sequence of independent random variables having full support probability measures on the state space, obtaining the NE of the resulting stochastic nonzero-sum dynamic game (NZSDG), then letting the intensity of the stochastic perturbations vanish, and see which of the infinitely many NE would survive this process. The resulting NE of the NZSDG would be the *perfect* one, which is also called *robust.* If all players have access to perfect state information, including memory, then another refinement would be to require that the NE of the original NZSDG also constitutes NE to any forward time-truncated one; such a NE is called a *strongly time consistent* one or one that is *sub-game perfect.* For this closed-loop information structure, robustness to vanishing stochastic perturbations actually leads to sub-game perfectness, and the resulting NE is called *closed-loop feedback* (CLFB). *Strong time consistency* captures here the property that if at any point in time any one of the players does not play according to the stipulated NE, the strategies are still in NE for the remaining portion of the game. A NE in general does not have such a strong property; it is only *weakly time consistent*, meaning that at any point in time, as long as the players have stuck by their NE policies in the past, there is no incentive for them to change their policies for the future [51].

In this section, we discuss the derivation of OL and CLFB NE for the class of NZSDGs formulated earlier in the Lecture. The former is weakly time consistent, whereas the latter is strongly time consistent.

11.3.1 Open-loop Nash equilibrium

One method of obtaining the OL NE solution(s) of the class of discrete-time games formulated is to view them as static infinite games and directly apply the results of Lecture 7. Toward this end we first note that it is possible to express L^i solely as functions of $\{u^j; j \in \mathbf{N}\}$ and the initial state x_1 whose value is known *a priori*, where u^j is defined as the aggregate control vector $(u_1^{j'}, u_2^{j'}, \ldots, u_K^{j'})'$. This then implies that, to every given set of functions $\{f_k, g_k^i; k \in \mathbf{K}\}$, there corresponds a unique function $\tilde{L}^i : X \times U^1 \times \cdots \times U^N \to \mathbf{R}$, which is the cost functional of $\mathbf{P}i$, $i \in \mathbf{N}$. Here, U^j denotes the aggregate control set of $\mathbf{P}j$, compatible with the requirement that if $u^j \in U^j$ then $u_k^j \in U_k^j$, $\forall k \in \mathbf{K}$, and the foregoing construction leads to a normal form description of the original game, which is no different from the class of infinite games treated in Lecture 7. Therefore, to obtain the open-loop Nash equilibria, we simply have to minimize $\tilde{L}^i(x_1, u^1, \ldots, u^{i-1}, \cdot, u^{i+1}, \ldots, u^N)$ over U^i, for each $i \in \mathbf{N}$, and then determine the intersection point(s) of the resulting reaction curves. In particular, if $\tilde{L}^i(x_1, u^1, \ldots, u^N)$ is continuous on $U^1 \times \cdots \times U^N$, strictly convex in u^i, and further if U^i are closed, bounded and convex, an open-loop Nash equilibrium (in pure strategies) exists.

Such an approach can sometimes lead to quite unwieldy expressions, especially if the number of stages in the game is large. An alternative derivation which partly removes this difficulty is the one that utilizes techniques of optimal control theory, by making explicit use of the stage-additive nature of the cost functionals and the specific structure of the extensive form description of the game, as provided by the state equation. There is in fact a close relationship between derivation of OL NE and the problem of solving (jointly) N optimal control problems, which can readily be observed from the inequalities defining the NE since each one of them describes an optimal control problem whose *structure* is not affected by the remaining players control vectors. Exploring this relationship a little further, we arrive at the following result.

Theorem 8 *For an N-person discrete-time infinite dynamic game, let*

(i) $f_k(\cdot, u_k^1, \ldots, u_k^N)$ be continuously differentiable on $\mathbf{R}^n$ $(k \in \mathbf{K})$

(ii) $g_k^i(\cdot, u_k^1, \ldots, u_k^N, \cdot)$ *be continuously differentiable on* $\mathbf{R}^n \times \mathbf{R}^n$ $(k \in \mathbf{K}, i \in \mathbf{N})$.

Then, if $\{\gamma^{i*}(x_1) = u^{i*}; i \in \mathbf{N}\}$ *provides an* open-loop Nash equilibrium solution *and* $\{x_{k+1}^*; k \in \mathbf{K}\}$ *is the corresponding state trajectory, there exists a finite sequence of n-dimensional (costate) vectors* $\{p_2^i, \ldots, p_{K+1}^i\}$ *for each* $i \in \mathbf{N}$ *such that the following relations are satisfied:*

$$x_{k+1}^* = f_k(x_k^*, u_k^{1*}, \ldots, u_k^{N*}), \quad x_1^* = x_1 \tag{78}$$

$$\gamma_k^{i*}(x_1) \equiv \arg \min_{u_k^i \in U_k^i} H_k^i(p_{k+1}^i, u_k^{1*}, \ldots, u_k^{i-1*}, u_k^i, u_k^{i+1*}, \ldots, u_k^{N*}, x_k^*) \tag{79}$$

$$\begin{aligned} p_k^i = \frac{\partial}{\partial x_k} f_k(x_k^*, u_k^{1*}, \ldots, u_k^{N*})' \Big[p_{k+1}^i + \Big(\frac{\partial}{\partial x_{k+1}} g_k^i(x_{k+1}^*, u_k^{1*}, \ldots, \\ u_k^{N*}, x_k^*\Big)'\Big] + \Big[\frac{\partial}{\partial x_k} g_k^i(x_{k+1}^*, u_k^{1*}, \ldots, u_k^{N*}, x_k^*)\Big]'; \\ p_{K+1}^i = 0, \quad i \in \mathbf{N}, \quad k \in \mathbf{K}, \end{aligned} \tag{80}$$

where

$$\begin{aligned} H_k^i(p_{k+1}, u_k^1, \ldots, u_k^N, x_k) \triangleq g_k^i(f_k(x_k, u_k^1, \ldots, u_k^N), u_k^1, \ldots, u_k^N, x_k) \\ + p_{k+1}^{i'} f_k(x_k, u_k^1, \ldots, u_k^N); \quad k \in \mathbf{K}, \quad i \in \mathbf{N}. \end{aligned} \tag{81}$$

Every such Nash equilibrium solution is weakly time consistent.

Proof Consider the NE inequality for $\mathbf{P}i$, which says that $\gamma^{i*}(x_1) \equiv u^{i*}$ minimizes $L^i(u^{1*}, \ldots, u^{i-1*}, u^i, u^{i+1*}, \ldots, u^{N*})$ over U^i subject to the state equation

$$x_{k+1} = f_k(x_k, u_k^{1*}, \ldots, u_k^{i-1*}, u_k^i, u_k^{i+1*}, \ldots, u_k^{N*}), \quad k \in \mathbf{K}.$$

But this is a standard optimal control problem for $\mathbf{P}i$ since u^{j*} $(j \in \mathbf{K}, j \neq i)$ are open-loop controls and hence do not depend on u^i. The result, then, follows directly from the minimum principle for discrete-time control systems [27]. ◇

Theorem 8 thus provides a set of necessary conditions (solvability of a set of coupled two-point boundary value problems) for the OL NE solution to satisfy; in other words, it produces candidate equilibrium solutions. In principle, one has to determine all solutions of this set of equations and further investigate which of these candidate solutions satisfy the original set of NE inequalities. If some further restrictions are imposed on f_k and g_k^i $(i \in \mathbf{N}, k \in \mathbf{K})$ so that the resulting cost functional $\tilde{L}^i$ (defined earlier, in this subsection) is convex in u^i for all $u^j \in U^j$, $j \neq i$, $j \in \mathbf{K}$, the latter phase of verification can clearly be eliminated, since then every solution set of (78)-(80) constitutes an OL NE solution. A specific class of problems for which this can be done, and the conditions involved expressed explicitly in terms of the parameters of the game, is the class of so-called "affine-quadratic" games which we first formally introduce below.

Definition 21 *An N-person discrete-time infinite dynamic game is of the* affine-quadratic *type if* $U_k^i = \mathbf{R}_i^m$ $(i \in \mathbf{N}, k \in \mathbf{K})$, *and* [28]

$$f_k(x_{k+1}, u_k^1, \ldots, u_k^N) = A_k x_k + \sum_{i \in \mathbf{N}} b_k^i u_k^i + c_k \tag{82}$$

$$g_k^i(x_{k+1}, u_k^1, \ldots, u_k^N, x_k) = \frac{1}{2}\left(x_{k+1}' Q_{k+1}^i x_{k+1} + \sum_{j \in \mathbf{N}} u_k^{j'} R_k^{ij} u_k^j \right) \tag{83}$$

where A_k, B_k, Q_{k+1}^i R_k^{ij} are matrices of appropriate dimensions, Q_{k+1}^i is symmetric, $R_k^{ii} > 0$, $c_k \in \mathbf{R}^n$ is a fixed vector sequence, and $k \in \mathbf{K}$, $i \in \mathbf{N}$.
An affine-quadratic game is of the linear quadratic *type if $c_k \equiv 0$.* $\diamond$

Theorem 9 *For an N-person affine-quadratic dynamic game with $Q_{k+1}^i \geq 0$ $(i \in \mathbf{N}, k \in \mathbf{K})$, let Λ_k, M_{k+1}^i $(k \in \mathbf{K}, i \in \mathbf{N})$ be appropriate dimensional matrices defined by*

$$\Lambda_k = I + \sum_{i \in \mathbf{N}} B_k^i [R_k^{ii}]^{-1} B_k^{i'} M_{k+1}^i \tag{84}$$

$$M_k^i = Q_k^i + A_k' M_{k+1}^i \Lambda_k^{-1} A_k; \quad M_{K+1}^i = Q_{K+1}^i. \tag{85}$$

If the matrices Λ_k $(k \in \mathbf{K})$, thus recursively defined, are invertible, the game admits a unique OL NE solution *given by*

$$\gamma_k^{i*}(x_1) \equiv u_k^{i*} = -[R_k^{ii}]^{-1} B_k^{i'} [M_{k+1}^i \Lambda_k^{-1} A_k x_k^* + \xi_k^i], \ (k \in \mathbf{K}, i \in \mathbf{N}) \tag{86}$$

where $\{x_{k+1}^*; k \in \mathbf{K}\}$ is the associated state trajectory determined from

$$x_{k+1}^* = \Lambda_k^{-1}[A_k x_k^* + \eta_k]; \quad x_1^* = x_1, \tag{87}$$

[28] The stagewise cost functions g_k^i can also to be taken to depend on x_k instead of x_{k+1}, but we prefer here the present structure (without any loss of generality) for convenience in the analysis to follow.

and ξ_k^i, η_k are defined by

$$\xi_k^i = M_{k+1}^i \Lambda_k^{-1} \eta_k + m_{k+1}^i\,, \tag{88}$$

$$\eta_k = c_k - \sum_{j\in\mathbf{N}} B_k^i [R_k^{ii}]^{-1} B_k^{i'} m_{k+1}^i\,, \tag{89}$$

with m_k^i recursively generated by

$$m_k^i = A_k'[m_{k+1}^i + M_{k+1}^i \Lambda_k^{-1}\eta_k]\,,\; m_{K+1}^i = 0,\; i\in\mathbf{N}, k\in\mathbf{K}. \tag{90}$$

Proof Since $Q_{k+1}\geq 0$, $R_k^{ii}>0$, $\tilde{L}^i(x_1,u^1,\ldots,u^N)$ is a strictly convex function of u^i for all $u^j\in\mathbf{R}^{m_jK}$ $(j\neq i, j\in\mathbf{N})$ and for all $x_1\in\mathbf{R}^n$. Therefore, every solution set of (78)-(80) provides an OL NE. Hence, the proof will be completed if we can show that (86) is the only candidate solution. First note that

$$\begin{aligned} H_k^i &= \frac{1}{2}\left(A_kx_k + c_k + \sum_{j\in\mathbf{N}} B_k^j u_k^j\right)' Q_{k+1}^i \left(A_kx_k + c_k + \sum_{j\in\mathbf{N}} B_k^j u_k^j\right) \\ &\quad + \frac{1}{2}\sum_{j\in\mathbf{N}} u_k^{j'} R_k^{ij} u_k^j + p_{k+1}^{i'}\left(A_kx_k + c_k + \sum_{j\in\mathbf{N}} B_k^j u_k^j\right), \end{aligned}$$

and since $Q_{k+1}^i\geq 0$, $R_k^{ii}>0$, minimization of this "Hamiltonian" over $u_k^i\in\mathbf{R}^{m_i}$ yields the unique relation

$$u_k^{i*} = -[R_k^{ii}]^{-1}B_k^{i'}[p_{k+1}^i + Q_{k+1}^i x_{k+1}^*], \tag{i}$$

where

$$x_{k+1}^* = A_kx_k^* + c_k + \sum_{i\in\mathbf{N}} B_k^i u_k^{i*};\quad x_1^* = x_1. \tag{ii}$$

Furthermore, the costate (difference) equation in (80) reads

$$p_k^i = A_k'[p_{k+1}^i + Q_{k+1}^i x_{k+1}^*];\quad p_{K+1}^i = 0. \tag{iii}$$

Let us start with $k=K$, in which case (i) becomes

$$u_K^{i*} = -[R_K^{ii}]^{-1}B_K^{i'}M_{K+1}^i x_{K+1}^*, \tag{iv}$$

and if both sides are first premultiplied by B_K^i and then summed over $i\in\mathbf{N}$ we obtain, by also making use of (ii) and (84),

$$x_{K+1}^* - A_Kx_K^* = (I-\Lambda_K)x_{K+1}^* + c_K$$

which further yields the unique relation

$$x^*_{K+1} = \Lambda_K^{-1}[A_K x^*_K + c_K]$$

which is precisely (87) for $k = K$. Substitution of this relation into (iv) then leads to (86) for $k = K$.

We now prove by induction that the unique solution set of (i)-(iii) is given by (86)-(87) and $p^i_k = A'_k[M^i_{k+1}x^*_{k+1} + m^i_{k+1}]$ $(i \in \mathbf{N}, k \in \mathbf{K})$. Let us assume that this is true for $k = l+1$ (already verified for $l = K-1$) and prove its validity for $k = l$.

First, using the solution $p^i_{l+1} = A'_{l+1}[M^i_{l+2}x^*_{l+2} + m^i_{l+2}]$ in (i) with $k = l$, we obtain, after several algebraic manipulations,

$$u^{i*}_l = -[R^{ii}_l]^{-1}B^{i'}_l[M^i_{l+1}x^*_{l+1}m^i_{l+1}]. \qquad (v)$$

Again, premultiplying this expression by B^i_l, and summing it over $i \in \mathbf{N}$ leads to, also in view of (ii), (84), and (89),

$$x^*_{l+1} = \Lambda_l^{-1}[A_l x^*_l + \eta_l]$$

which is (87). If this relation is used in (v), we obtain the unique control vectors (86), for $k = l$, and if it is further used in (iii) we obtain, in view of (90),

$$p^i_l = A'_l[M^i_{l+1}x^*_{l+1} + m^i_{l+1}]\,.$$

This then closes the induction argument, and thereby completes the proof of the theorem. $\diamond$

Remark 11 An alternative derivation for the OL NE solution of the affine-quadratic game is (as discussed earlier in this subsection in a general context) to convert it into a standard static quadratic game and then to make use of the available results on such games (cf. Proposition 13). By backward recursive substitution of the state vector from the state equation into the quadratic cost functionals, it is possible to bring the cost functional of $\mathbf{P}i$ into the structural form as given by (44), which further is strictly convex in $u^i = (u^{i'}_1, \ldots, u^{i'}_K)'$ because of assumptions $Q^i_k \geq 0$, $R^{ii}_k > 0$, $(k \in \mathbf{K})$.[29] Consequently, each player has a unique reaction curve, and the condition for existence

[29]The condition $Q^i_{k+1} \geq 0$ is clearly sufficient (along with $R^{ii}_k > 0$) to make $\tilde{L}^i$ strictly convex in u^i, but is by no means necessary. It can be replaced by weaker conditions (which ensure convexity) under which the statements of Theorem 9 and this remark are still valid; for details, see [1].

of a unique NE becomes equivalent to the condition for unique intersection of these reaction curves (cf. Proposition 12). The existence condition of Theorem 9, *i.e.* nonsingularity of Λ_k, $(k \in \mathbf{K})$, is precisely that condition, but expressed in a different (more convenient, recursive) form. $\diamond$

11.3.2 Closed-loop feedback Nash equilibrium

The imposition of sub-game perfectness or strong time consistency for NE solution directly leads to a recursive derivation which involves solutions of static N-person nonzero-sum games at every stage of the dynamic game. As a direct consequence of sub-game perfectnes, the feedback equilibrium solution depends only on x_k at stage k, and dependence on x_1 is only at stage $k = 1$.[30] By utilizing these properties, we readily arrive at the following theorem.

Theorem 10 *For an N-person discrete-time infinite dynamic game, the set of strategies $\{\gamma_k^{1*}(x_k); k \in \mathbf{K}, i \in \mathbf{N}\}$ provides a* feedback Nash equilibrium solution *if, and only if, there exist functions $V^i(k,\cdot)$: $\mathbf{R}^n \to \mathbf{R}, k \in \mathbf{K}, i \in \mathbf{N}$, such that the following recursive relations are satisfied:*

$$\begin{aligned} V^i(k,x) &= \min_{u_k^i \in \mathbf{U}_k^i} \; g_k^i(\tilde{f}_k^{i*}(x,u_k^i), \gamma_k^{1*}(x), \ldots, \gamma_k^{i-1*}(x), u_k^i, \gamma_k^{i+1*}(x) \\ &\qquad\qquad \ldots, \gamma_k^{N*}(x), x) + V^i(k+1, \tilde{f}_k^{i*}(x,u_k^i)] \\ &= g_k^i(\tilde{f}_k^{i*}(x,\gamma_k^{i*}(x)), \gamma_k^{1*}(x), \ldots, \gamma_k^{N*}(x), x) \\ &\quad + V^i(k+1, \tilde{f}_k^{i*}(x,\gamma_k^{i*}(x)); \quad V^i(K+1,x) = 0, \quad i \in \mathbf{N}, \end{aligned} \tag{91}$$

where

$$\tilde{f}_k^{i*}(x,u_k^i) \triangleq f_k(x, \gamma_k^{1*}(x), \ldots, \gamma_k^{i-1*}(x), u_k^i, \gamma_k^{i+1*}(x), \ldots, \gamma_k^{N*}(x)).$$

Every such equilibrium solution is strongly time consistent*, and the corresponding Nash equilibrium cost for $\mathbf{P}i$ is $V^i(1,x_1)$.*

Proof Let us start with the truncated game which has only one stage to go. Since we are looking for a NE for that game which is valid for all $\gamma_k^i \in \Gamma_k^i$, $i \in \mathbf{N}$, $k \leq K-1$, this necessarily implies that the NE property will have to hold for all values of state x_k which are reachable by utilization of some combination of these strategies. Let us denote that subset of $\mathbf{R}^n$ by X_K. Then, the corresponding set of Nash inequalities becomes equivalent to the problem of seeking Nash equilibria of an N-person static game with cost functionals

$$g_K^i(f_K(x_K, u_K^1, \ldots, u_K^N), u_K^1, \ldots, u_K^N, x_K), \quad i \in \mathbf{N}, \tag{i}$$

[30] This statement is valid also under the "closed-loop perfect state" information pattern. Note that the feedback equilibrium solution retains its equilibrium property also under the feedback information pattern.

which should be valid for all $x_K \in X_K$. This is precisely what (91) says for $k = K$, with a set of associated Nash equilibrium controls denoted by $\{\gamma_K^{i*}(x_K); i \in \mathbf{N}\}$ since they depend explicitly on $x_K \in X_K$, but not on the past values of the state (including the initial state x_1). Now, with these strategies substituted into (i), and looking into a 2-stage truncated game, a similar argument (as above) leads to the conclusion that we now have to solve a static Nash game with cost functionals

$$V^i(K, x_K) + g_{K-1}^i(x_K, u_{K-1}^1, \ldots, u_{K-1}^N, x_{K-1}), \quad i \in \mathbf{N},$$

where

$$x_K = f_{K-1}(x_{K-1}, u_{K-1}^1, \ldots, u_{K-1}^N),$$

and the Nash solution has to be valid for all $x_{K-1} \in X_{K-1}$ (where X_{K-1} is the counterpart of X_K at stage $k = K-1$). Here again, we observe that the Nash equilibrium controls can only be functions of x_{K-1}, and (91) with $k = K-1$ provides a set of necessary and sufficient conditions for $\{\gamma_{K-1}^{i*}(x_{K-1}): i \in \mathbf{N}\}$ to solve this static Nash game. The theorem then follows from a standard induction argument. Note that the "strong time consistency" property of the feedback Nash equilibrium, and the expression for the corresponding cost for each player, are direct consequences of the recursive nature of the construction of the solution. $\diamond$

The following corollary, which is the counterpart of Theorem 9 in the case of feedback Nash equilibrium, now follows as a special case of Theorem 10.

Preliminary Notation for Corollary 11.1. Let P_k^i ($i \in \mathbf{N}, k \in \mathbf{K}$) be appropriate dimensional matrices satisfying the set of linear matrix equations

$$[R_k^{ii} + B_k^{i'} Z_{k+1}^i B_k^i] P_k^i + B_k^{i'} Z_{k+1}^i \sum_{\substack{j \in \mathbf{N} \\ j \neq i}} B_k^j P_k^j = B_k^{i'} Z_{k+1}^i A_k, \; i \in \mathbf{N}, \tag{92}$$

where Z_k^i ($i \in \mathbf{N}$) are obtained recursively from

$$Z_k^i = F_k' Z_{k+1}^i F_k + \sum_{j \in \mathbf{N}} P_k^{j'} R_k^{ij} P_k^j + Q_k^i; \; Z_{K+1}^i = Q_{K+1}^i, \; i \in \mathbf{N}, \tag{93}$$

and

$$F_k \triangleq A_k - \sum_{i \in \mathbf{N}} B_k^i P_k^i, \quad k \in \mathbf{K}. \tag{94}$$

Furthermore, let $\alpha_k^i \in \mathbf{R}^{m_i}$ $(i \in \mathbf{N}, k \in \mathbf{K})$ be vectors satisfying the set of linear equations:

$$[R_k^{ii} + B_k^{i'} Z_{k+1}^i B_k^i]\alpha_k^i + B_k^{i'} Z_{k+1}^i \sum_{\substack{j \in \mathbf{N} \\ j \neq i}} B_k^j \alpha_k^j = B_k^{i'}(\zeta_{k+1}^i + Z_{k+1}^i c_k),\ i \in \mathbf{N}, \tag{95}$$

where ζ_k^i $(i \in \mathbf{N})$ are obtained recursively from

$$\zeta_k^i = F_k'(\zeta_{k+1}^i + Z_{k+1}^i \beta_k) + \sum_{j \in \mathbf{N}} P_k^{j'} R_k^{ij} \alpha_k^j;\ \zeta_{K+1}^i = 0, i \in \mathbf{N}, \tag{96}$$

and

$$\beta_k \triangleq c_k - \sum_{j \in \mathbf{N}} B_k^{j'} \alpha_k^j, \quad k \in \mathbf{K}. \tag{97}$$

Finally, let $n_k^i \in \mathbf{R}$ $(i \in \mathbf{N}$, $k \in \mathbf{K})$ be generated by

$$n_k^i = n_{k+1}^i + \frac{1}{2}|\beta_k|^2_{Z_{k+1}^i} + \zeta_{k+1}^{i'} \beta_k + \frac{1}{2} \sum_{j \in \mathbf{N}} |\alpha_k^j|^2_{R_k^{ij}},\ n_{K+1}^i = 0. \tag{98}$$

Corollary 11.1 *An N-person affine-quadratic dynamic game (cf. Definition 21) with $Q_{k+1}^i \geq 0$ $(i \in \mathbf{N}, k \in \mathbf{K})$ and $R_k^{ij} \geq 0$ $(i, j \in \mathbf{N}, j \neq i, k \in \mathbf{K})$ admits a* unique feedback Nash equilibrium solution *if, and only if, (92) and (95) admit unique solution sets $\{P_k^{i*}; i \in \mathbf{N}, k \in \mathbf{K}\}$ and $\{\alpha_k^{i*}; i \in \mathbf{N}, k \in \mathbf{K}\}$, respectively, in which case the equilibrium strategies are given by*

$$\gamma_k^{i*}(x_k) = -P_k^{i*} x_k - \alpha_k^{i*} \quad (k \in \mathbf{K}, i \in \mathbf{N}), \tag{99}$$

and the corresponding feedback Nash equilibrium cost for each player is:

$$J^i(\gamma^{1*}, \ldots, \gamma^{N*}) = V^i(1, x_1) = \frac{1}{2}|x_1|^2_{Z_1^i} + \zeta_1^{i'} x_1 + n_1^i, \quad (i \in \mathbf{N}). \tag{100}$$

Proof Starting with $k = K$ in the recursive equation (91), we first note that the functional to be minimized (for each $i \in \mathbf{N}$) is strictly convex, since $R_K^{ii} + B_K^{i'} Q_{K+1}^i B_K^i > 0$. Then, the first order necessary conditions for minimization are also sufficient and therefore we have (by differentiation) the unique set of equations

$$\begin{aligned} -[R_K^{ii} + B_K^{i'} Q_{K+1}^i B_K^i]\gamma_K^{i*}(x_K) - B_K^{i'} Q_{K+1}^i \textstyle\sum_{j \neq i}^{j \in \mathbf{N}} B_K^j \gamma_K^{j*}(x_k) \\ = B_K^{i'} Q_{K+1}^i [A_K x_K + c_K]; \quad i \in \mathbf{N} \end{aligned}$$

which readily leads to the conclusion that any set of Nash equilibrium strategies at stage $k = K$ has to be affine in x_K. Therefore, by substituting $\gamma_K^{i*} = -P_K^i x_K - \alpha_K^i$ $(i \in \mathbf{N})$ into the foregoing

equation, and by requiring it to be satisfied for all possible x_K, we arrive at (92) and (95) for $k = K$. Further substitution of this solution into (91) for $k = K$ leads to $V^i(K,x) = \frac{1}{2}x'(Z^i_K - Q^i_K)x + \zeta^{i'}_K x + n_K$; that is, $V^i(K,\cdot)$ has a quadratic structure at stage $k = K$. Now, if this expression is substituted into (91) with $k = K-1$, and the outlined procedure is carried out for $k = K-1$, and this so (recursively) for all $k \geq K-1$, one arrives at the conclusion that
(i) $V^i(k,x) = \frac{1}{2}x'(Z^i_k - Q^i_k)x + \zeta^{i'}_k x + n_k$ is the unique solution of the recursive equation (91) under the hypothesis of the corollary and by noting that $Z^i_k \geq 0$ $(i \in \mathbf{N}, k \in \mathbf{K})$, and
(ii) the minimization operation in (91) leads to the unique solution (99) under the condition of unique solvability of (92) and (95).

The expression for the cost, (100) follows directly from the expression derived for the "cost-to-go" $V^i(k,x)$. This, then, completes verification of Corollary 11.1. $\diamond$

Remark 12 The result of Corollary 11.1 as well as the verification given above extends readily to more general affine-quadratic dynamic games where the cost functions of the players contain additional terms that are linear in x_k, that is with g^i in (83) replaced by

$$g^i_k(x_{k+1}, u^1_k, \ldots, u^N_k, x_k) = \frac{1}{2}(x'_{k+1}[Q^i_{k+1}x_{k+1} + 2l^i_{k+1}] + \sum_{j\in\mathbf{N}} u^{j'}_k R^{ij}_k u^j_k),$$

where l^i_{k+1} $(k \in \mathbf{K})$ is a known sequence of n-dimensional vectors for each $i \in \mathbf{N}$. Then, the statement of Corollary 11.1 remains intact, with only the equation (96) that generates ζ^i_k now reading:

$$\zeta^i_k = F'_k(\zeta^i_{k+1} + Z^i_{k+1}\beta_k) + \sum_{j\in\mathbf{N}} P^{j'}_k R^{ij}_k \alpha^j_k + l_k;\ \zeta^i_{K+1} = l_{K+1}, i \in \mathbf{N},$$

and the cost-to-go functions admitting the compatibly modified form

$$V^i(k,x) = \frac{1}{2}x'(Z^i_k - Q^i_k)x + (\zeta^i_k - l^i_k)'x + n_k, \quad i \in \mathbf{N}.$$

$\diamond$

Remark 13 The "nonnegative definiteness" requirements imposed on Q_{k+1} and $R^{ij}_k(i,j \in \mathbf{N}, j \neq i; k \in \mathbf{K})$ are sufficient for strict convexity of the functionals to be minimized in (91), but they are by no means necessary. A set of less stringent (but more indirect) conditions would be

$$R_k^{ii} + B_k^{i'} Z_{k+1}^i B_k^i > 0 \quad (i \in \mathbf{N}, k \in \mathbf{K}),$$

under which the statement of Corollary 11.1 still remains valid. Furthermore, it follows from the proof of Corollary 11.1 that, if (92) admits more than one set of solutions, every such set constitutes a feedback Nash equilibrium solution, which is also strongly time consistent. $\diamond$

Remark 14 It is possible to give a precise condition for the unique solvability of the sets of equations (92) and (95) for P_k^i and α_k^i ($i \in \mathbf{N}, k \in \mathbf{K}$), respectively. The said condition (which is the same for both) is the invertability of matrices Φ_k, $k \in \mathbf{K}$, which are composed of block matrices, with the iith block given as $R_k^{ii} + B_k^{i'} Z_{k+1}^i B_k^i$ and the ijth block as $B_k^{i'} Z_{k+1}^i B_k^j$, where $i, j \in \mathbf{N}$, $j \neq i$. $\diamond$

12 Lecture 12: Dynamic Stackelberg games: Incentive strategies and mechanism design

12.1 Deterministic incentives

The idea of declaring a reward (or punishment) for a decision maker **P**1 according to his particular choice of action in order to induce a certain 'desired' behavior on the part of another decision maker **P**2 is known as an *incentive* (or in case of the punishment, as a threat). Mathematical formulation and analysis of such decision problems bear strong connections with the theory of Stackelberg games presented earlier, which is what we will be discussing in this section, for deterministic scenarios. Counterparts of these results in the stochastic case will be presented in the next section. Following the earlier convention, we call, in the above scenario, **P**1 the leader and **P**2 the follower. Then, the action outcome desired by the leader is:

$$(u^{1^t}, u^{2^t}) = \arg \min_{u^1 \in S^1, u_2 \in S^2} L^1(u^1, u^2). \tag{101}$$

The incentive problem can now be stated as: find a $\gamma^1 \in \Gamma^1$, where Γ^1 is an admissible subclass of all mappings from S^2 into S^1, such that

$$\begin{aligned} \arg\min_{u^2} L^2(\gamma^1(u^2), u^2) &= u^{2^t}, & (102) \\ \gamma^1(u^{2^t}) &= u^{1^t}. & (103) \end{aligned}$$

Note that (102) and (103) require choosing a set of m_1 scalar functions which together map S^2 into S^1 so as to satisfy $m_1 + m_2$ equations. If this set of m_1 functions has $m_1 + m_2$ or more parameters then we might in general accomplish this by choosing the parameters appropriately.

Incentive problems do arise in real life decision making. Think of **P**1 as a government and of **P**2 as a citizen. The income tax which **P**2 has to pay is a fraction (say k) of his income (before taxation) u^2. The amount of money that the government receives is $u^1 = ku^2$. It is up to **P**2 how hard to work and thus how much money to earn. The incentive here is $u^1 = \gamma^1(u^2) = ku^2$. The government will choose k so as to achieve certain goals, but it cannot choose its own income, u^1, directly. In reality, the γ^1-functions will often be nonlinear, but that does not take away the incentive phenomenon.

Example. Consider $L^1 = (u^1)^2 + (u^2)^2$ and $L^2 = (u^1 - 1)^2 + (u^2 - 1)^2$, where the u^i are scalars. By inspection, $u^{1^t} = u^{2^t} = 0$. Consider the choice $u^1 = ku^2$, with k approaching ∞ if necessary, as a possible incentive mechanism for $\mathbf{P}1$. The idea is that any choice of $u^2 \neq 0$ will make L^2 approach ∞ if $k \to \infty$ and thus force $\mathbf{P}2$ to choose u^2 arbitrarily close to u^{2^t} in his own interest. However, by substituting $u^1 = ku^2$ into L^2, it is easily shown that the minimizing action is $u^2 = (k+1)/(k^2+1)$ and consequently $u^1 = (k^2+k)/(k^2+1)$. Thus u^2 approaches $u^{2^t} = 0$ and u^1 approaches 1 (which is different from u^{1^t}) as k approaches ∞ and hence (103) is violated. Consequently 'infinite threat' as just described is not generally feasible. Besides, such a threat may not be credible in practice.

Let us now consider an incentive γ^1 of the form

$$u^1 = \gamma^1(u^2) = u^{1^t} + g(u^2, u^{2^t}), \tag{104}$$

where g is a function which satisfies $g(u^{2^t}, u^{2^t}) = 0$. With this restriction on g, equation (103) is automatically satisfied. Let us try the linear function $g = k(u^2 - u^{2^t})$. Equation (102) reduces to $(k+1)/(k^2+1) = 0$ and hence k must be equal to -1. Graphically, the incentive $u^1 = ku^2 = -u^2$ is a line through the team solution (u^{1^t}, u^{2^t}) and has only this point in common with the set of points (u^1, u^2) defined by $L^2(u^1, u^2) \leq L^2(u^{1^t}, u^{2^t})$. By announcing the γ^1-function, $\mathbf{P}1$ ensures that the solution (u^1, u^2) will lie on the line $u^1 = -u^2$ in the (u^1, u^2) plane, independent of the action of $\mathbf{P}2$. Being rational, $\mathbf{P}2$ will choose that point on this line which minimizes his cost function; such a choice is $u^2 = 0$. The relationship of this with the derivation of the Stackelberg solution discussed earlier should be clear now. Particularly, any (nonlinear, continuous or discontinuous) incentive policy $u^1 = \gamma^1(u^2)$ which passes through the point characterized by the team solution and has only this point in common with the set just described, will lead to the team solution for $\mathbf{P}1$. If we restrict ourselves to linear incentives, then the solution is unique in this example.

This example exhibits yet another feature. Note that with $k = -1$,

$$L^2(ku^2, u^2) = 2(u^2)^2 + 2 = L^1(u^{1^t}, u^2) + 2.$$

In other words, by this choice of incentive, the objectives of both players are identical (apart from a constant), thus fulfilling the old adagium 'if you wish other people to behave in your own interest, then make them see things your way'. In general making the cost functions identical by the choice of an appropriate γ^1-function will be too strong a requirement. A weaker form, however, which also leads to the team solution, is

$$\arg\min_{u^2} L^2(\gamma^1(u^2), u^2) = \arg\min_{u^2} L^1(u^{1^t}, u^2), \tag{105}$$

where it is assumed that γ^1 is of the form as described by (104). $\diamond$

Definition 22 *The incentive problem, as defined in this section, is* (linearly) incentive controllable*, if there exists a (linear) γ^1-function such that (102) and (103) are satisfied.* $\diamond$

Of course, not all problems are incentive controllable. What can **P**1 achieve in problems that are not incentive controllable? The method to be employed to answer this question will be described in the context of an example, given below. It should then be clear how the method would apply to more elaborate problems.

Example. Consider $L^1 = (u^1-4)^2+(u^2-4)^2$ and $L^2 = (u^1)^2+(u^2-1)^2$, where the u^i are scalars; $u^1 \in S^1 = [0,3]$ and $u^2 \in S^2 = [0,6]$. The team solution in this case is: $u^{1^t} = 3, u^{2^t} = 4$, and $L^1(u^{1^t}, u^{2^t}) = 2$. The worst possible outcome for **P**2, even if he minimizes his cost function with respect to his own decision variable, is

$$\min_{u^2\in S^2}[\max_{u^1\in S^1} L^2(u^1,u^2)].$$

This occurs for $u^2 = 1, u^1 = 3$ (call this point A in the plane), and $L^2(3,1) = 9$. Whatever choice **P**1 makes for γ^1, the cost for **P**2 will never be higher than 9. If **P**1 chooses $u^1 = \gamma^1(u^2) = 3$ on the interval $[0,6]$, then the outcome becomes $u^2 = 1, u^1 = 3$, (point A), and the costs for **P**1 and **P**2 become 9 and 10, respectively. This, however, is not optimal for **P**1. He should instead consider $\min_{u^1} L^1$ subject to $L^2 \leq 9$. (We will refer to the region defined by this inequality constraint as shaded region in the discussion below.) The solution of this minimization problem is $u^1 = 12/5, u^2 = 13/3$ (call this point B in the plane). Now, any γ^1-curve, in the rectangle $0 \leq u^1 \leq 3, 0 \leq u^2 \leq 6$, which has only the points A and B with the shaded region in common, would lead to a nonunique choice for **P**2; he might either choose $u^2 = 1$ or $u^2 = 13/3$. Both choices lead to $L^2 = 9$. The costs for **P**1 are respectively 9 and 4 for these choices. Therefore, **P**1 will choose a γ^1 function as just described, with one exception; it will have a little 'dip' in the shaded area near point B, such that the choice of **P**2 will be unique again. (A possible choice is: $\gamma^1(u^2) = 3$ for $0 \leq u^2 < 13/3 - \varepsilon$, where $\varepsilon > 0$ and $\gamma^1(u^2) = 12/5$ for $13/3 - \varepsilon \leq u^2 \leq 6$.) The outcome will now be a point u^1, u^2 near point B, just within the shaded area. **P**1 can keep his costs arbitrarily close to 4 (but not equal to 4). $\diamond$

Extensions of the foregoing analysis are possible in different directions, such as the multi-stage problems or problems with multiple hierarchies; another possibility is the many-followers case; these extensions can be found in [1].

12.2 Stochastic incentives

We start with two examples (two related games) and then proceed with some general results.

Example. Consider the 2-stage scalar stochastic dynamic game described by the state equations

$$\left.\begin{array}{rcl} x_2 & = & x_1 - u^2 + \theta_1 \\ x_3 & = & x_2 - u^1 + \theta_2 \end{array}\right\} \tag{106}$$

and cost functionals

$$\begin{aligned} L^1 &= (x_3)^2 + 2(u^1)^2 + (u^2)^2 \\ L^2 &= (x_3)^2 + (u^2)^2. \end{aligned}$$

Here, θ_1 and θ_2 are taken as independent random variables with mean zero and variances σ_1 and σ_2, respectively. The leader (**P**1) acts at stage 2 and has access to both x_1 and x_2, while the follower (**P**2) acts at stage 1 and has only access to x_1. We approach this problem within the context of SES after bringing it to normal form.

If $\gamma^i \in \Gamma^i$ denotes a general strategy of $\mathbf{P}i$ $(i = 1, 2)$, the expected (average) cost functional of **P**2 can be written as

$$\begin{aligned} J^2(\gamma^1, \gamma^2) &= E\{[x_2 - \gamma^1(x_2, x_1) + \theta_2]^2 + [\gamma^2(x_1)]^2\} \\ &= E\{[x_2 - \gamma^1(x_2, x_1)]^2 + [\gamma^2(x_1)]^2\} + \sigma_2 \end{aligned}$$

which has to be minimized over $\gamma^2 \in \Gamma^2$, to determine the optimal response of **P**2 to $\gamma^1 \in \Gamma^1$. Because of some difficulties involved in working with a general γ^1, we restrict the investigation here to a subclass of strategies in Γ^1 which are affine in x_2, that is, to strategies of the form[31]

$$\gamma^1(x_2, x_1) = \alpha x_2 + \beta x_1 \tag{107}$$

where α and β are free parameters which are yet to be determined. They will, in general, be dependent on x_1 which is, though, known *a priori* by both players.

[31] The apparent linear structure of the second term below is adopted for the sake of convenience in the analysis to follow; it could also have been taken as a single function β.

Under the structural restriction (107), J^2 admits a unique minimum, thus leading to the optimal response strategy (for the follower)

$$\gamma^{2\circ}(x_1) = \frac{(1-\alpha)(1-\alpha-\beta)}{1+(1-\alpha)^2} x_1 \tag{108}$$

which explicitly depends on the parameters α and β that characterize the leader's strategy. To determine their optimal values, we now substitute (107)-(108) into $J^1(\gamma^1, \gamma^2)$, together with the corresponding values of x_3 and x_2 from (106), to obtain the function F given below, which has to be minimized over $\alpha \in \mathbf{R}$, $\beta \in \mathbf{R}$ for fixed x_1:

$$F(\alpha, \beta) = \left\{ \frac{[1-\alpha-\beta]^2}{1+(1-\alpha)^2} + \frac{2[\alpha+2\beta-\alpha\beta]^2}{[1+(1-\alpha)^2]^2} \right\} x_1^2 + [(1-\alpha)^2 + \alpha^2]\sigma_1 + \sigma_2.$$

Let us now note that

(i) F is jointly continuous in (α, β), $F(\alpha, \beta) \geq 0$ $\forall (\alpha, \beta) \in \mathbf{R} \times \mathbf{R}$, and $F(\alpha, \beta) \to \infty$ as $|\alpha|, |\beta| \to \infty$. Therefore, we can restrict our search for a minimum on $\mathbf{R}^2$ to a closed and bounded subset of $\mathbf{R}^2$, and consequently there exists (by the Weierstrass theorem at least one pair (α^*, β^*) that minimizes F for any given pair (x_1, σ_1).

(ii) The optimum pair (α^*, β^*) depends on (x_1, σ_1), but not on σ_2, and it cannot be expressed analytically as a function of (x_1, σ_1). Hence, for each fixed (x_1, σ_1), $F(\alpha, \beta)$ has to be minimized numerically.

(iii) With (α^*, β^*) determined as above, the linear-in-x_2 strategy $\gamma^{1*}(x_2, x_1) = \alpha^*(x_1)x_2 + \beta^*(x_1)x_1$ is only a suboptimal Stackelberg strategy for the leader, since he may possibly achieve a better performance by announcing a strategy outside the "linear-in-x_2" class.

$\diamond$

We now ask the question: What if the leader has also access to the follower's past control actions, in addition to the state information based on which these actions were obtained? In this enlarged information structure, a given strategy of the leader will have multiple representations, thus opening the possibility of enforcement of a team solution (to the leader's advantage) by selecting an appropriate representation of the team-optimal strategy (of the leader). To illustrate this line of thought now, let us revisit the previous example, but with this enlarged information structure:

Example (continued). Consider the 2-stage scalar stochastic dynamic game example analyzed, but with the enlarged information structure that allows **P**1 to have access to u^2, in addition to x_1 and x_2. Note that since u^2 depends only on x_1, this enlarged information structure carries the same statistical information as the earlier one for each fixed (pure) policy of the follower; however, as we will see shortly, the informational redundancy that it generates will bring a substantial advantage to the leader.

In the spirit of the analysis of the example in the previous section for the deterministic case, let us first determine the best performance the leader would achieve if the follower were cooperating with him (in the minimization of the leader's expected cost function). The associated team problem is :

$$J^t \triangleq J^1(\gamma^{1^t}, \gamma^{2^t}) = \min_{\gamma^1 \in \Gamma^1} \min_{\gamma^2 \in \Gamma^2} J^1(\gamma^1, \gamma^2)$$

where

$$J^1(\gamma^1, \gamma^2) = E\{[x_2 - \gamma^1(x_2, x_1) + \theta_2]^2 + 2[\gamma^1(x_2, x_1)]^2 + [\gamma^2(x_1)]^2\}$$

Here, the cost shows dependence on $u^2 = \gamma^2(x_1)$ not only directly, but also through x_2 as given by (106), which has to be taken into account in the minimization. Furthermore, the strategy spaces Γ^1 and Γ^2 are taken as before, since the additional knowledge of u^2 for **P**1 does not help in further reducing the minimum team cost J^t. Now, this team problem is in fact a standard LQ stochastic control problem, and its solution can readily be obtained as:

$$\gamma^{1^t}(x_2, x_1) = \frac{1}{3}x_2, \quad \gamma^{2^t}(x_1) = \frac{5}{14}x_1 \tag{109}$$

which is the *unique* minimizing pair in $\Gamma^1 \times \Gamma^2$. It is not, however, unique in the enlarged strategy space for the leader, as (for example) the following parametrized strategy also constitutes an optimal solution, along with γ^{2^t} given above, for every $\alpha \in \mathbf{R}$:

$$\gamma^1_\alpha(x_2, x_1, u^2) = \frac{1}{3}x_2 + \alpha(u^2 - \frac{5}{14}x_1). \tag{110}$$

This in fact characterizes the complete class of linear (in x_2, x_1, u^2) optimal strategies, but of course there are also nonlinear ones—all leading to the same (minimum) expected value for the leader. We will refer to all these "minimum expected cost achieving" strategies *representations* of γ^{1^t} under the team-optimal solution $(\gamma^{1^t}, \gamma^{2^t})$. This is a rich family of strategies, among which we seek one with the additional property that if the follower instead minimizes his own expected cost function, then the strategy in Γ^2 that achieves this minimum is still γ^{2^t}. The corresponding strategy (representation) for the leader would then clearly constitute a global Stackelberg solution, leading to the best possible performance for him.

Let us now conduct the search in the family of linear representations (110), which leads to the quadratic optimization problem:

$$\min_{\gamma^2\in\Gamma^2} E\{[x_2 - \gamma^1_\alpha(x_2, x_1, u^2) + \theta_2]^2 + [u^2]^2\},$$

where

$$x_2 = x_1 - u^2 + \theta_1.$$

Since x_1 is independent of θ_1 and θ_2, which have zero mean, this problem is equivalent to the following deterministic optimization problem:[32]

$$\min_{v\in\ \mathbf{R}}\{[(\frac{2}{3})(x_1 - v) - \alpha(v - \frac{5}{14}x_1)]^2 + v^2\}$$

where we have written v for u^2, to simplify the notation. Now, a simple optimization shows that for the value of $\alpha = 8/27$, this optimization problem admits the unique solution $v = (5/14)x_1 \equiv \gamma^{2^t}(x_1)$, and hence the policy pair

$$\gamma^1(x_2, x_1, u^2) = \frac{1}{3}x_2 + \frac{8}{27}(u^2 - \frac{5}{14}x_1),\ \gamma^2(x_1) = \frac{5}{14}x_1$$

provides a SES. This is in fact the unique such solution in the linear class. ⋄

Stochastic decision problems of the type above, where the leader is allowed to have access to past actions of the follower are known as *stochastic incentive problems*. In these incentive problems, the information structure may not always be *nested* (for the leader), as in the example above, where the leader has access to all the information that the follower has access to (plus more). If, for instance, the leader has only access to x_2 and u^2, then we have a problem with a *nonnested* information structure, to which the methodology presented above does not apply, since the dynamic information for the leader no longer exhibits redundancy. Discussion of such problems, where the follower possesses *private information* not known to the leader, is beyond the scope of our coverage here; see *Ho, Luh and Olsder (1982)* and *Başar (1984, 1989a))*. For stochastic incentive problems

[32]This equivalence holds as far as its optimum solution goes (which is what we seek), but not for the corresponding minimum values.

with nested information, however, the methodology used in the Example can be developed into a general procedure as briefly discussed below for a special class of such problems.

Consider a two-person stochastic incentive problem with cost functions $L^1(u^1, u^2; \theta)$ and $L^1(u^1, u^2; \theta)$, for **P**1 (leader) and **P**2 (follower), respectively, where θ is some random vector with a known distribution function. Let $y^1 = h^1(\theta)$ be the measurement of **P**1 on θ, and $y^2 = h^2(\theta)$ be **P**2's measurement, with the property that what **P**2 knows is also known by **P**1 (but not necessarily *vice versa*).[33] Let Γ^i be the set of all measurable policies of the form $u^i = \gamma^i(y^i)$, $i = 1, 2$, and $\hat{\Gamma}^1$ be the set of all measurable policies of the form $u^1 = \gamma^1(y^1, u^2)$. Introduce the pair of policies

$$(\gamma^{1^t}, \gamma^{2^t}) \stackrel{\Delta}{=} \arg \min_{\gamma^1 \in \Gamma^1, \gamma^2 \in \Gamma^2} E_\theta \{L^1(\gamma^1(h^1(\theta)), \gamma^2(h^2(\theta)), \theta)\}$$

assuming that the underlying team problem admits a minimizing solution. Then, a *representation* of the leader's strategy γ^{1^t} under the pair above is an element of $\hat{\Gamma}^1$, say $\hat{\gamma}^1$, with the property

$$\hat{\gamma}^1(h^1(\theta), \gamma^{2^t}(h^2(\theta))) = \gamma^{1^t}(h^1(\theta)), \quad \text{a.s.}^{34} \tag{111}$$

The following result now readily follows:

Proposition 17 *For the stochastic incentive decision problem with nested information as formulated above, the pair* $(\hat{\gamma}^1, \gamma^{2^t})$ *constitutes a global Stackelberg solution, leading to the best possible outcome for the leader. Equivalently, if a strategy* $\hat{\gamma}^1 \in \hat{\Gamma}^1$ *exists satisfying (111), the stochastic decision problem is incentive controllable.*[35] $\diamond$

Remark 15 For the special class of LQG problems, where the decision variables (u^1, u^2) belong to finite-dimensional Euclidean spaces, θ is a Gaussian random vector, h^1 and h^2 are linear, and L^1 is jointly quadratic in the triple (u^1, u^2, θ), strictly convex in (u^1, u^2) for each θ, the team-optimal policies γ^{1^t} and γ^{2^t} exist, are unique and linear in y^1 and y^2, respectively (see any standard book in stochastic control. If, furthermore, L^2 is also a quadratic function, then except for some isolated cases one can restrict the search to linear representations of γ^{1^t}:

$$\hat{\gamma}^1(y^1, u^2) = \gamma^{1^t}(y^1) + P[u^2 - \gamma^{2^t}(y^2)] \tag{112}$$

[33] In mathematical terms, this requirement can be stated as the sigma-field generated by y^1 including the sigma-field generated by y^2.

[34] The equality should hold for almost all values of θ, under its assumed distribution function.

[35] The terminology we have used here is the natural counterpart (in the stochastic case) of the one introduced in Definition 22 for deterministic incentive problems.

where P is a matrix of appropriate dimensions. Now, invoking the condition (111) one can obtain an equation for P, whose solution (when used in (112)) leads to a linear incentive policy. This then makes the decision problem linear incentive controllable; for details see *(Başar, 1979d).* $\diamond$

The development above does not cover (even in the LQG framework) the most general class of dynamic nested stochastic incentive problems, because the measurements of the decision makers have been taken to be *static*—not depending on the past actions. If the leader's measurement at stage k depends on the past actions of the follower $(u_\ell^2, \ell < k)$, then the approach discussed above can easily be adjusted to apply to such multi-stage problems too. If, however, the follower also has access to the leader's past control actions, then because of the nestedness of the information structure for the leader (which does not allow for the follower to have access to all measurements of the leader) the associated dynamic team problem becomes what is called a *nonclassical stochastic control problem*, for which no general theory exists. Issues such as learning, inference, and filtering become of relevance then, whose treatment requires background in stochastic processes, information theory and control, much beyond the level of our coverage here.

13 Lecture 13: Differential games

Continuous-time infinite dynamic games, also known as *differential games* in the literature, constitute a class of decision problems wherein the evolution of the state is described by a differential equation and the players act throughout a time interval. Hence, as a counterpart of Definition 17, we can formulate such games of prespecified fixed duration as follows:

Definition 23 *A* quantitative N-person differential game *of prespecified fixed duration involves the following.*

(i) An index set $\mathcal{N} = \{1, \ldots, N\}$ *called the* players' set.

(ii) A time interval $[0, T]$ *which is specified* a priori *and which denotes the duration of the evolution of the game.*

(iii) An infinite set S_0 *with some topological structure, called the* trajectory space *of the game. Its elements are denoted as* $\{x(t), 0 \leq t \leq T\}$ *and constitute the permissible* state trajectories *of the game. Furthermore, for each fixed* $t \in [0, T]$, $x(t) \in S^0$, *where* S^0 *is a subset of a finite-dimensional vector space, say* $\mathbf{R}^n$.

(iv) An infinite set U^i *with some topological structure, defined for each* $i \in \mathcal{N}$ *and which is called the* control (action) space *of* $\mathbf{P}i$, *whose elements* $\{u^i(t), 0 \leq t \leq T\}$ *are the* control functions *or simply the* controls *of* $\mathbf{P}i$. *Furthermore, there exists a set* $S^i \subseteq \mathbf{R}^{m_i} (i \in \mathcal{N})$ *so that, for each fixed* $t \in [0, T]$, $u^i(t) \in S^i$.

(v) A differential equation

$$\frac{dx(t)}{dt} = f(t, x(t), u^1(t), \ldots, u^N(t)), x(0) = x_0, \tag{113}$$

whose solution describes the state trajectory *of the game corresponding to the* N*-tuple of control functions* $\{u^i(t),\ 0 \leq t \leq T\}$ $(i \in \mathcal{N})$ *and the given initial state* x_0.

(vi) A set-valued function $\eta^i(\cdot)$ *defined for each* $i \in \mathcal{N}$ *as*

$$\eta^i(t) = \{x(s), \quad 0 \leq s \leq \epsilon^i_t\}, \quad 0 \leq \epsilon^i_t \leq t, \tag{114}$$

where ϵ_t^i is nondecreasing in t, and $\eta^i(t)$ determines the state information gained and recalled by $\mathbf{P}i$ at time $t \in [0,T]$. Specification of $\eta^i(\cdot)$ (in fact, ϵ_t^i in this formulation) characterizes the information structure (pattern) *of $\mathbf{P}i$, and the collection (over $i \in \mathcal{N}$) of these information structures is the* information structure *of the game.*

(vii) A sigma-field N_t^i, in S_0, generated for each $i \in \mathcal{N}$ by the cylinder sets $\{x \in S_0, x(s) \in B\}$ where B is a Borel set in S^0 and $0 \leq s \leq \epsilon_t$. N_t^i, $t \geq t_0$, is called the information field *of $\mathbf{P}i$.*

(viii) A prespecified class Γ^i of mappings $\gamma^i : [0,T]\times S_0 \to S^i$, with the property that $u^i(t) = \gamma^i(t,x)$ is N_t^i-measurable (i.e. it is adapted to the information field N_t^i). Γ^i is the strategy space of $\mathbf{P}i$ and each of its elements γ^i is a permissible strategy for $\mathbf{P}i$.

(ix) Two functionals $q^i : S^0 \to \mathbf{R}, g^i : [0,T] \times S^0 \times S^1 \times \cdots \times S^N \to \mathbf{R}$ defined for each $i \in \mathcal{N}$, so that the composite functional

$$L^i(u^1,\ldots,u^N) = \int_0^T g^i(t,x(t),u^1(t),\ldots,u^N(t))\,\mathrm{d}t + q^i(x(T)) \tag{115}$$

is well defined [36] for every $u^j(t) = \gamma^j(t,x), \gamma^j \in \Gamma^j (j \in \mathcal{N})$, and for each $i \in \mathcal{N}$. L^i is the cost functional of $\mathbf{P}i$ in the differential game of fixed duration. $\diamond$

A differential game, as formulated above, is yet not well-defined unless we impose some additional restrictions on some of the terms introduced. In particular, we have to impose conditions on f and $\Gamma^i (i \in \mathcal{N})$, so that the differential equation (113) admits a unique solution for every N-tuple $\{u^i(t) = \gamma^i(t,x),\ i \in \mathcal{N}\}$, with $\gamma^i \in \Gamma^i$. A nonunique solution to (113) is clearly not allowed under the extensive form description of a dynamic game, since it corresponds to nonunique state trajectories (or game paths) and thereby to a possible nonuniqueness in the cost functions for a single N-tuple of strategies. We now provide below in Thm 11 a set of conditions under which this uniqueness requirement is fulfilled. But first we list down some information structures within the context of deterministic differential games, as a counterpart of Definition 18.

Definition 24 *In an N-person continuous-time deterministic dynamic game (differential game) of prespecified fixed duration $[0,T]$, we say that $\mathbf{P}i$'s information structure is*

(i) open-loop *(OL) pattern if $\eta^i(t) = \{x_0\}, t \in [0,T]$,*

[36] This term will be made precise in the sequel.

(ii) closed-loop perfect state *(CLPS) pattern if*

$$\eta^i(t) = \{x(s), 0 \le s \le t\}, \quad t \in [0,T],$$

(iii) ϵ-delayed closed-loop perfect state *(ϵDCLPS) pattern if*

$$\eta^i(t) = \begin{cases} \{x_0\}, & 0 \le t \le \epsilon \\ \{x(s), 0 \le s \le t-\epsilon\}, & \epsilon < t \end{cases}$$

where $\epsilon > 0$ is fixed,

(iv) memoryless perfect state *(MPS) pattern if $\eta^i(t) = \{x_0, x(t)\}$, $t \in [0,T]$,*[37]

(v) feedback (perfect state) *(FB) pattern if $\eta^i(t) = \{x(t)\}$, $t \in [0,T]$.* ⋄

Theorem 11 *Within the framework of Definition 23, let the information structure for each player be any one of the information patterns of Definition 24. Furthermore, let $S_0 = C^n[0,T]$. Then, if*

(i) $f(t,x,u^1,\ldots,t^N)$ is continuous in $t \in [0,T]$ for each $x \in S^0$, $i \in \mathcal{N}$,

(ii) $f(t,x,u^1,\ldots,u^N)$ is uniformly Lipschitz in $x,u^1,\ldots,u^N$; i.e. for some $k > 0$,[38]

$$\begin{aligned} &|f(t,x,u^1,\ldots,u^N) - f(t,\bar{x},\bar{u}^1,\ldots,\bar{u}^N)| \\ &\qquad \le k \max_{0\le t\le T}\{|x(t)-\bar{x}(t)| + \sum_{i\in\mathcal{N}} |u^i(t)-\bar{u}^i(t)|\}, \\ &\qquad x(\cdot),\bar{x}(\cdot) \in \mathbf{C}^n[0,T]; u^i(\cdot),\bar{u}^i(\cdot) \in U^i \quad (i \in \mathcal{N}), \end{aligned}$$

(iii) for $\gamma^i \in \Gamma^i$ $(i \in \mathcal{N})$, $\gamma^i(t,x)$ is continuous in t for each $x(\cdot) \in \mathbf{C}^n[0,T]$ and uniformly Lipschitz in $x(\cdot) \in \mathbf{C}^n[0,T]$,

the differential equation (113) admits a unique solution (i.e. a unique state trajectory for every $\gamma^i \in \Gamma^i$ $(i \in \mathcal{N})$, so that $u^i(t) = \gamma^i(t,x)$, and furthermore this unique trajectory is continuous.

Proof It follows from a standard result on the existence of unique continuous solutions to differential equations. ⋄

The SPE, NE, and SES concepts introduced earlier for finite games are equally valid for (continuous-time) differential games if we bring them into equivalent normal form. To this end, we

[37] Note that (iv) and (v) are not covered by (114) in Definition 23.
[38] $|v|$ denotes here the Euclidean norm for the vector v.

start with the extensive form description of a differential game, as provided in Definition 23 and under the hypotheses of Theorem 11, and for each fixed N-tuple of strategies $\{\gamma^i \in \Gamma^i; i \in \mathcal{N}\}$ we obtain the unique solution of the functional differential equation

$$\frac{dx(t)}{dt} = f(t, x(t), \gamma^1(t,x), \ldots, \gamma^N(t,x)), x(0) = x_0,$$

and determine the corresponding action (control) vectors $u^i(\cdot) = \gamma^i(\cdot, x), i \in \mathcal{N}$. Substitution of these into (115), together with the corresponding unique state trajectory, thus yields an N-tuple of numbers $\{L^i; i \in \mathcal{N}\}$, for each choice of strategies by the players — assuming of course that functions g^i $(i \in \mathcal{N})$ are integrable, so that (115) are well-defined. Therefore, we have mappings $J^i : \Gamma^1 \times \cdots \times \Gamma^N \to \mathbf{R}$ $(i \in \mathcal{N})$ for each fixed initial state vector x_0, which we call the *cost functional* of $\mathbf{P}i$ in a differential game in normal form. These cost functionals, together with the strategy spaces $\{\Gamma^i, i \in \mathcal{N}\}$ of the players, then constitute the equivalent normal form description of the differential game, which is the right framework to introduce noncooperative equilibrium solution concepts, as we have done earlier for other classes of dynamic games.

13.1 Open-loop Nash equilibrium

The results to be described in this subsection are counterparts of those obtained for the discrete-time version; and in order to display this relationship explicitly, we shall present them here in the same order as their counterparts in Lecture 11. We therefore first have the counterpart of Theorem 8 in the continuous time.

Theorem 12 *For an N-person differential game of prescribed fixed duration $[0, T]$, let*

(i) $f(t, \cdot, u^1, \ldots, u^N)$ be continuously differentiable on $\mathbf{R}^n$, $\forall t \in [0, T]$,

(ii) $g^i(t, \cdot, u^1, \ldots, u^N)$ and $q^i(\cdot)$ be continuously differentiable on $\mathbf{R}^n$, $\forall t \in [0, T], i \in \mathbf{N}$.

Then, if $\{\gamma^{i*}(t, x_0) = u^{i*}(t);\ i \in \mathbf{N}\}$ provides an open-loop Nash equilibrium solution*, and $\{x^*(t),\ 0 \le t \le T\}$ is the corresponding state trajectory, there exist N* costate functions *$p^i(\cdot)[0, T] \to \mathbf{R}^n, i \in \mathbf{N}$, such that the following relations are satisfied:*

$$\dot{x}^*(t) = f(t, x^*(t), u^{1*}(t), \ldots, u^{N*}(t)); \quad x^*(0) = x_0 \tag{116}$$

$$\begin{aligned} \gamma^{i*}(t, x_0) \quad \equiv \quad & u^{i*}(t) \arg \min_{u^i \in S^i} H^i(t, p^i(t), x^*(t), u^{1*}(t), \ldots, \\ & \ldots, u^{i-1*}(t), u^i, u^{i+1*}(t), \ldots, u^{N*}(t)) \end{aligned} \tag{117}$$

$$
\begin{aligned}
\dot{p}^{i'}(t) &= -\frac{\partial}{\partial x}H^i(t,p^i(t),x^*,u^{1*}(t),\ldots,u^{N*}(t)) \\
p^{i'}(T) &= \frac{\partial}{\partial x}q^i(x^*(T)), \quad i\in\mathbf{N},
\end{aligned}
\tag{118}
$$

where

$$
H^i(t,p,x,u^1,\ldots,u^N) \triangleq g^i(t,x,u^1,\ldots,u^N)+p^{i'}f(t,x,u^1,\ldots,u^N),
$$

$$
t\in[0,T], \quad i\in\mathbf{N}.
$$

Every such Nash equilibrium solution is weakly time consistent.

Proof Follows the same lines as in the proof of Theorem 8, but now the minimum principle for continuous-time control systems is used [27]. $\diamond$

Remark 16 One class of differential games for which the necessity condition of Theorem 12 is satisfied is that with *weakly coupled* players, that is one with the following state equation and cost functions (taking N=2, without any loss of generality):

$$
\begin{aligned}
\dot{x}_1(t) &= f_1(t,x_1(t),u^1(t))+\epsilon f_{12}(t,x_2(t)); \quad x_1(0)=x_{10} \\
\dot{x}_2(t) &= f_2(t,x_2(t),u^2(t))+\epsilon f_{21}(t,x_1(t)); \quad x_2(0)=x_{20}
\end{aligned}
\tag{119}
$$

and cost functionals

$$
\begin{aligned}
L^i(u^1,u^2;\epsilon) = \int_0^T & [g^{ii}(t,x_i(t),u^i(t))+\epsilon g^{ij}(t,x_j(t),u^j(t))]\,\mathrm{d}t \\
& +q^{ii}(x_i(T))+\epsilon q^{ij}(x_j(T)); \quad j\neq i, \quad i,j=1,2,
\end{aligned}
\tag{120}
$$

where ϵ is a sufficiently small scalar. Under some appropriate convexity (on g^{ij}'s) and differentiability conditions, it can be shown (see *Srikant and Başar, 1992*) [52] that there exists an $\epsilon_0>0$ such that for all $\epsilon\in(-\epsilon_0,\epsilon_0)$, the differential game admits a unique OL NE solution that is stable with respect to Gauss-Seidell or Jacobi iterations (see Definition 14 for terminology). This solution can be obtained by expanding the state and control corresponding to the OL NE solution in power series in terms of ϵ,

$$
x^*(t;\epsilon)=\sum_{k=0}^{\infty}x^{(k)}(t)\epsilon^k, \quad u^{i*}(t;\epsilon)=\sum_{k=0}^{\infty}u_i^{(k)}(t)\epsilon^k,
$$

substituting these into (116)-(118), along with a similar expansion for $p^i(t)$, and solving for the different terms $x^{(k)}$ and $u_i^{(k)}$, $k=0,1,\ldots$, iteratively. It turns out that $u_i^{(0)}$ $(i=1,2)$ are the (open-loop) optimal controls associated with the decoupled optimal control problems:

$$
\begin{aligned}
\dot{x}_i &= f_i(t,x_i,u^i(t)), \quad x_i(0)=x_{i0} \\
L^i(u^i) &= \int_0^T g^{ii}(t,x_i(t),u^i(t))\,\mathrm{d}t+q^{ii}(x_i(T)); \quad i=1,2,
\end{aligned}
$$

and $x^{(0)}$ is the corresponding state trajectory, with $x^{(0)} = (x_1^{(0)'}, x_2^{(0)'})'$. For $k \geq 1$, $u_1^{(k)}$ and $u_2^{(k)}$ are obtained by solving some appropriate linear-quadratic optimal control problems (see, *Srikant and Başar, 1991*) [53]). Hence this approach decomposes the original two-player differential game into two nonlinear optimal control problems (the *zero*th order problems) and a sequence of iteratively constructed linear-quadratic control problems. Halting this iteration at the k'th step yields an ϵ^k-approximate open-loop Nash equilibrium solution. $\diamond$

As indicated earlier, Theorem 12 provides a set of necessary conditions for the OL NE solution to satisfy, and therefore it can be used to generate candidate solutions. For the special class of "affine-quadratic" differential games, however, a unique candidate solution can be obtained in explicit terms, which can further be shown to be an OL NE solution under certain convexity restrictions on the cost functionals of the players:

Definition 25 *An N-person differential game of fixed prescribed duration is of the* affine-quadratic *type if $S^i = \mathbf{R}^{m_i}$ $(i \in \mathbf{N})$ and*

$$\begin{aligned} f(t,x,u^1,\ldots,u^N) &= A(t)x + sum_{i\in\mathbf{N}} B^i(t)u^i + c(t) \\ g^i(t,x,u^1,\ldots,u^N) &= \frac{1}{2}(x'Q^i(t)x + sum_{j\in\mathbf{N}} u^{j'} R^{ij}(t)u^j) \\ q^i(x) &= \frac{1}{2}x'Q^i_f x, \end{aligned}$$

where $A(\cdot)$, $B^i(\cdot)$, $Q^i(\cdot)$, $R^{ij}(\cdot)$ are matrices of appropriate dimensions, $c(\cdot)$ is an n-dimensional vector, all defined on $[0,T]$, and with continuous entries $(i,j \in \mathbf{N})$. Furthermore $Q^i_f, Q^i(\cdot)$ are symmetric, and $R^{ii}(\cdot) > 0$ $(i \in \mathbf{N})$.
An affine-quadratic game is of the linear quadratic *type if $c \equiv 0$.* $\diamond$

Theorem 13 *For an N-person affine-quadratic differential game with $Q^i(\cdot) \geq 0$, $Q^i_f \geq 0$ $(i \in \mathbf{N})$, let there exist a unique solution set $\{M^i; i \in \mathbf{N}\}$ to the coupled matrix Riccati differential equations*

$$\begin{aligned} \dot{M}^i + M^iA + A'M^i + Q^i - M^i \textstyle\sum_{j\in\mathbf{N}} B^j(R^{jj})^{-1}B^{j'}M^j &= 0; \\ M^i(T) = Q^i_f \quad (i \in \mathbf{N}). & \end{aligned} \tag{121}$$

Then, the differential game admits a unique open-loop Nash equilibrium solution *given by*

$$\gamma^{i*}(t,x_0) \equiv u^{i*}(t) = -R^{ii}(t)^{-1}B^{i'}(t)[M^i(t)x^*(t) + m^i(t)] \quad (i \in N) \tag{122}$$

where $\{m^i(\cdot),\ i \in \mathbf{N}\}$ *solve uniquely the set of linear differential equations:*

$$\begin{array}{rcl} \dot{m}^i + A'm^i + M^i c - M^i \sum_{j\in\mathbf{N}} B^j (R^{jj})^{-1} B^{j'} m^j & = & 0; \\ m^i(T) = 0 \quad (i \in \mathbf{N}), & & \end{array} \tag{123}$$

and $x^*(\cdot)$ *denotes the (Nash) equilibrium state trajectory, generated by*

$$\begin{array}{rcl} x^*(t) & = & \Phi(t,0)x_0 + \int_0^t \Phi(t,\sigma)\eta(\sigma)\,\mathrm{d}\sigma \\ \frac{\mathrm{d}}{\mathrm{d}t}\Phi(t,\sigma) & = & F(t)\Phi(t,\sigma); \quad \Phi(\sigma,\sigma) = I \\ F(t) & \stackrel{\Delta}{=} & A - \sum_{i\in\mathbf{N}} B^i R^{ii^{-1}} B^{i'} M^i(t), \\ \eta(t) & \stackrel{\Delta}{=} & c(t) - \sum_{i\in\mathbf{N}} B^i R^{ii^{-1}} B^{i'} m^i(t). \end{array} \tag{124}$$

Proof For the affine-quadratic differential game, and in view of the additional restrictions $Q^i(\cdot) \geq 0$, $Q^i_f \geq 0$, $L^i(u^1,\ldots,u^N)$ is a strictly convex function of $u^i(\cdot)$ for all permissible control functions $u^j(\cdot)$ $(j \neq i, j \in \mathbf{N})$ and for all $x_0 \in \mathbf{R}^n$. This then implies that Theorem 12 is also a sufficiency result and every solution set of the first order conditions provides an open-loop Nash solution. We, therefore, now show that the solution given in Theorem 13 is the only candidate solution.

First note that the Hamiltonian is

$$H^i(t,p,x,u^1,\ldots,u^N) = \frac{1}{2}(x'Q^i x + \sum_{j\in\mathbf{N}} u^{j'} R^{ij} u^j) + p^{i'}(Ax + c + \sum_{j\in\mathbf{N}} B^j u^j)$$

whose minimization with respect to $u^i(t) \in \mathbf{R}^{m_i}$ yields the unique relation

$$u^{i*}(t) = -R^{ii}(t)^{-1} B^i(t)' p^i(t), \quad i \in \mathbf{N}. \tag{i}$$

Furthermore, the costate equations are

$$\dot{p}^i = -Q^i x^* - A' p^i; \quad p^i(T) = Q^i_f x(T) \quad (i \in \mathbf{N}),$$

and the optimal state trajectory is given by

$$\dot{x}^* = Ax^* + c - \sum_{i\in\mathbf{N}} B^i R^{ii^{-1}} B^{i'} p^i; \quad x^*(0) = x_0. \tag{ii}$$

This set of differential equations constitutes a two-point boundary value problem, the solution of which can be written, without any loss of generality, as $\{p^i(t) = M^i(t)x^*(t) + m^i(t),\ i \in \mathbf{N}; x^*(t), t \in [0,T]\}$ where $M^i(\cdot)$ are $(n \times n)$-dimensional matrices and $m^i(\cdot)$ are n-dimensional vectors. Now, substituting $p^i = M^i x^* + m^i$ ($i \in \mathbf{N}$) into the costate equations, we arrive at the conclusion that M^i ($i \in \mathbf{N}$) and m^i ($i \in \mathbf{N}$) should then satisfy (121) and (123), respectively. The expressions for the open-loop Nash strategies follow readily from (i) by substituting $p^i = M^i x^* + m^i$, and likewise the associated state trajectory (124) follows from (ii). $\diamond$

13.2 Closed-loop feedback Nash equilibrium

In order to eliminate informational nonuniqueness in the derivation of Nash equilibria under dynamic information (specifically under the MPS and CLPS patterns) we refine the Nash solution concept further, by requiring it to be strongly time consistent or sub-game perfect. Such a consideration leads to the concept of a "feedback Nash equilibrium solution" which is introduced below.

Definition 26 *For an N-person differential game of prescribed fixed duration $[0, T]$ and with memoryless perfect state (MPS) or closed-loop perfect state (CLPS) information pattern, an N-tuple of strategies $[\gamma^{i*} \in \Gamma^i; i \in \mathbf{N}\}$[39] constitutes a* feedback Nash equilibrium solution *if there exist functionals $V^i(\cdot, \cdot)$ defined on $[0, T] \times \mathbf{R}^n$ and satisfying the following relations for each $i \in \mathbf{N}$:*

$$
\begin{aligned}
V^i(T, x) &= q^i(x) \\
V^i(t, x) &= \int_t^T g^i(s, x^*(s), \gamma^{1*}(s, \eta_s), \ldots, \gamma^{i*}(s, \eta_s), \ldots, \gamma^{N*}(s, \eta_s))\, \mathrm{d}s \\
&\quad + q^i(x^*(T)) \\
&\le \int_t^T g^i(s, x^i(s), \gamma^{1*}(s, \eta_s), \ldots, \gamma^{i-1*}(s, \eta_s), \gamma^i(s, \eta_s), \\
&\quad \gamma^{i+1*}(s, \eta_s), \ldots, \gamma^{N*}(s, \eta_s))\, \mathrm{d}s + q^i(x^i(T)),\ \forall \gamma^i \in \Gamma^i, x \in \mathbf{R}^n
\end{aligned} \tag{125}
$$

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

$$
\begin{aligned}
\dot{x}^i(s) &= f(s, x^i(s), \gamma^{1*}(s, \eta_s), \ldots, \gamma^{i-1*}(s, \eta_s), \gamma^i(s, \eta_s), \\
&\quad \gamma^{i+1*}(s, \eta_s), \ldots, \gamma^{N*}(s, \eta_s));\ x^i(t) = x \\
\dot{x}^*(s) &= f(s, x^*(s), \gamma^{1*}(s, \eta_s), \ldots, \gamma^{i*}(s, \eta_s), \ldots, \gamma^{N*}(s, \eta_s));\ x^*(t) = x
\end{aligned} \tag{126}
$$

and η_s, stands for either the data set $\{x(s), x_0\}$ or $\{x(\sigma), \sigma \le s\}$, depending on whether the information pattern is MPS or CLPS. $\diamond$

One salient feature of the concept of feedback Nash equilibrium (FNE) solution introduced above is that if an N-tuple $\{\gamma^{i*}; i \in \mathbf{N}\}$ provides a FNE to an N-person NZS differential game with

[39] Here Γ^i is chosen to be compatible with the associated information pattern.

duration $[0,T]$, its restriction to the time interval $[t,T]$ provides a FNE to the same differential game defined on the shorter time interval $[t,T]$, with the initial state taken as $x(t)$, and this being so for all $0 \le t \le T$; hence, a FNE is strongly time consistent. An immediate consequence of this observation is that, under either MPS or CLPS information pattern, *feedback Nash equilibrium strategies will depend only on the time variable and the current value of the state*, but not on memory (including the initial state x_0). It is not difficult to see that FNE is indeed a NE.

If the value functions V^i $(i \in \mathbf{N})$ are continuously differentiable in both arguments, then N partial differential equations, related to the Hamilton-Jacobi-Bellman (HJB) equation of optimal control, replace (125):

Theorem 14 . *For an N-person NZS differential game of prescribed fixed duration $[0,T]$, and under either MPS or CLPS information pattern, an N-tuple of strategies $\{\gamma^{i*} \in \Gamma^i; i \in \mathbf{N}\}$ provides a* FNE *if there exist functions $V^i : [0,T] \times \mathbf{R}^n \to \mathbf{R}$, $i \in \mathbf{N}$, satisfying the partial differential equations*

$$\begin{aligned} -\frac{\partial V^i(t,x)}{\partial t} &= \min_{u^i \in S^i} \left[\frac{\partial V^i(t,x)}{\partial x} \tilde{f}^{i*}(t,x,u^i) + \tilde{g}^{i*}(t,x,u^i) \right] \\ &= \frac{\partial V^i(t,x)}{\partial x} \tilde{f}^{i*}(t,x,\gamma^{i*}(t,x)) + \tilde{g}^{i*}(t,x,\gamma^{i*}(t,x)) \\ V^i(T,x) &= q^i(x), \quad i \in \mathbf{N}, \end{aligned} \tag{127}$$

where

$$\begin{aligned} \tilde{f}^{i*}(t,x,u^i) &\triangleq f(t,x,\{\gamma^*_{-i}(t,x),u^i)\}) \\ \tilde{g}^{i*}(t,x,u^i) &\triangleq g^i(t,x,\{\gamma^*_{-i}(t,x),u^i)\}) \\ \{\gamma^*_{-i}(t,x),u^i)\} &\triangleq \gamma^{1*}(t,x),\ldots,\gamma^{i-1*}(t,x),u^i,\gamma^{i+1*}(t,x),\ldots,\gamma^{N*}(t,x). \end{aligned}$$

Every such equilibrium solution is strongly time consistent*, and the corresponding Nash equilibrium cost for $\mathbf{P}i$ is $V^i(0,x_0)$.*

For the class of N-person affine-quadratic differential games (cf. Definition 25), it is possible to obtain an explicit solution for (127), which is quadratic in x. This also readily leads to a set of FNE strategies which are expressible in closed-form.

Corollary 13.1 *For an N-person affine-quadratic differential game with $Q^i(\cdot) \geq 0$, $Q^i_f \geq 0$, $R^{ij}(\cdot) \geq 0$ $(i,j \in \mathbf{N}, i \neq j)$ let there exist a set of matrix valued functions $Z^i(\cdot) \geq 0$, $i \in \mathbf{N}$, satisfying the following N coupled matrix Riccati differential equations:*

$$\begin{array}{c} \dot{Z}^i + Z^i \tilde{F} + \tilde{F}' Z^i + \sum_{j \in \mathbf{N}} Z^j B^j {R^{jj}}^{-1} R^{ij} {R^{jj}}^{-1} B^{j\prime} Z^j + Q^i = 0; \\ Z^i(T) = Q^i_f, \end{array} \tag{128}$$

where

$$\tilde{F}(t) \triangleq A(t) - \sum_{i \in \mathbf{N}} B^i(t) R^{ii}(t)^{-1} B^i(t)' Z^i(t). \tag{129}$$

Then, under either the MPS or CLPS information pattern, the differential game admits a feedback Nash equilibrium solution, affine in the current value of the state, given by

$$\gamma^{i*}(t,x) = -R^{ii}(t)^{-1} B^i(t)' [Z^i(t) x(t) + \zeta^i(t)], \quad (i \in \mathbf{N}), \tag{130}$$

where ζ^i $(i \in \mathbf{N})$ are obtained as the unique solution of the coupled linear differential equations

$$\dot{\zeta}^i + \tilde{F}' \zeta^i + \sum_{i \in \mathbf{N}} Z^j B^j {R^{jj}}^{-1} R^{ij} {R^{jj}}^{-1} B^{j\prime} \zeta^j + Z^i \beta = 0; \; \zeta^i(T) = 0, \tag{131}$$

with

$$\beta \triangleq c - \sum_{j \in \mathbf{N}} B^i {R^{ii}}^{-1} B^{i\prime} \zeta^i. \tag{132}$$

The corresponding values of the cost functionals are

$$J^{i*} = V^i(0, x_0) = \frac{1}{2} x_0' Z^i(0) x_0 + x_0' \zeta^i(0) + n^i(0), \quad (i \in \mathbf{N}), \tag{133}$$

where $n^i(\cdot)$ $(i \in \mathbf{N})$ are obtained from

$$\dot{n}^i + \beta' \zeta^i + \frac{1}{2} \sum_{j \in \mathbf{N}} \zeta^j B^j {R^{jj}}^{-1} R^{ij} {R^{jj}}^{-1} B^{j\prime} \zeta^j = 0; \; n^i(T) = 0, \tag{134}$$

Proof Simply note that, under the condition of solvability of the set of matrix Riccati differential equations (128), the partial differential equation (127) admits a solution in the form $V^i(t,x) = \frac{1}{2} x' Z^i(t) x + x' \zeta^i(t) + n^i(t)$ $(i \in \mathbf{N})$ with the corresponding minimizing controls being given by (130). The "nonnegative definiteness" requirement imposed on $Z^i(\cdot)$ is a consequence of the fact that $V^i(t,x) \geq 0$ $\forall x \in \mathbf{R}^n, t \in [0,T]$, this latter feature being due to the eigenvalue restrictions imposed *a priori* on $Q^i(\cdot)$, Q^i_f and $R^{ij}(\cdot)$. Finally, the corresponding "Nash" values for the cost functionals follow from the definition of $V(t,x)$ (see equation (125)). $\diamond$

Remark 17 The foregoing corollary provides only one set of FNE strategies for the affine-quadratic game under consideration, and it does not attribute any uniqueness feature to this solution set. However, in view of the discrete-time counterpart of this result (cf. Corollary 11.1), one would expect the solution to be unique under the condition that (128) admits a unique solution set; but, in order to verify this, one has to show that it is not possible to come up with other (possibly nonlinear) solutions that satisfy (125), and hitherto this has remained an unresolved problem. What can be shown, though, is uniqueness of FNE when the players are restricted at the outset to affine memoryless strategies. $\diamond$

Remark 18 As in the case of Remark 12, the result above extends readily to more general affine-quadratic dynamic games where the cost functions of the players contain additional terms that are linear in x, that is with g^i and q^i in Definition 25 replaced, respectively, by

$$g^i = \frac{1}{2}\left(x'[Q^i(t)x + 2l^i(t)] + \sum_{j\in\mathbf{N}} u_k^{j'} R_k^{ij} u_k^j \right); \quad q^i = \frac{1}{2} x'(Q_f^i x + 2l_f^i),$$

where $l^i(\cdot)$ is a known n-dimensional vector-valued function, continuous on $[0,T]$, and l_f^i is a fixed n-dimensional vector, for each $i \in \mathbf{N}$. Then, the statement of the Corollary as well its derivation remain intact, with only the differential equation (131) that generates $\zeta^i(\cdot)$ now reading:

$$\dot{\zeta}^i + \tilde{F}'\zeta^i + \sum_{j\in\mathbf{N}} Z^j B^j R^{jj^{-1}} R^{ij} R^{jj^{-1}} B^{j'} \zeta^j + Z^i \beta + l^i = 0; \zeta^i(T) = l_f^i.$$

$\diamond$

Remark 19 For general nonlinear-nonquadratic differential games wherein the players are weakly coupled through the system equation as well as the cost functions, the features observed in Remark 16 for the OL NE can also be derived for the CLFB NE, but now we have to use the sufficiency result of Theorem 14. Again confining ourselves to the two-player case, we take the system equation and the cost functions, as well as the expansion of the state to be in the same structural form as in Remark 16, and only replace the expansion for $u^i(t;\epsilon)$ by a similar expansion for the FB strategy:

$$\gamma^{i*}(t,x;\epsilon) = \sum_{k=0}^{\infty} \gamma_i^{(k)}(t,x)\epsilon^k.$$

Invoking a similar expansion on V^i,

$$V^{i*}(t,x;\epsilon) = \sum_{k=0}^{\infty} V_i^{(k)}(t,x)\epsilon^k,$$

it can be shown (see, *Srikant and Başar, 1991*) [53]) that $V_i^{(0)}$ $(i = 1, 2)$, the *zero*th order terms, satisfy decoupled Hamilton-Jacobi-Bellman equations (associated with the optimal control problems obtained by setting $\epsilon = 0$), and the higher order terms, $V_i^{(k)},\ k \geq 1$, involve simple cost evaluations subject to some state equation constraints. Furthermore, the higher order feedback strategies, $\gamma_i^{(k)},\ k \geq 1$, are obtained from linear equations. More explicitly, for the *zero*th order we have:

$$\begin{aligned} -\frac{\partial}{\partial t} V_i^{(0)}(t, x) &= \min_{u^i \in S^i} \left[\frac{\partial}{\partial x_i} V_i^{(0)}(t, x_1, x_2) f_i(t, x_i, u^i) + g^{ii}(t, x_i, u^i) \right] \\ &= \frac{\partial}{\partial x_i} V_i^{(0)}(t, x_1, x_2) f_i(t, x_i, \gamma_i^{(0)}(t, x)) + g^{ii}(t, x_i, \gamma_i^{(0)}(t, x)) \\ & V_i^{(0)}(T, x) = q^{ii}(x_i), \quad i = 1, 2, \end{aligned} \tag{135}$$

and to *first* order $V_i^{(0)}$ satisfies:

$$\begin{aligned} -\frac{\partial}{\partial t} V_i^{(1)}(t, x) &= \frac{\partial}{\partial x_i} V_i^{(1)} f_i(t, x_i, \gamma_i^{(0)}(t, x)) + \frac{\partial}{\partial x_i} V_i^{(0)} f_{ij}(x_j) \\ &= \frac{\partial}{\partial x_j} V_i^{(1)} f_j(t, x_j, \gamma_j^{(0)}(t, x)) + g^{ij}(t, x_j, \gamma_j^{(0)}(t, x)) \\ V_i^{(1)}(T, x) = q^{ij}(x_j), \quad & j \neq i,\ i, j = 1, 2, \end{aligned} \tag{136}$$

and $\gamma_i^{(1)}$ is obtained from the linear equation

$$\frac{\partial}{\partial u^i} f_i(t, x_i, \gamma_i^{(0)}(t, x)) \left[\frac{\partial}{\partial t} V_i^{(1)}(t, x) \right]' + \frac{\partial}{\partial x_i} V_i^{(0)} f_i(t, x_i, \gamma_i^{(0)}(t, x)) \gamma_i^{(1)}(t, x) \tag{137}$$

$$+ \frac{\partial}{\partial x_i} V_i^{(0)} f_{ij}(x_j) + \frac{\partial^2}{\partial u^i} g^{ii}(t, x_i, \gamma_i^{(0)}(t, x)) \gamma_i^{(1)}(t, x)) = 0,\ i, j = 1, 2. \tag{138}$$

Note that (135) is the Hamilton-Jacobi-Bellman (HJB) equation associated with the optimal control problem obtained by setting $\epsilon = 0$ in the the differential game with weakly coupled players, and hence $\gamma_i^{(0)}$ is the feedback representation of the open-loop policy $u_i^{(0)}$ obtained in Remark 16, on a common *zero*th order trajectory $x_i^{(0)}$. Hence, to *zero*th order, there exists a complete equivalence between open-loop and feedback Nash equilibrium solutions in weakly coupled differential games—as in the case of optimal control. For higher order terms, however, no such correspondence exists in nonzero-sum differential games, which is one explanation for the nonequivalence of the equilibrium trajectories under different information structures—a feature we have observed throughout these *Notes.* $\diamond$

14 Lecture 14: Stochastic differential games: Nash equilibrium

This lecture deals with a class of stochastic differential games, and derivation of NE in that context. It also introduces a different type of cost function not discussed heretofore in these *Notes*, that of *exponentiated integral*; these are known as *risk-sensitive* problems. The lecture will follow the contents of the following paper:

T. Başar, "Nash equilibria of risk-sensitive nonlinear stochastic differential games," J. Optimization Theory and Applications, 100(3):479-498, March 1999.

The paper is reproduced below in its entirety as far as its content goes, starting with its abstract.

Abstract. We consider in this paper a class of risk-sensitive stochastic nonzero-sum differential games with parameterized nonlinear dynamics and parameterized cost functions. The parametrization is such that if all or some of the parameters are set equal to some nominal values, then the differential game either becomes equivalent to a risk-sensitive stochastic control (RSSC) problem, or decouples into several independent RSSC problems, which in turn are equivalent to a class of stochastic zero-sum differential games. This framework allows us to study the sensitivity of the Nash equilibrium (NE) of the original stochastic game to changes in the values of these parameters, and to relate the NE (which is generally difficult to compute, and establish existence and uniqueness for – at least directly) to solutions of RSSC problems, which are relatively easier to obtain. It also allows us to quantify the sensitivity of solutions to RSSC problems (and thereby to nonlinear H_∞-control problems) to unmodelled subsystem dynamics controlled by multiple players.

14.1 Introduction

Most formulations of stochastic control, team, and stochastic differential game problems have been of the *risk-neutral* type, where the cost functions to be minimized are the expected values of stage additive loss functions. Not all behavior, however, can be captured by risk-neutral cost functions — a dictum that has its roots in the statistical decision theory literature of the 1950's. One way of capturing "risk-seeking" or "risk-averse" behavior is to exponentiate a stage additive loss function positively or negatively before expectation, which has first found its way into the control literature through the works of Jacobson [54] and Whittle [55, 56] who have dealt primarily with linear-quadratic models. This last decade of the twentieth century has witnessed a revival of interest in risk-sensitive formulations of control problems in both linear and nonlinear frameworks, with the main driving force being the established relationship between risk-sensitive stochastic control and a class of H_∞ control (worst-case design) problems. For some representative papers on this relationship, which manifests itself in different forms depending on the structure of the problem at hand, and on more recent work that takes this relationship further in different directions, we can cite [57, 58, 59, 60, 61, 62, 63].

Our objective in this paper is to take this line of research a step further, and lift it to the level of nonzero-sum stochastic differential games where each player has a different risk-sensitive cost function with a possibly different risk-sensitivity index, and where the optimality concept is that of Nash equilibrium. We do this by providing a set of sufficient conditions for risk-sensitive Nash equilibrium to satisfy, and by establishing connections with other risk-sensitive stochastic control problems or risk-neutral stochastic differential games, either directly or as appropriate limits. In addition, we develop a computational scheme for Nash equilibria of "weakly-coupled" or "almost-team" stochastic differential games.

Problem formulation is given in the next section, which is followed by the treatment of risk-sensitive stochastic control problems (as a special case) in Section 14.3. Section 14.4 first presents sufficient conditions for the existence of risk-sensitive Nash equilibria, and subsequently specializes these conditions to a few limiting cases, for which there exist recursive computational schemes. The paper ends with the concluding remarks of Section 14.5.

14.2 Problem Formulation

We consider the class of two-player stochastic differential games whose n-dimensional state evolves according to the Itô stochastic differential equation

$$dx(t) \quad = \quad f(t, x(t), u_1(t), u_2(t); \xi)\ dt + \sqrt{\epsilon} D(t, x(t); \xi)\ db(t), \quad x(0) = x_0, \tag{139}$$

where $u_i(t) \in U_i$ is the control input determined by Player i ($i = 1, 2$), with U_i being a subset of the p_i-dimensional Euclidean space, $\mathbf{R}^{p_i}$; $\{b(t), t \geq 0\}$ is a standard m-dimensional Brownian motion process, with $b(0) = 0$ with probability 1; and ξ and ϵ are small positive parameters which will play a role in the analysis to be carried out later in Section 14.4. The control processes $\{u_1(t), t \geq 0\}$, $\{u_2(t), t \geq 0\}$ are adapted to the sigma-field generated by $x(t)$, $t \geq 0$, with the underlying probability space being the triplet $(\Omega, \mathcal{F}, \mathcal{P})$. Let the state-feedback control law (policy) of Player i be denoted by $\mu_i \in \mathcal{M}_i$, so that $u_i(t) = \mu_i(t; x(t))$.

The cost functional to be minimized by Player i is $J_i(\mu_i, \mu_j; 0, x_0; \zeta)$, $j \neq i$, $i, j = 1, 2$, $\zeta := (\epsilon, \xi; \delta_1, \delta_2)$, where

$$J_i(\mu_i, \mu_j; t, x; \zeta) \quad = \quad \delta_i \log E \left\{ \exp \frac{1}{\delta_i} [q_i(x(t_f); \xi) + \int_t^{t_f} g_i(s, x_s, u_{is}, u_{js}; \xi) ds] \right\}, \tag{140}$$

where $\delta_i > 0$ is the *risk-sensitivity index*[40] for Player i, and $x = x(t)$. This cost functional is known as the risk-sensitive cost functional or the exponentiated integral cost.

We assume the following standard regularity conditions on f, q_1, q_2, g_1, g_2 D, μ_1 and μ_2, for a finite horizon $[0, t_f]$:

(i) f, g_1 and g_2 are $\mathcal{C}^1$ in $(t, x, u_1, u_2) \in [0, t_f] \times \mathbf{R}^n \times U_1 \times U_2$, q_1 and q_2 are $\mathcal{C}^2$ in $x \in \mathbf{R}^n$, and both q_i and g_i, $i = 1, 2$, are nonnegative;

(ii) Entries of the matrix D are $\mathcal{C}^2$ in $(t, x) \in [0, t_f] \times \mathbf{R}^n$ for each $\xi \geq 0$, and $DD' > 0$;

(iii) f, f_x, g_i, g_{ix}, $i = 1, 2$, are uniformly bounded on $[0, t_f] \times \mathbf{R}^n \times U_1 \times U_2$, and q_i, q_{ix}, $i = 1, 2$, are uniformly bounded on $[0, t_f] \times \mathbf{R}^n \times$.[41]

(iv) U_1 and U_2 are closed and bounded;

(v) $\mu_i : [0, t_f] \times \mathbf{R}^n \to U_i$ is piecewise continuous in t and Lipschitz continuous in x, $i = 1, 2$.

Now, equipped with the state dynamics (139) and cost functionals (140), we seek a noncooperative Nash equilibrium (μ_1^*, μ_2^*), satisfying the pair of inequalities [1]:

$$J_i(\mu_i^*, \mu_j^*; 0, x_0; \zeta) \quad \leq \quad J_i(\mu_i, \mu_j^*; 0, x_0; \zeta), \;\; \forall \mu_i \in \mathcal{M}_i, j \neq i, i, j = 1, 2; \tag{141}$$

or the stronger version (strongly time-consistent equilibrium):

$$J_i(\mu_i^*, \mu_j^*; t, x; \zeta) \leq J_i(\mu_i, \mu_j^*; t, x; \zeta), \;\; \forall x \in \mathbf{R}^n, t \in [0, t_f), \mu_i \in \mathcal{M}_i, j \neq i, i, j = 1, 2. \tag{142}$$

Three limiting cases are of interest here:

[40] In the literature this terminology is used sometimes for the reciprocal of this quantity.

[41] The notation g_{ix} stands for the partial derivative of the scalar-valued function g_i with respect to its argument x. This convention applies throughout the paper, also with respect to the time variable t.

(C1) **Case when $\epsilon = 0$.** In this case the stochastic differential game degenerates to a deterministic differential game, where the exponentiation of the integral cost loses its significance. If, however, ϵ and δ_i go to zero at the same rate (leading to the so-called *large deviation* limit), then exponentiation still plays an important role, as we will see shortly.

(C2) **Case when $\xi = 0$.** The parameter ξ is a weak-coupling parameter, so that

$$f(t, x, u_1, u_2; 0) = [\tilde{f}_1'(t, x, u_1) \;\; \tilde{f}_2'(t, x_2, u_2)]' \tag{143}$$

$$g_i(t, x, u_i, u_j; 0) = \tilde{g}_i(t, x_i, u_i), \quad i = 1, 2 \tag{144}$$

$$q_i(x; 0) = \tilde{q}(x_i), \quad i = 1, 2 \tag{145}$$

$$D(t, x; 0) = [\widetilde{D}_1'(t, x_1) \; \widetilde{D}_2'(t, x_2)]', \qquad \text{with } \widetilde{D}_1 \widetilde{D}_2' = 0, \forall (t, x) \in [0, t_f] \times \mathbf{R}^n, \tag{146}$$

where $\dim(f_i) = n_i$, $n_1 + n_2 = n$, and $\dim(\widetilde{D}_i) = n_i \times m$. Note that when $\xi = 0$, the original two-player differential game decouples into two separate risk-sensitive stochastic control (RSSC) problems, one for each player, with Player i controlling only an n_i-dimensional state dynamics.

(C3) **Case when $\delta_1 = \delta_2$, $\mathbf{q_1} \equiv \mathbf{q_2}$, and $\mathbf{g_1} \equiv \mathbf{g_2}$.** In this case, the game degenerates to a single RSSC problem, this time of state dimension n.

In Section 14.4, we will study properties of the Nash equilibrium in the neighborhoods of the limiting values above (such as, $\xi = 0$ and $\delta_1 = \delta_2$), which will reveal some structural features, leading to a recursive computation of the Nash equilibrium.

This completes the formulation of the problem under consideration. We now turn, in the next two sections, to detailed analyses of the differential game and its various limiting cases.

14.3 Risk-Sensitive Stochastic Control

Let us first consider the structurally simpler RSSC problem, which will provide useful insight into the solution of the risk-sensitive nonzero-sum stochastic differential game (RS NZSDG). Clearly, in the formulation of Section 14.2, letting $u = (u_1, u_2)$ be a single composite control, $q_1 \equiv q_2 =: q$, $g_1 \equiv g_2 =: g$, and $\delta_1 = \delta_2 =: \delta$ will convert the RS NZSDG into a RSSC problem, which is what we first study here. In this development we will also suppress the dependences of f, q, and g on ξ.

Following [57] and [7], let $\psi(t;x)$ be the value function associated with

$$E\left\{\exp\frac{1}{\delta}[q(x(t_f))+\int_t^{t_f} g(s,x(s),u(s))ds]\right\},$$

that is,

$$\psi(t;x)=\inf_{\mu\in\mathcal{M}} E\left\{\exp\frac{1}{\delta}[q(x(t_f))+\int_t^{t_f} g(s,x(s),u(s))ds]\right\},$$

subject to (139), with $x(t)=x$, and $u(s)=\mu(s;x(s))$, where μ stands for the composite control law (μ_1,μ_2), and $\mathcal{M}$ stands for the composite policy space. Note that

$$V(t;x) \quad := \quad \inf_{\mu\in\mathcal{M}} J(\mu;t,x)=\delta\log\psi(t;x). \tag{147}$$

Assumptions (i)-(v) imply that $\psi(t;x)$ is $\mathcal{C}^1$ in t and $\mathcal{C}^2$ in x, so that it can be differentiated in the Itô sense:

$$d\psi(t;x) \quad = \quad \left[\psi_t(t;x)+\psi_x(t;x)f(t,x,u)+\frac{\epsilon}{2}Tr[\psi_{xx}DD']\right] \quad dt.$$

This, together with the dynamic programming (DP) principle (Ref. 12):

$$\inf_{u\in\mathcal{U}}\left\{d\psi(t;x)+\frac{1}{\delta}\ g(t,x,u)\ \psi(t;x)\ dt\right\} \quad = \quad 0,$$

leads to the DP partial differential equation (PDE)

$$\begin{aligned}\psi_t(t;x_t)+\frac{\epsilon}{2}Tr[\psi_{xx}DD']+\inf_{u\in U}\left\{\psi_x(t;x)f(t,x,u)+\frac{1}{\delta}g(t,x,u)\psi(t;x)\right\}=0,\\ \psi(t_f;x)=\exp\frac{1}{\delta}q(x),\end{aligned} \tag{148}$$

which is a uniformly parabolic PDE (since $DD'>0$), and admits a unique bounded positive solution. Now, in view of (147), we have

$$\psi_t=\frac{1}{\delta}V_t\psi,\quad \psi_x=\frac{1}{\delta}V_x\psi,\quad \psi_{xx}=\frac{1}{\delta}V_{xx}\psi+\frac{1}{\delta^2}V_x'V_x\psi.$$

Substituting these in (148), and dividing throughout by ψ/δ, leads to the Hamilton-Jacobi-Bellman (HJB) equation for the RSSC problem:

$$\begin{aligned}-V_t(t;x) \quad &= \quad \inf_{u\in U}\{V_x(t;x)f(t,x,u)+g(t,x,u)\}+\frac{\epsilon}{2\delta}\left|D'V_x'(t;x)\right|^2+\frac{\epsilon}{2}Tr[V_{xx}DD'],\\ V(t_f;x) \quad &\equiv \quad q(x).\end{aligned} \tag{149}$$

3.1. Affine-Quadratic Case. We now consider the special class of problems where the set U is taken to be $\mathbf{R}^p$,[42] D is independent of x, the vector field f is affine in u, and the incremental cost g is quadratic in u:

$$f(t,x,u) = f_0(t,x) + B(t,x)u, \tag{150}$$

$$g(t,x,u) = g_0(t,x) + |u|^2, \tag{151}$$

This structure allows us to carry out the minimization in (149) explicitly, to lead to the following HJB equation, where we have suppressed the dependence of V on t and x:

$$\begin{aligned} -V_t &= V_x\ f_0(t,x) + g_0(t,x) - \frac{1}{4}\left[\left|B'V_x'\right|^2 - \frac{1}{\gamma^2}\left|D'V_x'\right|^2\right] + \frac{\epsilon}{2}Tr[V_{xx}DD'], \\ V(t_f;x) &\equiv q(x). \end{aligned} \tag{152}$$

Here the parameter γ is defined by

$$\gamma := \sqrt{\delta/2\epsilon},$$

and the optimizing control law is:

$$u^*(t) = \mu^*(t;x) = -\frac{1}{2}B'(t,x)V_x'(t;x), \ \ 0 \le t \le t_f. \tag{153}$$

This derivation has thus led to a sufficiency result, which says that if the HJB equation (152) admits a $\mathcal{C}^2$ solution, then the control law given by (153) provides an optimal solution to the corresponding RSSC problem.

3.2. LEQG Problem. Now, a further special case is the linear-exponential-quadratic-Gaussian (LEQG) problem, where f_0 in (150) is linear in x, and g_0 in (151) and q are nonnegative quadratic in x:

$$f_0(t,x) = A(t)x; \qquad g_0(t,x) = x'Q(t)x,\ Q(t) \ge 0;\ \ q(x) = x'Q_f x,\ Q_f \ge 0, \tag{154}$$

where both $A(\cdot)$ and $Q(\cdot)$ have continuous entries. In this case, (152) admits an explicit solution (whenever it exists), given by

$$V(t;x) = x'Z(t)x + \ell^\epsilon(t), \quad t \ge 0 \tag{155}$$

[42]This is not strictly a special case of the earlier formulation since there U was taken to be compact, but for the affine-quadratic problems considered here the compactness assumption can be relaxed, and the earlier derivation can be carried out without much difference from the one above.

where $Z(\cdot)$ is the nonnegative definite solution of the (generalized) Riccati differential equation (RDE):

$$\dot{Z} + A'Z + ZA + Q - Z\left(BB' - \frac{1}{\gamma^2}DD'\right)Z = 0, \qquad Z(t_f) = Q_f, \tag{156}$$

and ℓ^ϵ is given by

$$\ell^\epsilon(t) \;=\; \epsilon \int_t^{t_f} Tr[Z(s)D(s)D'(s)]ds. \tag{157}$$

The optimal control policy is

$$u^*(t) \;=\; \mu^*(t;x) = -B'(t)Z(t)\,x, \qquad 0 \le t \le t_f. \tag{158}$$

This solution to the LEQG problem was first given in [54], where it was also recognized that the RDE (156) is precisely of the type that arises in two-player linear-quadratic deterministic zero-sum differential games, which is known to have finite escape if γ is sufficiently small. Hence, if γ is sufficiently small (which means that either δ is sufficiently small for fixed $\epsilon > 0$ or ϵ is sufficiently large for fixed $\delta > 0$), then the LEQG problem will not have an optimal solution, and in fact the expected cost will be infinite for all admissible control policies.

3.3. A Class of Stochastic Zero-Sum Differential Games. The relationship alluded to above between the LEQG problem and a class of LQ zero-sum differential games in fact holds at a much more general level, that involves the original nonlinear RSSC problem and a class of stochastic zero-sum differential games.

Consider the two-player state dynamics described by the Itô stochastic differential equation:

$$dx(t) \;=\; f(t,x(t),u(t))\,dt + Dw(t)\,dt + \sqrt{\epsilon}Ddb(t), \;\; x(0) = x_0, \tag{159}$$

where $u(t)$, $t \ge 0$, as defined earlier, is the control input of Player 1, and $w(t)$, $t \ge 0$, is the control input of Player 2, taking values in $\mathbf{R}^m$. Both controls are adapted to the sigma-field generated by $x(t)$, $t \ge 0$, and $b(t)$ is a standard m-dimensional Brownian motion process as before. Consider the risk-neutral cost function:

$$\tilde{J}(\mu,\nu;t,x(t)=x) \;=\; E\left\{q(x(t_f)) + \int_t^{t_f} g(s,x(s),u(s))\,ds - \gamma^2\int_t^{t_f} |w(s)|^2 ds\right\} \tag{160}$$

to be minimized by Player 1 and maximized by Player 2, where $\nu : [0, t_f] \times \mathbf{R}^n \to \mathbf{R}^n$ is the state-feedback control policy of Player 2, which is piecewise continuous in t and Lipschitz continuous in x. Let

$$\overline{W}(t; x) = \inf_{\mu} \sup_{\nu} \tilde{J}(\mu, \nu; t, x)$$

denote the upper-value function associated with this stochastic zero-sum differential game (ZSDG). Then, if it is $\mathcal{C}^1$ in t and $\mathcal{C}^2$ in x, under assumptions (i)-(v) it satisfies the Hamilton-Jacobi-Isaacs (HJI) equation:

$$\inf_{u \in \mathcal{U}} \sup_{w \in \mathbf{R}^n} \left\{ \overline{W}_t + \overline{W}_x (f + Dw) + g - \gamma^2 |w|^2 \right\} + \frac{\epsilon}{2} Tr[\overline{W}_{xx} DD'] = 0, \qquad W(t_f; x) \equiv q(x).$$

Since the dependences on u and w above are separable, Isaacs condition holds [1], and hence the upper-value function is in fact the value function (for the game), satisfying (after carrying out the maximization with respect to w):

$$-W_t(t; x) = \inf_{u \in U} \{ W_x(t; x) f(t, x, u) + g(t, x, u) \}$$

$$+ \frac{1}{4\gamma^2} |D' W_x'(t; x)|^2 + \frac{\epsilon}{2} Tr[W_{xx}(t; x) DD'], \quad W(t_f; x) \equiv q(x). \qquad (161)$$

Note that this PDE is identical to (149), and hence the most general RSSC problem is equivalent to a particular stochastic ZSDG, in the sense that the value function of the two problems are identical, and the optimal control law in the RSSC problem is identical with the saddle-point controller of Player 1 in the stochastic ZSDG. This equivalence also carries over to the deterministic case, which for the ZSDG corresponds to simply setting $\epsilon = 0$ (in (161), as well as in the state equation), while for the RSSC problem it corresponds to letting $\delta \to 0$ as $\epsilon \to 0$, at such a rate so that $\delta / 2\epsilon = \gamma^2$.[43]

14.4 RS-NZSDG: Nash Equilibrium and Its Series Expansion

Now returning to the risk-sensitive nonzero-sum differential game (RS NZSDG) of Section 14.2, we first write down a natural counterpart of (149):

$$-V_{it} = \inf_{u_i \in U_i} \left\{ V_{ix} \bar{f}^i + \bar{g}_i \right\} + \frac{\epsilon}{2\delta_i} |V_{ix} D|^2 + \frac{\epsilon}{2} Tr[V_{ixx} DD'], \ V_i(t_f, x; \xi) \equiv q_i(x; \xi). \qquad (162)$$

[43] Otherwise, simply letting $\epsilon = 0$, will convert the RSSC problem to a purely deterministic one, where exponentiation of the cost will be immaterial, and hence any connection with ZSDGs will be lost.

where

$$
\begin{aligned}
\bar{f}^1 &:= f(t, x, u_1, \mu_2^*(t; x); \xi) && (163)\\
\bar{f}^2 &:= f(t, x, \mu_1^*(t; x), u_2; \xi) && (164)\\
\bar{g}^i &:= g_i(t, x, u_i, \mu_j^*(t; x); \xi), \quad j \neq i,\ i, j = 1, 2, && (165)
\end{aligned}
$$

and (μ_1^*, μ_2^*) is a pair of candidate Nash equilibrium policies. We now immediately have the following theorem.

Theorem 4.1. Consider the RS NZSDG formulated in Section 14.2, with ζ fixed. Assume that there exists a pair $(\mu_1^*, \mu_2^*) \in \mathcal{M}_1 \times \mathcal{M}_2$ such that

(i) the coupled pair of PDE's (162) admits a unique pair of bounded nonnegative solutions, $V^i(t, x; \zeta),\ i = 1, 2$;

(ii) $\mu_i^*(t; x)$ minimizes $V_{ix}\bar{f}^i + \bar{g}_i$ over U_i, $i = 1, 2$.

Then, the pair (μ_1^*, μ_2^*) is a Nash equilibrium solution, and

$$
J_i(\mu_i^*, \mu_j^*; t, x; \zeta) \equiv V_i(t, x; \zeta), \quad i = 1, 2. \tag{166}
$$

Proof: If $u_2(t) = \mu_2^*(t; x(t))$ is substituted into (139) and (140) with $i = 1$, then the HJB equation of the resulting RSSC problem, which is (149), is precisely (162) with $i = 1$. Since this PDE admits a solution by (i), and a control policy achieving the minimum in (162) is μ_1^* by (ii), it follows that μ_1^* is in equilibrium with μ_2^*, and V_1 is the value function for Player 1. A similar argument applied to Player 2 leads to the conclusion that μ_2^* is in equilibrium with μ_1^* and V_2 is the value function for Player 2. $\diamond$

4.1. Relationship With a Risk-Neutral Game. We have seen in Section 14.3 that the RSSC problem is equivalent to a particular worst-case stochastic control problem (or equivalently a stochastic ZSDG) with a risk-neutral cost function. A question of interest is whether a similar type of a game with risk-neutral cost function can be found for the risk-sensitive NZSDG of this section. Indeed such a worst-case NZSDG exists, as elucidated below.

Let us first modify the dynamics (139) to read:

$$dx(t) = f(t_1 x(t), u_1(t), u_2(t); \xi)dt + D(t, x(t); \xi)w(t)dt + \sqrt{\epsilon} D(t, x(t); \xi)db(t), x(0) = x_0, \quad (167)$$

where $w(t)$, $t \geq 0$, is an unknown additional input, which can be viewed as the control of a *third* player, taken to be generated by a policy $\nu : [0, t_f] \times \mathbf{R}^n \to \mathbf{R}^m$ that is continuous in t and Lipschitz continuous in x. The cost functional of Player i, $i = 1, 2$, is now taken as $L_i(\mu_i, \mu_j; 0, x_0; \zeta)$, $j \neq i, j = 1, 2$, where for each fixed $t \in [0, t_f]$,

$$L_i(\mu_i, \mu_j; t, x; \zeta) = \sup_{w=\nu(\cdot)} E\left\{ q_i(x(t_f); \xi) + \int_t^{t_f} \left[g_i(s, x(s), u_i(s), u_j(s); \xi) - \gamma_i^2 |w(s)|^2 \right] ds \right\}, \quad (168)$$

where $x(s), s \geq t$, is generated by (167) with $x(t) = x$, and

$$\gamma_i \quad := \quad \sqrt{\delta_i / 2\epsilon}. \quad (169)$$

Note that L_i is not an exponentiated cost function, but rather one that involves (causal) maximization of a risk-neutral cost function over the unknown input w. It is also to be noted that even though there is a single disturbance input w to the state dynamics (167), the actual maximizing choice in (168) will in general be different for different i; that is, we are not solving for a Nash equilibrium in this three-player game. However, viewed as a two-player game, with individually worst-case cost functions (168) for each of the two players, the Nash equilibrium is then well defined, in exactly the same way as in (141), with of course L_i replacing J_i. Letting (μ_1^*, μ_2^*) be such an equilibrium, and introducing the following counterpart of V_i given by (166):

$$W_i(t, x; \zeta) \quad := \quad L_i(\mu_i^*, \mu_j^*; t, x(t) = x; \zeta), \quad i, j = 1, 2, \; i \neq j, \quad (170)$$

one can readily show, using the arguments of Section 14.3, that provided that W_i's are $\mathcal{C}^2$ they satisfy the coupled pair of PDEs:

$$\begin{aligned} -W_{it} &= \inf_{u_i \in U_i} \left\{ W_{ix} \bar{f}^i + \bar{g}_i \right\} + \frac{\epsilon}{2\delta_i} |W_{ix} D|^2 + \frac{\epsilon}{2} Tr[W_{ixx} D D'], \; W_i(t_f, x; \zeta) \equiv q_i(x; \xi), \quad (171) \\ i &= 1, 2, \end{aligned}$$

where $\bar{f}^i$, $\bar{g}_i$ are defined as before (by (163)-(165)), and the minimum is achieved at $u_i(t) = \mu_i(t; x(t))$, $i = 1, 2$. Note that (171) is identical with (162), and therefore any Nash equilibrium solution of the RS NZSDG constitutes a Nash equilibrium for the stochastic DG formulated above, and *vice versa*. One limiting case of particular interest here is the following: Let $\gamma_i > 0$, $i = 1, 2$, be fixed, and let $\epsilon \to 0$ in (171) (or equivalently in (162)) as well as in (167), which captures the large deviation limit. The resulting coupled PDEs, in this limiting case, are:[44]

$$-V_{it} \quad = \quad \inf_{u_i \in U_i} \left\{ V_{ix}\bar{f}^i + \bar{g}_i \right\} + \frac{1}{4\gamma_i^2}|V_{ix}D|^2,\ V_i(t_f, x; \zeta) \equiv q_i(x; \xi),\ i = 1, 2. \tag{172}$$

Each of these PDEs corresponds to a HJI equation that arises in deterministic nonlinear H_∞ control [7], meaning that the deterministic system

$$\dot{x} \quad = \quad f(t, x, \mu_1^*(t; x), \mu_2^*(t; x); \xi) + Dw;\ \ x_0 = 0, \tag{173}$$

satisfies under an output determined by the i-th criterion a disturbance attenuation level γ_i, i.e.,

$$\int_0^{t_f} g_i(t, x(t), \mu_i^*(t; x(t)), \mu_j^*(t; x(t)); \xi)dt \quad \leq \quad \gamma_i^2 \int_0^{t_f} |w(t)|^2 dt,\ i = 1, 2. \tag{174}$$

Since this argument holds for both $i = 1$ and 2, the Nash equilibrium provides guaranteed levels of disturbance attenuation with respect to two different outputs from the same disturbance input w. As indicated earlier in Section 14.3, even for the case when there is a single player, (172) may not admit a bounded solution for all $\gamma > 0$, and hence in (174) one is forced to restrict the pair (γ_1, γ_2) to a region $\Gamma \subset \mathbf{R}^2$, for the PDEs (172) to admit solutions (in the generalized, viscosity, sense). We will call Γ the *set of achievable performance levels for the multicriteria disturbance attenuation problem* with noncooperative controllers. This is related to (though is not identical with) *joint feasibility of risk sensitivity indices* in the formulation of the RS NZSDG. For the most general NZSDG, the restrictions on these indices (for the existence of a Nash equilibrium) will be coupled and cannot be determined independently of each other, and therefore Γ will not be a rectangular set. However, for weakly coupled games, that is when $\xi > 0$ is sufficiently small, a more transparent characterization can be obtained — as the decoupling analysis presented next will exhibit.

4.2. Sequential Decoupling of the Nash Equilibrium Solution. We had mentioned earlier that setting $\xi = 0$ in the system dynamics and the cost functionals decouples the original RS NSZDG into two separate RSSC problems, and hence it would be of interest to see to what extent the solutions to these RSSC problems provide a Nash equilibrium solution to the original

[44]This is only a formal limit, as (172) may not admit a smooth solution even though (171) does for every $\epsilon > 0$. In this case the solution will be interpreted in the *viscosity* sense [64].

game, and what would the correction terms be like, when ξ is sufficiently small. Toward this end, and to capture the *weakly-coupled* nature of the problem, we make the structural assumption that f, q_i's, and g_i's are given by, for $i = 1, 2$,

$$f = (f_1' \;\; f_2')', \qquad \dim(f_i) = n_i \tag{175}$$

$$f_i \;=\; a_i(x_i) + B_i(x_i)u_i + \xi(a_{ii}(x_i) + a_{ij}(x_j)), \;\; j \neq i, \; j = 1, 2, \tag{176}$$

$$g_i \;=\; c_i(x_i) + |u_i|^2 + \xi(c_{ii}(x_i) + c_{ij}(x_j)), \;\; j \neq i, \; j = 1, 2, \tag{177}$$

$$q_i \;=\; e_i(x_i) + \xi(e_{ii}(x_i) + e_{ij}(x_j)), \;\; j \neq i, \; j = 1, 2, \tag{178}$$

where x_1 is the n_1-dimensional first sub-vector of x, x_2 is the n_2-dimensional second sub-vector of x, and we have suppressed the dependence on the time variable t in (175) and (177). We further take D to be independent of ξ,[45] and

$$D \;=\; [D_1' \;\; D_2']', \text{ with } D_1 D_2' = 0. \tag{179}$$

Note that all this partitioning is consistent with (143)-(146).

Now, using this structure in (162) leads to the simplification (with $j \neq i, \; i, j = 1, 2$)

$$\begin{aligned} -V_{it} \;=\; & \inf_{u_i \in U_i} \{V_{ix_i}[a_i(x_i) + B_i(x_i)u_i + \xi(a_{ii}(x_i) + a_{ij}(x_j))] \\ & +c_i(x_i) + |u_i|^2 + \xi(c_{ii}(x_i) + c_{ij}(x_j))\Big\} \\ & +V_{ix_j}[a_j(x_j) + B_j(x_j)\mu_j^*(t; x) + \xi(a_{jj}(x_j) + a_{ji}(x_i))] \\ & +\frac{\epsilon}{2\delta_i}|V_{ix_i}D_i|^2 + \frac{\epsilon}{2\delta_i}|V_{ix_j}D_j|^2 + \frac{\epsilon}{2}Tr[V_{ix_ix_i}D_iD_i'] \\ & +\frac{\epsilon}{2}Tr[V_{ix_jx_j}D_jD_j'], \quad V_i(t_f, x; \xi) \equiv e_i(x_i) + \xi(e_{ii}(x_i) + e_{ij}(x_j)), \end{aligned} \tag{180}$$

[45] This assumption is not essential for the validity of the results to follow, but is made here for the sake of simplicity, without much loss of conceptual generality.

and the control that minimizes the right-hand side of (180) is

$$\mu_i^*(t,x;\xi) \quad = \quad -\frac{1}{2}B_i'(x_i)V_{ix_i}'(t,x;\xi), \quad i=1,2. \tag{181}$$

Substituting this back into (180) finally leads to (suppressing all the arguments), for $i=1,2$,

$$-V_{it} = V_{ix_i}a_i - \frac{1}{4}|B_i'V_{ix_i}'|^2 + c_i + \frac{\epsilon}{2\delta_i}|V_{ix_i}D_i|^2 + \frac{\epsilon}{2}Tr[V_{ix_ix_i}D_iD_i']$$

$$+\xi[V_{ix_i}(a_{ii}+a_{ij}) + c_{ii} + c_{ij} + V_{ix_j}(a_{jj}+a_{ji})] - \frac{1}{2}V_{ix_j}B_jB_j'V_{jx_j}' + V_{ix_j}a_j$$

$$+\frac{\epsilon}{2\delta_i}|V_{ix_j}D_j|^2 + \frac{\epsilon}{2}Tr[V_{ix_jx_j}D_jD_j'], \;\; V_i(t,x;\xi) \equiv e_i(x_i) + \xi(e_{ii}(x_i)+e_{ij}(x_j)). \tag{182}$$

We now assume that the solution of (182) admits a power series expansion around $\xi = 0$, which we write as

$$V_i(t,x;\xi) \quad = \quad V_i^{(0)}(t,x_i) + \xi V_i^{(1)}(t,x) + \xi^2 V_i^{(2)}(t,x) + o(\xi^2). \tag{183}$$

Using this in (181) leads to a natural expansion of μ_i^* in terms of ξ:

$$\mu_i^*(t,x;\xi) \quad = \quad \mu_i^{(0)}(t,x_i) + \xi\mu_i^{(1)}(t,x) + \xi^2\mu_i^{(2)}(t,x) + o(\xi^2), \tag{184}$$

where

$$\mu_i^{(k)}(t,x) \quad = \quad -\frac{1}{2}B_i'(x_i)V_{ix_2}^{(k)}(t,x)', \quad k=0,1,2. \tag{185}$$

Now, substituting (183) into (182), and equating like powers of ξ, leads to the following equations for iterative computation of $V_i^{(k)}$'s:

Zeroth-order terms

$$-V_{it}^{(0)} \quad = \quad V_{ix_i}^{(0)}a_i - \frac{1}{4}|V_{ix_i}^{(0)}B_i|^2 + \frac{\epsilon}{2\delta_i}|V_{ix_i}^{(0)}D_i|^2 + \frac{\epsilon}{2}\mathrm{Tr}[V_{ix_ix_i}^{(0)}D_iD_i'] + c_i(x_i),$$

$$V_i^{(0)}(t_f,x) \quad \equiv \quad e_i(x_i). \tag{186}$$

First-order terms

$$-V_{it}^{(1)} \quad = \quad V_{ix_i}^{(1)}[a_i - \frac{1}{2}B_iB_i'V_{ix_i}^{(0)\prime} + \frac{\epsilon}{\delta_i}D_iD_i'V_{ix_i}^{(0)\prime}] + V_{ix_i}^{(0)}[a_{ii}(x_i)+a_{ij}(x_j)] + c_{ii}(x_i)$$

$$+c_{ij}(x_j) + V_{ix_j}^{(1)}[a_j - B_jB_j'V_{ix_j}^{(0)'}] + \frac{\epsilon}{2}\mathrm{Tr}[V_{ix_ix_i}^{(1)}D_iD_i' + V_{ix_jx_j}^{(1)}D_jD_j'],$$

$$V_i^{(1)}(t_f, x) \quad \equiv \quad e_{ii}(x_i) + e_{ij}(x_j). \tag{187}$$

Second-order terms

$$\begin{aligned} -V_{it}^{(2)} \quad = \quad & V_{ix_i}^{(2)}[a_i - \frac{1}{2}B_iB_i'V_{ix_i}^{(0)'} + \frac{\epsilon}{\delta_i}D_iD_i'V_{ix_i}^{(0)'}] + V_{ix_i}^{(1)}[a_{ii}(x_i) + a_{ij}(x_j) + \frac{\epsilon}{2\delta_i}D_iD_i'V_{ix_i}^{(1)'}] \\ & + V_{ix_j}^{(1)}[a_{jj}(x_j) + a_{ji}(x_i) + \frac{\epsilon}{2\delta_i}D_jD_j'V_{ix_j}^{(1)'}] + V_{ix_j}^{(2)}[a_j - \frac{1}{2}B_jB_j'V_{jx_j}^{(0)'}] \\ & - \frac{1}{2}V_{ix_j}^{(1)}B_jB_j'V_{ix_j}^{(1)'} - \frac{1}{4}|V_{ix_i}^{(1)}B_i|^2 + \frac{\epsilon}{2}\mathrm{Tr}[V_{ix_ix_i}^{(2)}D_iD_i' + V_{ix_jx_j}^{(2)}D_jD_j'], \end{aligned} \tag{188}$$

$$V_i^{(2)}(t_f, x) \quad \equiv \quad 0.$$

Note that as to be expected the leading terms above (the *zeroth order* terms) provide solutions to two independent RSSC problems, one for each player. It is important to notice here that $V_i^{(0)}$, and thereby $\mu_i^{(0)}$ in (185), have only x_i as their arguments, and not x_j. For the higher-order correction terms, however, the story is quite different. First, the PDEs satisfied by $V_i^{(1)}$ and $V_i^{(2)}$ are both linear, and hence no restrictions are imposed on the ranges of feasible values for δ_i's beyond those imposed at the *zeroth* stage for existence of unique solutions to the zeroth order PDEs. Second, both $V_i^{(1)}$ and $V_i^{(2)}$ depend, in general, not only on x_i but also on x_j, thus making the correction terms in the control laws of each player depend on the entire state vector. This structure as well as that of the feasibility set Γ becomes more transparent in the linear-quadratic case, where for $j \neq i$, $i, j = 1, 2$,

$$\begin{aligned} a_i(t, x_i) = A_i(t)x_i; \quad a_{ii}(t, x_i) = A_{ii}(t)x_i; \quad a_{ij}(t, x_j) = A_{ij}(t)x_j; \quad B_i(t, x) = B_i(t); \\ D_i(t, x) = D_i(t); \quad c_i(t, x_i) = x_i'Q_i(t)x_i; \quad c_{ii}(t, x_i) = x_i'Q_{ii}(t)x_i; \quad c_{ij}(t, x_j) = x_j'Q_{ij}(t)x_j; \\ e_i(x_i) = x_i'Q_{if}x_i; \quad e_{ii}(x_i) = x_i'Q_{iif}x_i; \quad e_{ij}(x_j) = x_j'Q_{ijf}x_j, \\ Q_i(\cdot) > 0, \quad Q_{ii} \geq 0, \quad Q_{ij}(\cdot) \geq 0, \quad Q_{if} \geq 0, \quad Q_{iif} \geq 0, \quad Q_{ijf} \geq 0. \end{aligned}$$

Under this structure, the solutions to (186), (187) and (188) are, respectively,

$$V_i^{(0)}(t, x_i) = x_i'Z_{ii}^{(0)}(t)x_i + \ell_i^{(0)}(t), \;\; V_i^{(1)}(t, x) = x'Z_{ii}^{(1)}x + \ell_i^{(1)}(t), \;\; V_i^{(2)}(t, x) = x'Z_i^{(2)}(t)x + \ell_i^{(2)}(t),$$

where $Z_{ii}^{(0)} \geq 0$ is of dimensions $n_i \times n_i$, whereas $Z_i^{(1)} \geq 0$ are of dimensions $n \times n$. The first-one solves the generalized (game) Riccati differential equation:

$$-\dot{Z}_{ii}^{(0)} = A_i' Z_{ii}^{(0)} + Z_{ii}^{(0)} A_i - Z_{ii}^{(0)} \left(B_i B_i' - \frac{2\epsilon}{\delta_i} D_i D_i' \right) Z_i^{(0)} + Q_i, \quad Z_{ii}(0) = Q_{if} \tag{189}$$

which is of the type (156), and admits a unique nonnegative definite solution provided that $\delta_i := \sqrt{\delta_i / 2\epsilon}$ is above a certain threshold. $Z_i^{(1)}$ and $Z_i^{(2)}$, on the other hand, satisfy some linear matrix differential equations which always admit unique solutions provided that $Z_{ii}^{(0)}$ exist. Hence the set Γ is in this case rectangular, and is determined solely by the existence condition of the zeroth-order Riccati differential equation (189).

4.3. Perturbation Around a Common Risk-Sensitivity Index. We now focus on the role played by the risk-sensitivity indices δ_1 and δ_2, and in particular the effect of a deviation (for one player) from a common risk-sensitivity index for a nominal team problem. Toward this end, and to obtain some explicit expressions for the controllers, we take

$$f = a(x) + B_1(x) u_1 + B_2(x) u_2 \tag{190}$$

$$g_i = c(x) + |u_1|^2 + |u_2|^2, \; q_i = q(x) \tag{191}$$

$$\delta_1 = \delta, \qquad \delta_2 = \delta + \Delta\delta. \tag{192}$$

Under this structure, (162) reads:

$$\begin{aligned} -V_{it} &= V_{ix} a - \frac{1}{4} |B_i' V_{ix}'|^2 + \frac{\epsilon}{2\delta_i} |V_{ix} D|^2 + \frac{\epsilon}{2} Tr[V_{ix} D D'] \\ &\quad + c(x) + \frac{1}{4} |B_j' V_{jx}'|^2 - \frac{1}{2} V_{ix} B_j B_j' V_{jx}', \; V_i(t_f, x; \delta_1, \delta_2) \equiv q(x), \end{aligned} \tag{193}$$

and the Nash equilibrium controllers are, for $i = 1, 2$,

$$\mu_i^*(t, x; \delta_1, \delta_2) = -\frac{1}{2} B_i'(x) V_{ix}'(t, x; \delta_1, \delta_2). \tag{194}$$

Again let us assume a power series expansion around $\delta_1 = \delta_2 = \delta$, with respect to δ_2, of $V_i(t, x; \delta_1, \delta_2)$:

$$V_i(t, x; \delta_1, \delta_2) = V_i^{(0)}(t, x; \delta) + V_i^{(1)}(t, x; \delta)\Delta\delta + V_i^{(2)}(t, x; \delta)(\Delta\delta)^2 + o((\Delta\delta)^2) \tag{195}$$

which results, in view of (194), in a similar expansion for μ^{i*}'s:

$$\mu^{i*}(t,x;\delta_1,\delta_2) \quad = \quad \mu_i^{(0)}(t,x;\delta) + \mu_i^{(1)}(t,x;\delta)\Delta\delta + \mu_i^{(2)}(t,x;\delta)(\Delta\delta)^2 + o((\Delta\delta)^2) \tag{196}$$

where

$$\mu_i^{(k)}(t,x;\delta) \quad = \quad -\frac{1}{2}B_i'(x)V_{ix}^{(k)}(t,x;\delta)', \quad i=1,2. \tag{197}$$

Substitution of (195) into (193), and equating like powers of $\Delta\delta$, leads to:

Zeroth-order terms

$$V_1^{(0)}(t,x;\delta) \quad \equiv \quad V_2^{(0)}(t,x;\delta) =: V^{(0)}(t,x;\delta), \tag{198}$$

where $V^{(0)}$ satisfies the HJB equation for the team problem:

$$\begin{aligned} -V_t^{(0)} \quad &= \quad V_x^{(0)}a - \frac{1}{4}|B_1'V_x^{(0)'}|^2 - \frac{1}{4}|B_2'V_x^{(0)}|^2 + c(x) + \frac{\epsilon}{2\delta}|D'V_x^{(0)}|^2 \\ &\quad + \frac{\epsilon}{2}Tr[V_{xx}^{(0)}DD'], \\ V^{(0)}(t_f,x;\delta) \quad &\equiv \quad q(x). \end{aligned} \tag{199}$$

First-order terms

$$\begin{aligned} -V_{2t}^{(1)} \quad &= \quad V_{2x}^{(1)}\left[a - \frac{1}{2}\left(B_1B_1' + B_2B_2' - \frac{2\epsilon}{\delta}DD'\right)V_x^{(0)'}\right] \\ &\quad + \frac{\epsilon}{2}Tr[V_{2xx}^{(1)}DD'] - \frac{\epsilon}{2\delta^2}|V_x^{(0)}D|^2, \quad V_2^{(1)}(t_f,x;\delta) \equiv 0 \end{aligned} \tag{200}$$

$$V_1^{(1)}(t,x;\delta) \quad \equiv \quad 0. \tag{201}$$

Second-order terms

$$\begin{aligned} -V_{1t}^{(2)} \quad &= \quad V_{1x}^{(2)}\left[a - \frac{1}{2}\left(B_1B_1' + B_2B_2' - \frac{2\epsilon}{\delta}DD'\right)V_x^{(0)'}\right] \\ &\quad + \frac{1}{4}V_{2x}^{(1)}B_2B_2'V_{2x}^{(1)'} + \frac{\epsilon}{2}Tr[V_{1xx}^{(2)}DD'], \quad V_1^{(2)}(t_f,x;\delta) \equiv 0 \end{aligned} \tag{202}$$

$$\begin{aligned} -V_{2t}^{(2)} \quad &= \quad V_{2x}^{(2)}\left[a - \frac{1}{2}\left(B_1B_1' + B_2B_2' - \frac{2\epsilon}{\delta}DD'\right)V_x^{(0)'}\right] \\ &\quad - \frac{\epsilon}{\delta^2}V_x^{(0)}DD'V_{2x}^{(1)'} - \frac{1}{4}V_{2x}^{(1)}B_2B_2'V_{2x}^{(1)'} + \frac{\epsilon}{2\delta}V_{2x}^{(1)}DD'V_{2x}^{(1)'} \\ &\quad + \frac{\epsilon}{2}Tr[V_{2xx}^{(2)}DD'] + \frac{\epsilon}{2\delta^3} \mid V_x^{(0)}D \mid^2, \quad V_2^{(2)}(t_f,x;\delta) \equiv 0. \end{aligned} \tag{203}$$

Note that again as to be expected the leading (*zeroth-order*) terms define a common RSSC problem. There is no first-order correction to μ_1 or V_1, since both $V_1^{(1)}$ and $\mu_1^{(1)}$ are zero, but there are first-order corrections to μ_2 and V_2, with $V_2^{(1)}$ satisfying a linear differential equation. The fact that $V_1^{(1)} \equiv 0$ is to be expected as the change in the value of δ_2 has a more immediate effect on Player 2's optimum cost, than on Player 1's. Regarding second-order correction terms, however, both $V_1^{(2)}$ and $V_2^{(2)}$ are nonzero, and they are both obtained from solutions of linear (though coupled) differential equations. Hence, no additional restrictions are imposed on δ through these higher-order terms, and an iterative computation of all correction terms is again possible.

These features are better seen if we consider the special class of linear-quadratic RS NZSDGs, where

$$a(x) = Ax, \quad B_1(x) = B_1, \quad B_2(x) = B_2 \tag{204}$$

$$c(x) = x'Qx, \quad Q \geq 0; \tag{205}$$

$$q(x) = x'Q_f x, \quad Q_f \geq 0; \tag{206}$$

with all matrices, A, B_1, B_2, Q, being independent of x. Then, the *zeroth*-order Nash equilibrium solution is

$$\mu_i^{(0)}(t;x) = -B_i' Z^{(0)}(t)x, \quad i = 1, 2, \tag{207}$$

where $Z^{(0)}$ satisfies the RDE:

$$-\dot{Z}^{(0)} = Z^{(0)}A + A'Z^{(0)} - Z^{(0)}\left[B_1B_1' + B_2B_2' - \frac{2\epsilon}{\delta}DD'\right]Z^{(0)} + Q,\ Z^{(0)}(t_f) = Q_f, \tag{208}$$

which admits a unique nonnegative definite solution if either δ is sufficiently large or ϵ is sufficiently small.

For $i = 1$, the first-order correction to (207) is zero, whereas for $i = 2$ the first-order correction to the same is:

$$\mu_2^{(1)}(t,x) = -B_2' Z_2^{(1)}(t)x, \tag{209}$$

where $Z_2^{(1)}$ satisfies the Lyapunov differential equation:

$$-\dot{Z}_2^{(1)} = Z_2^{(1)}F^{(0)} + F^{(0)\prime}Z_2^{(1)} - \frac{2\epsilon}{\delta}Z^{(0)}DD'Z^{(0)},\ Z_2^{(1)}(t_f) = 0 \tag{210}$$

where

$$F^{(0)} \quad := \quad A - \left(B_1 B_1' + B_2 B_2' - \frac{2\epsilon}{\delta} DD' \right) Z^{(0)}. \tag{211}$$

Finally, the second-order corrections to (207) are

$$\mu_i^{(2)}(t, x) \quad = \quad -B_i' Z_i^{(2)}(t) x, \tag{212}$$

where $Z_1^{(2)}$ and $Z_2^{(2)}$ satisfy

$$-\dot{Z}_1^{(2)} \quad = \quad Z_1^{(2)} F^{(0)} + F^{(0)\prime} Z_1^{(2)} + Z_2^{(1)} B_2 B_2' Z_2^{(1)}, \;\; Z_1^{(2)}(t_f) = 0 \tag{213}$$

$$\begin{aligned} -\dot{Z}_2^{(2)} \quad = \quad & Z_2^{(2)} F^{(0)} + F^{(0)\prime} Z_2^{(2)} - \frac{2\epsilon}{\delta^2} [Z^{(0)} DD' Z_1^{(1)} + Z_2^{(1)} DD' Z^{(0)}] \\ & - Z_2^{(1)} \left[B_2 B_2' - \frac{2\epsilon}{\delta} DD' \right] Z_2^{(1)} + \frac{2\epsilon}{\delta^3} Z^{(0)} DD' Z^{(0)}, \;\; Z_2^{(2)}(t_f) = 0, \end{aligned} \tag{214}$$

which are linearly coupled Lyapunov differential equations, and hence admit unique solutions. Therefore, as long as the Riccati differential equation at the zeroth stage admits a nonnegative-definite solution (which is guaranteed if the quantity $\gamma = \sqrt{\delta/2\epsilon}$ exceeds a precomputable threshold), the expansion exists and the controllers for both players can be constructed recursively to any degree of accuracy.

14.5 Conclusions

In this paper, we have obtained a set of sufficient conditions for Nash equilibria of two-player nonzero-sum nonlinear stochastic differential games with risk-sensitive performance indices, and have established connections with stochastic games of the risk-neutral type. We have also studied some limiting cases where either the risk-sensitive nonzero-sum differential game degenerates to a single RSSC problem or decouples into two independent risk-sensitive stochastic control problems. Furthermore, we have investigated small noise limits, where the game solution becomes asymptotically equivalent to that of a multi-criteria nonlinear H_∞ control problem. In addition to providing insight into the structure of the equilibrium solutions, these limiting cases also lead to sequential computational schemes for the construction of Nash equilibria of the risk-sensitive nonzero-sum differential game. The analysis further provides a quantification of the sensitivity of the optimal solutions of risk-sensitive stochastic team problems with respect to unmodeled weak-coupling in the system dynamics or unmodeled discrepancy in the risk-sensitivity indices among different players.

Even though the paper has dealt with two-player nonzero-sum differential games only, both the approach and the results find immediate extensions to the many-player case; this has not been carried out here simply not to burden the development with additional notational complexity. For a similar reason, we have also not included higher than second-order terms in the expansion of the Nash equilibrium solution in terms of the small parameters, which can readily be carried out, with also the system and the cost functions admitting such higher-order expansions. There are, however, other extensions, which require substantial additional work, and stand out as promising future research directions. Among these are:

(i) Extending the results of this paper to the infinite-horizon case;

(ii) Allowing in the problem formulation not only weak spatial coupling, but also fast-slow dynamics (as in [65, 66]);

(iii) Obtaining the counterparts of these results under other information structures, such as sampled-data or imperfect state.

15 Lecture 15: Stochastic differential games: Feedback Stackelberg equilibrium

I have so far discussed in the context of dynamic and differential games under mainly two types of information structures (OL and CL), the solution concept of NE (and as a special case SPE), but not the Stackelberg equilibrium concept (except in the context of incentive design), and in this lecture I discuss the latter for not only deterministic but also stochastic dynamic/differential games. I will focus on the solution concept of *feedback Stackelberg equilibrium*, which is different from the Stackelberg equilibrium under CL FB information pattern, where the latter is not even weakly time consistent, while the former is strongly time consistent (and this makes it appealing). It is one where leadership is exercised stagewise (in a discrete-time format) or at each point in time (in the continuous-time case), which makes its computation more manageable. I will also discuss a game model where leadership could be time-varying, governed by the outcome of an exogenous stochastic process.

The lecture will essentially be based on the contents of the book chapter:

[15-1] T. Başar and A. Haurie, "Feedback equilibria in differential games with structural and modal uncertainties," in J. B. Cruz, Jr., editor, Advances in Large Scale Systems, chapter 1, pp. 163–201. JAE Press Inc., Connecticut, May 1984.

A scanned reprint of this article has been attached.

The lecture will also cover some recent work on mixed leadership, when a player can act both as a leader and a follower (by controlling different instrument variables); this will be based on the journal paper:

[15-2] T. Başar, A. Bensoussan, and S.P. Sethi, "Differential games with mixed leadership: the open-loop solution," J. Applied Mathematics and Computation, 2010 (to appear).

A pre-print of this second paper has also been attached.

16 Lecture 16: CDMA power control as an application of NZSG theory

This lecture will cover the application of non-cooperative nonzero-sum game theory discussed in the previous lectures in CDMA (code division multiple access) power control where the users control their uplink power level in a non-cooperative way. Some pricing schemes will also be discussed in this context. The presentation will be based on the contents of three journal papers:

[16-1] T. Alpcan, T. Başar, R. Srikant, and E. Altman, "CDMA uplink power control as a noncooperative game," Wireless Networks, 8:659-670, 2002.

[16-2] T. Alpcan, T. Başar, and S. Dey, "A power control game based on outage probabilities for multicell wireless data networks," IEEE Transactions on Wireless Communications , 5(4):890-899, April 2006.

[16-3] X. Fan, T. Alpcan, M. Arcak, T. J. Wen, and T. Başar, "A passivity approach to game-theoretic CDMA power control," Automatica , 42(11):1837-1847, November 2006.

Reprints of all three papers are attached.

17 Lecture 17: Routing and flow control

This lecture will cover the application of non-cooperative nonzero-sum game theory discussed in the previous lectures in routing and flow control in communication networks. The presentation will be based on the contents of three journal papers:

[**17-1**] E. Altman and T. Başar, "Multi-user rate-based flow control," IEEE Transactions on Communications, 46(7):940-949, July 1998.

[**17-2**] E. Altman, T. Başar, and R. Srikant, "Nash equilibria for combined flow control and routing in networks: Asymptotic behavior for a large number of users," IEEE Trans. Automatic Control, 47(6):917-930, June 2002.

Reprints of both papers are attached.

18 Lecture 18: Pricing of services as an application of hierarchical game theory

This lecture will cover the application of hierarchical game theory and incentive design schemes (including pricing) to the interaction between service providers and users in a networking environment. It will also introduce in this context large population games. The presentation will be based on the contents of three papers:

[18-1] T. Başar and R. Srikant, "Revenue-maximizing pricing and capacity expansion in a many-users regime," Proc. IEEE Infocom 2002,Vol. 1, pp. 294-301, New York, New York, June 23-27, 2002

[18-2] H. Shen and T. Başar, "Optimal nonlinear pricing for a monopolistic network service provider with complete and incomplete information," IEEE J. on Selected Areas in Communications (JSAC) Special Issue: Non-Cooperative Behavior in Networking , 25(6):1216-1223, June 2007.

[18-3] H. Shen and T. Başar, "Pricing under information asymmetry for a large population of users," Telecommunication Systems, 2010 (to appear).

Reprints of the first two papers and a preprint of the third one are attached.

19 Lecture 19: Game-theoretic approach to network security

This lecture will be on the application of non-cooperative game theory and the fictitious play algorithm to network security and intrusion detection. The presentation will be based on the contents of three papers:

[19-1] T. Alpcan and T. Başar, "A game theoretic analysis of intrusion detection in access control systems," Proc. IEEE Conference on Decision and Control (Dec 14-17, 2004; Nassau, The Bahamas), pp. 1568-1573.

[19-2] K. C. Nguyen, T. Alpcan, and T. Başar, "Security games with decision and observation errors," Proc. 2010 American Control Conference (ACC 2010), Baltimore, Maryland, June 30 - July 2, 2010.

[19-3] K. C. Nguyen, T. Alpcan, and T. Başar, "Fictitious play with time-invariant frequency update for network security," Proc 2010 IEEE Multi-Conference on Systems and Control (MSC 2010), Yokohama, Japan, September 8-10, 2010.

Reprints or pre-prints of these three papers are attached.

A comprehensive treatment of game-theoretic approach to security can be found in the forthcoming book by *Alpcan and Başar* [37].

20 Lecture 20: Auctioning divisible resources

This lecture will discuss the application of the Nash equilibrium concept in auctioning divisible resources. The presentation will be based on the contents of three papers:

[**20-1**] R.T. Maheswaran and T. Başar, "Social welfare of selfish agents: motivating efficiency for divisible resources," Proc. IEEE Conference on Decision and Control (Dec 14-17, 2004; Nassau, The Bahamas), pp. 1550-1555.

[**20-2**] R.T. Maheswaran and T. Başar, "On revenue generation when auctioning network resources," Proc. Joint 44th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC'05, Dec 12-15, 2005; Seville, Spain), pp. 7466-7471.

[**20-3**] R.T. Maheswaran and T. Başar, "Efficient signal proportional allocation (ESPA) mechanisms: Decentralized social welfare maximization for divisible resources," IEEE J. on Selected Areas in Communications (JSAC) Special Issue: Price-Based Access Control and Economics for Communication Networks, 24(5):1000-1009, May 2006.

Reprints of the three papers are attached.

21 Appendix: List of sixteen papers attached

[15-1] T. Başar and A. Haurie, "Feedback equilibria in differential games with structural and modal uncertainties," in J. B. Cruz, Jr., editor, Advances in Large Scale Systems, chapter 1, pp. 163–201. JAE Press Inc., Connecticut, May 1984.

[15-2] T. Başar, A. Bensoussan, and S.P. Sethi, "Differential games with mixed leadership: the open-loop solution," J. Applied Mathematics and Computation, 2010 (to appear).

[16-1] T. Alpcan, T. Başar, R. Srikant, and E. Altman, "CDMA uplink power control as a noncooperative game," Wireless Networks, 8:659-670, 2002.

[16-2] T. Alpcan, T. Başar, and S. Dey, "A power control game based on outage probabilities for multicell wireless data networks," IEEE Transactions on Wireless Communications , 5(4):890-899, April 2006.

[16-3] X. Fan, T. Alpcan, M. Arcak, T. J. Wen, and T. Başar, "A passivity approach to game-theoretic CDMA power control," Automatica , 42(11):1837-1847, November 2006.

[17-1] E. Altman and T. Başar, "Multi-user rate-based flow control," IEEE Trans. Communications, 46(7):940-949, July 1998.

[17-2] E. Altman, T. Başar, and R. Srikant, "Nash equilibria for combined flow control and routing in networks: Asymptotic behavior for a large number of users," IEEE Trans. Automat Control, 47(6):917-930, June 2002.

[18-1] T. Başar and R. Srikant, "Revenue-maximizing pricing and capacity expansion in a many-users regime," Proc. IEEE Infocom 2002, Vol. 1, pp. 294-301, New York, New York, June 23-27, 2002

[18-2] H. Shen and T. Başar, "Optimal nonlinear pricing for a monopolistic network service provider with complete and incomplete information," IEEE J. Selected Areas in Communications (JSAC) Special Issue: Non-Cooperative Behavior in Networking , 25(6):1216-1223, June 2007.

[18-3] H. Shen and T. Başar, "Pricing under information asymmetry for a large population of users," Telecommunication Systems, 2010 (to appear).

[19-1] T. Alpcan and T. Başar, "A game theoretic analysis of intrusion detection in access control systems," Proc. IEEE Conf Decision and Control (Dec 14-17, 2004; Nassau, The Bahamas), pp. 1568-1573.

[19-2] K. C. Nguyen, T. Alpcan, and T. Başar, "Security games with decision and observation errors," Proc. 2010 American Control Conference (ACC 2010), Baltimore, Maryland, June 30 - July 2, 2010.

[19-3] K. C. Nguyen, T. Alpcan, and T. Başar, "Fictitious play with time-invariant frequency update for network security," Proc 2010 IEEE Multi-Conference on Systems and Control (MSC 2010), Yokohama, Japan, September 8-10, 2010.

[20-1] R.T. Maheswaran and T. Başar, "Social welfare of selfish agents: motivating efficiency for divisible resources," Proc. IEEE Conf Decision and Control (Dec 14-17, 2004; Nassau, The Bahamas), pp. 1550-1555.

[20-2] R.T. Maheswaran and T. Başar, "On revenue generation when auctioning network resources," Proc. Joint 44th IEEE Conf Decision and Control and European Control Conference (CDC-ECC'05, Dec 12-15, 2005; Seville, Spain), pp. 7466-7471.

[20-3] R.T. Maheswaran and T. Başar, "Efficient signal proportional allocation (ESPA) mechanisms: Decentralized social welfare maximization for divisible resources," IEEE J. on Selected Areas in Communications (JSAC) Special Issue: Price-Based Access Control and Economics for Communication Networks, 24(5):1000-1009, May 2006.

References

[1] T. Başar and G.J. Olsder, Dynamic Noncooperative Game Theory, 2nd edition, Classics in Applied Mathematics, SIAM, Philadelphia, 1999.

[2] G. Owen, Game Theory, 3rd edition, Academic Press, 1995.

[3] D. Fudenberg and J. Tirole, Game Theory, MIT Press, 1991.

[4] N. H. Vorob'ev, Game Theory, Springer-Verlag, Berlin, 1977.

[5] W. Saad, Z. Han, M. Debbah, A. Hjorungnes, and T. Başar, "Coalitional game theory for communication networks [A tutorial]," IEEE Signal Processing Magazine, Special Issue on Game Theory, 26(5):77-97, September 2009.

[6] R. Isaacs, Differential Games, Kruger, NY, 2nd ed., 1975 (First edition: Wiley, NY, 1965).

[7] T. Başar and P. Bernhard, H^∞ Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach, Birkhäuser, Boston, MA, August 1995. (New paperback edition in 2008)

[8] J. M. Smith and G.R. Price, "The logic of animal conflict," Nature, **246**:15-18, 1973.

[9] J. M. Smith, "The theory of games and the evolution of animal conflicts," J. Theoretical Biology, **47**:209-221, 1974.

[10] J. M. Smith, Evolution and the Theory of Games, Cambridge University Press, 1982.

[11] D.P. Bertsekas, Nonlinear Programming, 2nd ed., Athena Scientific, Boston, MA, 1999.

[12] D.P. Bertsekas, Dynamic Programming and Optimal Control, Athena Scientific, Boston, MA, 2007.

[13] D.P. Bertsekas and J.N. Tsitsiklis, Parallel and Distributed Computation: Numerical Methods, Athena Scientific, Boston, MA, 1997.

[14] J. von Neumann and O. Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, Princeton, NJ, 2nd ed., 1947 (first edition: 1944).

[15] J. von Neumann, "Zur theorie der Gesellschaftspiele," Mathematische Annalen, 100:295-320, 1928.

[16] J.F. Nash, Jr., "Equilibrium points in n-person games," Proc. National Academy of Sciences, 36(1):48-49, January 15, 1950.

[17] J.F. Nash, Jr., "Non-cooperative games," Annals of Math., 54(2):286–295, September 1951.

[18] J.B. Rosen, "Existence and uniqueness of equilibrium points for concave N-person games," Econometrica, 33(3):520-534, July 1965.

[19] G. Owen, "Existence of equilibrium pairs in continuous games," Internat J Game Theory, 5:97-105, 1974.

[20] H.von Stackelberg, Marktform und Gleichgewicht, Springer Verlag, Vienna, 1934, (An English translation appeared in 1952 entitled "The Theory of the Market Economy," published by Oxford University Press, Oxford, England.)

[21] D. Chazan and W. Miranker, "Chaotic relaxation," Linear Algebra and Applications, 2:199-222, 1969.

[22] T. Başar, "Relaxation techniques and on-line asynchronous algorithms for computation of noncooperative equilibria," J. Economic Dynamics and Control, 71:531–549, 1987.

[23] S. Li and T. Başar, "Distributed learning algorithms for the computation of noncooperative equilibria," Automatica, 23(4):523–533, July 1987.

[24] T. Başar and S. Li, "Distributed algorithms for the computation of Nash equilibria in linear stochastic differential games," SIAM Journal on Control and Optimization, 27(3):563–578, May 1989.

[25] E. Altman and T. Başar, "Multi-user rate-based flow control," IEEE Transactions on Communications, 46(7):940-949, July 1998.

[26] T. Alpcan, T. Başar, R. Srikant, and E. Altman, "CDMA uplink power control as a noncooperative game," Wireless Networks, 8:659-670, October 2002.

[27] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko, The Mathematical Theory of Optimal Processes, Interscience Publishers, New York, N.Y., 1962.

[28] G. W. Brown, "Iterative solutions of games by fictitious play," in Activity Analysis of Production and Allocation, T. C. Koopmans, Ed. New York: Wiley, 1951, pp. 374376.

[29] J. Robinson, "An iterative method for solving a game," Annals of Math., 54(2):296–301, September 1951.

[30] K. Miyasawa, "On the convergence of learning processes in a 2×2 non- zero-sum two person game," Economic Research Program, Princeton Univ., Princeton, NJ, Tech. Rep. 33, 1961.

[31] D. Monderer and L. S. Shapley, "Fictitious play property for games with identical interests," J. Econ. Theory, 68:258265, 1996.

[32] L. S. Shapley, "Some Topics in Two-person Games," in Advances in Game Theory, L. S. Shapley, M. Dresher, and A. W. Tucker, Eds. Princeton, NJ: Princeton Univ. Press, 1964, pp. 129.

[33] J. S. Shamma and G. Arslan, "Unified convergence proofs of continuous-time fictitious play," IEEE Transactions on Automatic Control, 49(7):1137-1142, July 2004.

[34] J. S. Shamma and G. Arslan, "Dynamic fictitious play, dynamic gradient play, and distributed convergence to Nash equilibria," IEEE Transactions on Automatic Control, 50(3):312-327, March 2005.

[35] K. C. Nguyen, T. Alpcan, and T. Başar, "Security games with decision and observation errors," Proc. 2010 American Control Conference, Baltimore, Maryland, June 30 - July 2, 2010.

[36] K. C. Nguyen, T. Alpcan, and T. Başar, "Fictitious Play with Time-Invariant Frequency Update for Network Security," Proc. of the IEEE Multi-Conference on Systems and Control (MSC) and IEEE CCA, Yokohama, Japan, September 2010.

[37] T. Alpcan and T. Başar, Network Security: A Decision and Game Theoretic Approach, Cambridge University Press, Cambridge, England, 2011 (a pre-publication copy available at ¡http://www.tansu.alpcan.org/book.php¿)

[38] R. Selten, "Reexamination of the perfectness concept for equilibrium points in extensive games," Internat. J. Game Theory, 4:25-55, 1975.

[39] T. Başar, "On the uniqueness of the Nash solution in linear-quadratic differential games," Internat. J. Game Theory, 5:65-90, 1976.

[40] R.B. Myerson, R. Selten, "Refinement of the Nash equilibrium concept," Internat. J. Game Theory, 7:73-80, 1978.

[41] E. van Damme, "A relationship between perfect equilibria in extensive games and proper equilibria in normal form," R. Selten, "Reexamination of the perfectness concept for equilibrium points in extensive games," Internat. J. Game Theory, 13:1-13, 1984.

[42] E. van Damme, Stability and Perfection of Equilibria, Springer-Verlag, New York, 1987.

[43] M. Kreps and R. Wilson, "Sequential equilibria," Econometrics, 50:863-894, 1982.

[44] E. Kohlberg and J.-F. Mertens, "On the strategic stability of equilibria," Econometrica, 54:1003-1037, 1986.

[45] T. Başar, "An equilibrium theory for multi-person decision making with multiple probabilistic models," IEEE Transactions on Automatic Control, AC-30(2):118-132, February 1985.

[46] T. Başar, "A counter example in linear-quadratic games: Existence of non-linear Nash solutions, J. Optimization Theory and Applications, 14(4):425–430, October 1974.

[47] T. Başar, "Hierarchical decision making under uncertainty," in Dynamic Optimization and Mathematical Economics, P.T. Liu, ed., Plenum Press, New York, NY, pp. 205-221, 1979.

[48] T. Başar, "Affine incentive schemes for stochastic systems with dynamic information," SIAM J. Control, 22:199-210, 1984.

[49] T. Başar, "Stochastic incentive problems with partial dynamic information and multiple levels of hierarchy," European J. Political Economy, 5:203–217, 1989.

[50] Y.C. Ho, P.B. Luh, and G.J. Olsder, "A control-theoretic view on incentives," Automatica, 18:167-180, 1982.

[51] T. Başar, "Time consistency and robustness of equilibria in noncooperative dynamic games," in Dynamic Policy Games in Economics, F. Van der Ploeg and A. de Zeeuw, eds., North Holland, Amsterdam, pp. 9-54, 1989.

[52] R. Srikant and T. Başar, "Sequential decomposition and policy iteration schemes for M-player games with partial weak coupling," Automatica, 28(1):95-106, January 1992.

[53] R. Srikant and T. Başar, "Iterative computation of noncooperative equilibria in nonzero-sum differential games with weakly coupled players," J. Optimization Theory and Applications, 71(1):137-168, October 1991.

[54] D. H. Jacobson, "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games," *IEEE Transactions on Automatic Control*, vol. AC-18, pp. 124–131, 1973.

[55] P. Whittle, "Risk-sensitive linear/quadratic/Gaussian control," *Advances in Applied Probability*, vol. 13, pp. 764–777, 1981.

[56] P. Whittle and J. Kuhn, "A Hamiltonian formulation of risk-sensitive linear-quadratic-Gaussian control," *International Journal of Control*, vol. 43, no. 1, pp. 1–12, 1986.

[57] W. H. Fleming and W. M. McEneaney, "Risk sensitive control and differential games," *Springer LN in Control and Info. Sci.*, no. 184, pp. 185–197, 1992.

[58] M. R. James, J. Baras, and R. J. Elliott, "Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems," *IEEE Transactions on Automatic Control*, vol. AC-39, pp. 780–792, April 1994.

[59] P. Whittle, *Risk-Sensitive Optimal Control.* Chichester, NY: John Wiley and Sons, 1990.

[60] P. Whittle, "A risk-sensitive maximum principle: The case of imperfect state observation," *IEEE Transactions on Automatic Control*, vol. AC-36, pp. 793–801, 1991.

[61] Z. Pan and T. Başar, "Model simplification and optimal control of stochastic singularly perturbed systems under exponentiated quadratic cost," *SIAM Journal of Control and Optimization*, vol. 34, no. 5, pp. 1734–1766, September 1996.

[62] Z. Pan and T. Başar, "Backstepping controller design for nonlinear stochastic systems under a risk-sensitive cost criterion," *Proc. 1997 American Control Conference*, Albuquerque, New Mexico, pp. 1278-1282, June (4-6) 1997.

[63] T. Runolfsson, " The equivalence between infinite-horizon optimal control of stochastic systems with exponential-of-integral performance index and stochastic differential games," *IEEE Transactions on Automatic Control*, vol. AC-39, pp. 1551–1563, 1994.

[64] W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*, vol. 25. Springer-Verlag, 1993.

[65] Z. Pan and T. Başar, "Time-scale separation and robust controller design for uncertain nonlinear singularly perturbed systems under perfect state measurements," *International J. Robust and Nonlinear Control*, vol. 6, no. 7, pp. 585–608, August-September 1996.

[66] Z. Pan and T. Başar, "Robust control of dynamic systems under structural perturbation," *Proc. 12th Symp. on Energy Engineering Sciences*, Argonne, IL, pp. 31-39, April 1994.