JOURNAL OF RESEARCH of the National Bureau of Standards
Vol. 85, No. 5, September-October 1980

# Player Aggregation in Noncooperative Games*

**A. J. Goldman† and D. R. Shier****

**National Bureau of Standards, Washington, D.C. 20234**

**March 29, 1978**

A condition is given, under which subsets of the players of a noncooperative game can be combined into "aggregate players" without changing the set of equilibrium-point solutions of the game. The condition is that an individual player's payoff does not depend on the strategy choices of the other players forming the same aggregate player. "Approximate" versions of this result are also formulated and proven.

Key words: Aggregation; equilibrium; game theory; mathematical economics; noncooperative games.

## 1. Introduction

In classroom discussions of game theory, the set of players is typically "given." For applications, however, the number and identities of the players may well be matters for the judgement of the mathematical modeler. The number of stakeholder groups with distinguishable interests in the situation under study, may be so great that their treatment as individual groups would impose an unacceptable complexity of analysis and/or a forbidding burden of data-gathering. Even where this is not the case, the interests and likely actions of the members of some subsets of the players may appear sufficiently parallel (though not perfectly so) as to warrant combining each such subset into a single "aggregate player," in the expectation that the greater clarity of insights from analysis of the smaller (aggregated) game will more than compensate for the concomitant loss of finer detail.[1]

It seems natural, therefore, to investigate, from a mathematical viewpoint, the consequences of such aggregation. The present paper constitutes one such investigation. It is restricted to noncooperative games and to the equilibrium-point notion of "solution"; for completeness, these concepts are defined in section 2 below, where the process of aggregation is also formalized. A closely related concept of aggregation (group equilibrium) is investigated in [3].[2]

In section 3, we present a condition under which aggregation does not change a game's set of solutions. Stripped of its formal trappings, that condition is really rather transparent. Under aggregation, individual players of the original game become able to coordinate their choice of strategy with the choices of the other individual players who make up the same aggregate player. The condition ensures that no advantage can be gained from this new capability, by stipulating that each individual player's payoff (in the original game) is independent of the strategy choices by the other individuals comprising the same aggregate player.

Clearly, the condition of section 3 is a fairly strong one, and it does not capture the notion of aggregating players with parallel (rather than independent) interests. However, the present investigation constitutes, it is hoped, a useful stimulus towards achieving a more realistic formulation and analysis of such aggregated games.

The result in section 3 applies, in particular, to a recent inspector-inspectee game [1,2] in which the inspectee decides whether or not to "cheat" at each of a number of sites which may be examined by the inspector. The implication is that the results of that game's analysis remain essentially unchanged if the inspectee player is disaggregated, even "all the way" to a set of individual "site managers."

*AMS subject classification: 90D10, 90D40, 90D45.

†Current address: Department of Mathematical Sciences, Johns Hopkins University, Baltimore, MD 20218.

**Center for Applied Mathematics, National Engineering Laboratory.

[1] This was the case for several cartel-analysis papers presented at the NSF-sponsored Workshop on Applications of Game Theoretical Analysis to Energy Policy, Rice University, August 1975. The first writer gratefully acknowledges that Workshop as a stimulus for this paper.

[2] Numbers in brackets indicate literature references at the end of this paper.

Section 4 extends the preceding material to "approximate solutions," while section 5 takes up the case in which the condition described above is satisfied only approximately. These topics reflect an expectation that in applied contexts, many mathematical relationships will not (or cannot be known to) hold exactly.

## 2. Games, equilibria, aggregation

Let $n \geqslant 1$ be an integer, and $N = \{1,2,...,n\}$. An $n$-person noncooperative game $G = (X,f)$ consists of an $n$-tuple $(X_1,...,X_n)$ of nonempty sets $X_i$ with Cartesian product $X$, and an $n$-tuple $f = (f_1,...,f_n)$ of functions $f_i : X \rightarrow R_i$ where $R_i$ is a set equipped with an irreflexive binary relation $\varrho_i$. Here $X_i$ is interpreted as the set of strategies or actions open to the $i$-th player, $f_i$ as that player's "payoff function", $R_i$ as the set of possible payoffs or outcomes experienced by that player, and $\varrho_i$ as the relation of (strict) preference by that player among outcomes. The fact that the domain of $f_i$ is $X$, rather than $X_i$, expresses the idea that each player's payoff depends not only on what strategy he chooses, but also on the choices made by other players.

For any $x\varepsilon X$, and $i\varepsilon N$, and any $x_i \varepsilon X_i$, we denote by $(x,i,x_i)$ the member of $X$ obtained from $x$ by changing its $i$-th coordinate to $x_i$. With this notation, a "solution" concept can be defined: $x^0 \varepsilon X$ is called an *equilibrium point* ($EP$) for game $G$ if, for every $i\varepsilon N$ and every $x_i \varepsilon X_i$, the relation

$$f_i(x^0,i,x_i)\varrho_i f_i(x^0) \tag{1}$$

is *false*. That is, if one thinks of the coordinates of $x^0$ as the players' "current" choices of strategies, then no player has an incentive to deviate unilaterally from his current choice. Since the game is regarded as "noncooperative," only unilateral shifts come into consideration, and so the falsity of all relations (1) is sufficient to describe the "stability" of $x^0$. If $n=1$, an $EP$ is simply a strategy that yields a preference-maximal outcome for the (sole) player.

Next we describe an "aggregation" of game $G$. Let $m$ be an integer with $1 \leqslant m \leqslant n$, and let $M = \{1,2,...,m\}$. An $m$-player aggregation $G[B,F]$ of $G$ is specified by the following structure. $B = \{B_1,...,B_m\}$ is a partition of $N$ into nonempty sets; note that the relation $i\varepsilon B_{j(i)}$ defines a function $j : N \rightarrow M$. Let $S_j$ be the Cartesian product of the sets $\{R_i : i\varepsilon B_j\}$; also let $F = (F_1,...,F_m)$ be an $m$-tuple of functions $F_j : S_j \rightarrow T_j$ where each set $T_j$ is equipped with an irreflexive binary relation $\tau_j$, and each function $F_j$ is strictly monotone in each of its arguments. This last condition means that for any $s_j \varepsilon S_j$, for any $i\varepsilon B_j$ with $r_i$ the $i$-th coordinate of $s_j$, and for any $r_i' \varepsilon R_i$,

$$r_i' \varrho_i r_i \text{ implies } F_j(s_j,i,r_i')\tau_j F_j(s_j). \tag{2}$$

This structure defines an $m$-person game as follows. The "players" are $\{B_j : j\varepsilon M\}$. The set of strategies of $B_j$ is $Y_j$, the Cartesian product of $\{X_i : i\varepsilon B_j\}$. Note that the Cartesian product of the players' strategy sets, i.e. of $\{Y_j : j\varepsilon M\}$, is the same set $X$ as for the original game; this permits the symbols "$x$" and "$y$" to be used interchangeably, and corresponds to the idea that we are dealing with aggregation of players and payoffs, but not of strategies. (The same observation justifies the later unambiguous use of notation $(x,j,y_j)$, as an extension of the previous symbol $(x,i,x_i)$.) In the aggregated game, the payoff function for player $B_j$ is $g_j : X \rightarrow T_j$, defined by[3]

$$g_j(x) = F_j[\{f_i(x) : i\varepsilon B_j\}]. \tag{3}$$

The definition of an $EP$ for game $G[B,F]$ is directly analogous to that for $G$.

## 3. The limited-dependence condition

The game $G$ will be said to satisfy the *limited-dependence condition* ($LDC$), relative to partition $B$ of $N$, if for each $i\varepsilon N$ the payoff function $f_i$ is independent of the arguments $\{x_k \varepsilon X_k : k\varepsilon B_{j(i)}\text{-}\{i\}\}$.

[3] In the following notation, the argument of $F_j$ is the member $s_j \varepsilon S_j$ whose $i$-th coordinate, for $i\varepsilon B_j$, is $f_i(x)$.

THEOREM 1: *If the* LDC *holds, and* $x^0$ *is an* EP *for* G[B,F], *then* $x^0$ *is an* EP *for* G.

PROOF: Suppose, to the contrary, that (1) holds for some $i\varepsilon N$ and some $x_i\varepsilon X_i$. By the *LDC*, $f_k(x^0,i,x_i) = f_k(x^0)$ for all $k\varepsilon B_{j(i)}$-$\{i\}$. We can now apply (2) with $j = j(i)$, with $s_j$ having coordinates $\{f_k(x^0) : k\varepsilon B_{j(i)}\}$, and with $r_i' = f_i(x^0,i,x_i)$. The result, using (3), is

$$g_{j(i)}(x^0,i,x_i)\tau_{j(i)}g_{j(i)}(x^0),$$

contradicting the hypothesis that $x^0$ is an *EP* of $G[B,F]$.

THEOREM 2: *Assume each* $\varrho_i$ *is complete and each* $\tau_j$ *is a (strict) partial order. If the* LDC *holds, and* $x^0$ *is an* EP *for* G, *then* $x^0$ *is an* EP *for* G[B,F].

PROOF: Suppose, to the contrary, that there exist $j\varepsilon M$ and $y_j\varepsilon Y_j$ such that $g_j(x^0,j,y_j)\ \tau_j g_j(x^0)$. Denote the coordinates of $y_j$ by $\{x_i : i\varepsilon B_j\}$; then it follows from the *LDC* that $f_i(x^0,j,y_j) = f_i(x^0,i,x_i)$ for each $i\varepsilon B_j$, and so

$$F_j[\{f_i(x^0,i,x_i) : i\varepsilon B_j\}]\ \tau_j\ F_j[\{f_i(x^0) : i\varepsilon B_j\}]. \tag{4}$$

Since $\varrho_i$ is complete and irreflexive for each $i\varepsilon B_j$, $B_j$ has a tripartite partition $B_j = B_j^+ \cup B_j^- \cup B_j^0$ where

$$B_j^+ = \{i\varepsilon B_j : f_i(x^0,i,x_i)\varrho_i f_i(x^0)\},$$
$$B_j^- = \{i\varepsilon B_j : f_i(x^0)\varrho_i f_i(x^0,i,x_i)\} - B_j^+,$$
$$B_j^0 = B_j - B_j^+ - B_j^- = \{i\varepsilon B_j : f_i(x^0,i,x_i) = f_i(x^0)\}.$$

We will show that $B_j^+$ is nonempty, implying a contradiction of the hypothesis that $x^0$ is an *EP* for *G*.

Suppose then that $B_j^+$ is empty. Denote the coordinates of $x^0$ by $x_i^0\varepsilon X_i$, and with an obvious extension of previous notation, define $x'\varepsilon X$ by

$$x' = (x^0, B_j^-, \{x_i : i\varepsilon B_j^-\}) = ((x^0,j,y_j), B_j^0, \{x_i^0 : i\varepsilon B_j^0\}).$$

It follows from the LDC that

$$f_i(x') = f_i(x^0) \qquad (\text{all } i\varepsilon B_j^0), \tag{5}$$

$$f_i(x') = f_i(x^0,j,y_j) \qquad (\text{all } i\varepsilon B_j^-). \tag{6}$$

It is clear that if $B_j^- = \phi$, then $x' = x^0$ and so

$$g_j(x^0) = g_j(x'), \tag{7}$$

while if $B_j^- \neq \phi$ it follows from the definition of $B_j^-$, the monotonicity of $F_j$ and the transitivity of $\tau_j$, that

$$g_j(x^0)\tau_j g_j(x'). \tag{8}$$

Also, it follows from (5) and the definition of $B_j^0$ that

$$g_j(x') = g_j(x^0,j,y_j). \tag{9}$$

Combining (9) with whichever of (7) or (8) applies yields a contradiction to the initial assumption that $g_j(x^0,j,y_j)\tau_j g_j(x^0)$. Thus $B_j^+$ is nonempty, as desired. This completes the proof of Theorem 2.

We are indebted to colleague S. Haber for pointing out that Theorem 2's requirement of completeness for every $\varrho_i$ can be relaxed as follows. Recall that elements $u$ and $u'$, in the domain of binary relation $\varrho$, are called *$\varrho$-incomparable* if neither $u\varrho u'$ nor $u'\varrho u$ holds. (For example, if $\varrho$ is irreflexive then equality of elements implies their incomparability.) Let us call the aggregation scheme $[B,F]$ *incomparability-preserving* if, for all $j\varepsilon M$, whenever $s_j = \{u_i : i\varepsilon B_j\}$ and $s_j' = \{u_i' : i\varepsilon B_j\}$ are members of $S_j$ such that $u_i$ and $u_i'$ are $\varrho_i$-incomparable for all $i\varepsilon B_j$, it follows that $F_j(s_j)$ and $F_j(s_j')$ are $\tau_j$-incomparable. Then Theorem 2, without the assumption of completeness for the $\varrho_i$'s, holds for incomparability-preserving aggregations. To adapt the preceding proof so as to establish this generalization, omit the last expression in the definition of $B_j^0$, so that

$$B_j^0 = \{i\varepsilon B_j : f_i(x^0,i,x_i) \text{ and } f_i(x^0) \text{ are } \varrho_i\text{-incomparable}\}.$$

Relations (7) and (8) are proved as before, and either of them together with the initial assumption $g_j(x^0,j,y_j)$ $\tau_j g_j(x^0)$ implies that $g_j(x^0,j,y_j)\tau_j g_j(x')$. This however, together with (5), (6) and the definition of $B_j^0$, yields a contradiction to the hypothesis that the aggregation is incomparability-preserving.

Taken together, Theorems 1 and 2 assert that under mild restrictions on the $\varrho_i$'s and $\tau_j$'s, the *LDC* is a *sufficient* condition for the "aggregation" transition from $G$ to $G[B,F]$ to leave the set of equilibrium-point "solutions" unchanged. (It is a condition on the pair $(G,B)$, yielding the desired invariance for every choice of $F$.)

However, this sufficient condition is *not* also a necessary condition. An example which does not satisfy the *LDC*, but for which the set of equilibrium points is unchanged by aggregation, can be based on the game $G$ shown in figure 1. Here $n=2$, $X_1=\{A,B\}$, $X_2=\{a,b\}$, $\varrho_1$ and $\varrho_2$ are the numerical ">" relation, and the payoff functions $f_1$ and $f_2$ are identical ($f_1=f_2=\bar{f}$) with

$$\bar{f}(A,a)=2, \bar{f}(A,b)=\bar{f}(B,a)=1, \bar{f}(B,b)=0.$$

The only *EP* of $G$ is $(A,a)$. Now consider any aggregation $G[B,F]$ with $B_1=\{1,2\}$, so that $m=1$. Since $F_1$ is monotone, $G[B,F]$ will also have $(A,a)$ as its only equilibrium point. Thus the solution-set is unchanged by the aggregation, although the *LDC* does not hold.

The idea of this example can readily be extended to examples with $m>1$. It seems doubtful that a "nice" necessary and sufficient condition, verifiable without having to solve the game $G$ (which would defeat the purpose of the aggregation), can be found.

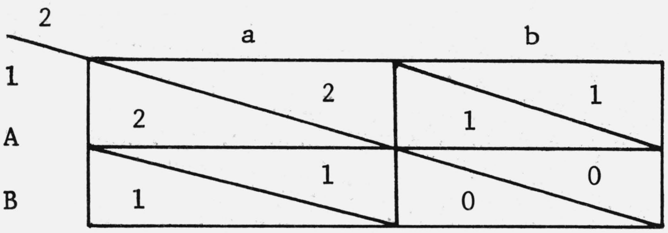


*Figure 1:* An Example.

## 4. Approximate equilibrium points

Since the topics of this section and section 5 deal with quantitative rather than qualitative relationships, we now take all sets $R_i$ and $T_j$ to be the set $R$ of real numbers, and think of the relations $\varrho_i$ and $\tau_j$ as the ordinary numerical "greater than" relation. The payoff-aggregation functions $F_j$ will for simplicity be taken to be summations, i.e. (3) becomes

$$g_j(x) = \Sigma\{f_i(x) : i\varepsilon B_j\} \tag{8}$$

For each $i\varepsilon N$ and each $x\varepsilon X$, the quantity

$$M_i(x) = \sup\{f_i(x,i,x_i) : x_i\varepsilon X_i\} - f_i(x)$$

is nonnegative. If $\delta = (\delta_1,\ldots,\delta_n)$ is an $n$-tuple of positive real numbers, and if $x^0\varepsilon X$ satisfies

$$M_i(x^0) \leqslant \delta_i \qquad (\text{all } i\varepsilon N), \tag{9}$$

then $x^0$ will be called a $\delta$-*EP* of game $G$. Approximate *EP*'s of $G[B,F]$ are defined analogously.

THEOREM 3: *Assume* n-*tuple* $\delta$ *and* m-*tuple* $\delta'$ *satisfy* $\delta_i \geqslant \delta'_{j(i)} > 0$ *for all* $i\varepsilon$N. *If the* LDC *holds and* $x^0$ *is a* $\delta'$-EP *of* G[B,F], *then* $x^0$ *is a* $\delta$-EP *of* G.

PROOF: Suppose, to the contrary, that some $i\varepsilon N$ and $x_i \varepsilon X_i$ satisfy

$$f_i(x^0,i,x_i) - f_i(x^0) > \delta_i.$$

Define $y_{j(i)} \varepsilon Y_{j(i)}$ to have coordinates $x_i \varepsilon X_i$ and $x_k^0 \varepsilon X_k$ for all $k\varepsilon B_{j(i)} - \{i\}$. Then by (8) and the *LDC*,

$$\begin{aligned} g_{j(i)}(x^0,j(i),y_{j(i)}) - g_{j(i)}(x^0) &= \Sigma\{f_k(x^0,i,x_i) - f_k(x^0) : k\varepsilon B_{j(i)}\} \\ &= f_i(x^0,i,x_i) - f_i(x^0) > \delta_i \geqslant \delta'_{j(i)}, \end{aligned}$$

contradicting the hypothesis that $x^0$ is a $\delta'$-*EP* of $G[B,F]$.

THEOREM 4: *Assume positive* n-*tuple* $\delta$ *and* m-*tuple* $\delta'$ *satisfy* $\delta'_j \geqslant \Sigma\{\delta_i : i\varepsilon B_j\}$ *for all* $j\varepsilon$M. *If the* LDC *holds and* $x^0$ *is a* $\delta$-EP *of* G, *then* $x^0$ *is a* $\delta'$-EP *of* G[B,F].

PROOF: Choose any $j\varepsilon M$ and any $y_j \varepsilon Y_j$; let the coordinates of $y_j$ be $\{x_i \varepsilon X_i : i\varepsilon B_j\}$. By hypothesis,

$$f_i(x^0,i,x_i) - f_i(x^0) \leqslant \delta_i \qquad (\text{all } i\varepsilon B_j),$$

which by the LDC can be rewritten

$$f_i(x^0,j,y_j) - f_i(x^0) \leqslant \delta_i \qquad (\text{all } i\varepsilon B_j).$$

Summing over all $i\varepsilon B_j$ and applying (8), we obtain

$$g_j(x^0,j,y_j) - g_j(x^0) \leqslant \Sigma\{\delta_i : i\varepsilon B_j\} \leqslant \delta'_j$$

for all $j\varepsilon M$ and $y_j \varepsilon Y_j$, establishing the desired result.

Assuming the *LDC* holds, Theorem 3 provides a "degree of approximation" for $x^0 \varepsilon X$ as an (approximate) *EP* of $G$, in terms of its "degree of approximation" as an *EP* of $G[B,F]$. Theorem 4 does the reverse. The two theorems are not intended to apply simultaneously to the same pair $(\delta,\delta')$, and do not so apply except in the trivial case (all $|B_j| = 1$) of "no aggregation".

## 5. The approximate LDC

In this section, the notation $Z_i$ will be used for the Cartesian product of the sets $\{X_k : k\varepsilon B_{j(i)} - \{i\}\}$. Note that for any $j\varepsilon M$ and $i\varepsilon B_j$, each $y_j \varepsilon Y_j$ can be uniquely represented as $y_j = (x_i,z_i)$ with $x_i \varepsilon X_i$ and $z_i \varepsilon Z_i$.

Observe that the *LDC* is equivalent to the following condition: for each $x\varepsilon X$, each $j\varepsilon M$, each $i\varepsilon B_j$, and each $y_j = (x_i,z_i) \varepsilon Y_j$,

$$f_i(x,j,y_j) = f_i(x,i,x_i).$$

This suggests the following definition. Let $\lambda = (\lambda_1,\ldots,\lambda_n)$ be an $n$-tuple of positive numbers. Then we say that the *LDC* $\lambda$-*holds* if for each $x\varepsilon X$, each $j\varepsilon M$, each $i\varepsilon B_j$, and each $y_j = (x_i,z_i) \varepsilon Y_j$,

$$|f_i(x,j,y_j) - f_i(x,i,x_i)| \leqslant \lambda_i.$$

THEOREM 5: *Assume* n-*tuple* $\delta$ *and* m-*tuple* $\delta'$ *satisfy* $0 < \delta'_{j(i)} \leqslant \delta_i - \Sigma\{\lambda_k : k \varepsilon B_{j(i)} - \{i\}\}$ *for all* $i\varepsilon N$. *If the* LDC $\lambda$-*holds and* $x^0$ *is a* $\delta'$-EP *of* G[B,F], *then* $x^0$ *is a* $\delta$-EP *of* G.

THEOREM 6: *Assume positive* n-*tuple* $\delta$ *and* m-*tuple* $\delta'$ *satisfy* $\delta'_j \geqslant \Sigma\{\delta_i + \lambda_i : i\varepsilon B_j\}$ *for all* $j\varepsilon M$. *If the* LDC $\lambda$-*holds and* $x^0$ *is a* $\delta$-EP *of* G, *then* $x^0$ *is a* $\delta'$-EP *of* G[B,F].

The proofs of Theorems 5 and 6 are straightforward extensions of those of Theorems 3 and 4, respectively, and therefore are omitted.

## 6. References


[1] Goldman, A. J. and Pearl, M. H., The dependence of inspection-system performance on levels of penalties and inspection resources, *J. Research National Bureau of Standards,* Vol. *80B* (1976), 189-236.

[2] Pearl, M. H. and Goldman, A. J., A game-theoretic model of inspection-resource allocation, to appear in *J. Research National Bureau of Standards.*

[3] Szidarovszky, F. and Galantai, A., On new concepts of game theory, in C. A. Brebbia (ed.), *Applied Numerical Modelling,* Halsted Press (John Wiley & Sons, New York, 1978).