A Game-Theoretic Approach to Multipolar Stability*

Rudolf Avenhaus and Bernhard von Stengel

Universität der Bundeswehr München
Neubiberg, Germany

Shmuel Zamir

The Hebrew University of Jerusalem
Jerusalem, Israel

1. Problem Formulation and Overview

This paper proposes a simple and tentative mathematical description of multipolar military stability that uses concepts of cooperative game theory. Its purpose is to propose a plausible approach and to stimulate a discussion. At this stage, we do not pretend to present here a complete theory of stability. The paper is the successor to an earlier approach by Avenhaus, Cao, von Stengel and Zamir (1992), where the idea was to use game theory for analyzing stability not only quantitatively but also conceptually: Game-theoretic solution concepts, like the core for cooperative games, or the Nash equilibrium for noncooperative games, are based on some notion of stability, which may suggest to consider the corresponding outcome as "stable" in some sense.

As the discussions throughout the symposium of the present proceedings volume indicated, there is a need to clarify what is actually meant by stability. The model presented in the following is based on the simple operational definition that "no participant has an incentive to attack another one". Stability will not be defined as a special instance of a game-theoretic solution concept, since for a starting point of a discussion, as this paper is meant to be, this would be too technical. Game theory will be used in a part of the model that defines the peaceful interaction of the players as a cooperative game. In this game, the "power" of a player is the target of a potential aggressor, who tries to increase his power by a successful military action, and risks losing it (or part of it) by a defeat. This power is measured by the Shapley value, which is a suitable and nontrivial concept for this purpose.

The approach can be summarized as follows: A cooperative game of n players, representing the nations among which stability is studied, describes their non-military interactions. The game reflects the economic and political importance of the individual players as well as the rules of collective decision making, like in the security

* This research was supported in part by the German-Israeli Foundation for Research and Development (GIF) and in part by the Volkswagen foundation. Jerome Bracken and Michel Rudnianski suggested an approach that takes into account gains as well as losses as possible outcomes of a military conflict.

council of the United Nations or in the European Parliament. In this game, the *power* of a player may be represented by his Shapley value. If a player (called attacker) goes to war against another player (called defender), the outcome of the war will change the game and thereby the power of each player. The situation is considered stable if no player can expect an increase in power from an attack. Thereby, a war is assumed to have two possible outcomes, which change the game so that the post-war power of the attacker will be larger if he is successful, and smaller if he is not. The mathematical expectation of the power of the attacker, determined by the probability of winning, is then taken to determine his incentive of attack. If this expectation is not larger than his original power, the situation between the respective players is considered as stable, otherwise it is unstable. Overall stability exists if the situation is stable for any pair of players, each taken as attacker and defender.

At first glance, this concept of stability looks as if it were simply one of "pairwise bipolar stability", thereby not taking into account the multipolar situation with more than two players. The Shapley value, however, does take into account *all possible coalitions*. Furthermore, military coalitions may also be incorporated as natural extensions of the model. For simplicity, however, only pairwise confrontations are considered at this first stage of the study, since the case of a single aggressor seems to be the most frequent case actually observed.

Also, defensive alliances are not considered explicitly, but the likelihood of a collective reaction to an aggressive act may be taken into account implicitly. To illustrate this, consider the question why a militarily strong nation normally does not try to conquer a weak nation even if the risk of losing such a war is very small. Why is it beyond imagination in today's peaceful Western Europe that, say, France invades and annexes Luxembourg? Here, the political and military aspects are highly intertwined. First, the potential increase in power to France would be very small, even if the international community or other states would not react by military force. In the model, the change in the game is such that the probability of winning would have to be very high, say 95 percent. Second, even if the actual probability should be that high in military terms, one might nevertheless argue that the situation is stable because of the imponderabilities to the aggressor, since, with some chance, the international community would react militarily and expel the invader. This could be modeled by a reduced probability of winning that incorporates some estimate of the likelihood of a collective military reaction of the other nations. In a chaotic situation among the states, where war gains are not likely to be undone by the interference of stronger powers or alliances, the winning probability may very well be determined mainly by military factors.

The outlined approach has the following crucial components. First, there is the peaceful interaction of the players modeled by a cooperative game. The game used here, which will be explained technically in the second section, is defined — for the sake of simplicity — as a weighted majority game, essentially describing a system of voting. The Shapley value in this game counts how often the (weighted) vote of the player is crucial in reaching the required majority and thereby represents his power. In this game, the "weight" of a player defines his importance, as reflected by, for example, the gross national product of the respective nation, its geographic area, population, etc.. It also reflects — in some admittedly vague way — factors such as international stature, or the political influence in the international community, for example as in a decision making body like the United Nations security council. Thus, the game describes the status quo, which will be changed by a successful or unsuccessful attack. In particular, the weight of the attacking player is assumed to increase (in some proportion to the weight of the victim party) if the attack is successful, and to decrease if it is not. That is, the weights are redistributed after a successful or failed attack.

The second crucial part of the model concerns the probability of a successful attack. For each attacker-defender pair, a threshold for this probability is easily computed from the attacker's Shapley values before and after a successful or failed attack. For example, if, in the given game, the attacker has a Shapley value of 0.2, which increases to 0.25 if he successfully attacks another player, and drops to zero if he loses, the situation between these two players is stable if, and only if, the probability of winning does not exceed 0.8. These threshold probabilities for a successful attack, for all pairs of players, form the output matrix of the game-theoretic analysis. Usually, for a possible attack of a player with large weight against a player with small weight, this threshold is high. In other words, the probability of winning has to be high in order to make the attack worthwhile. On the other hand, since a country of small weight may gain significantly from successfully attacking one with a larger weight, the respective threshold is usually low (see the examples in section 4). Thus, the military force structures, about which nothing has been said so far, should be such that these critical probabilities for a successful attack are not exceeded, given the political environment. If follows that countries with large weight should have larger forces, since there is more power to be protected, and countries with less weight should have smaller forces, in order to deny them the chance to conquer the power of an important country.

In summary, the game-theoretic model yields, for each pair of conceivable antagonists (players) in a multipolar international system, mutual threshold values for the probability of a successful attack, based on the resulting gain or loss of power. For the sake of stability, these threshold probabilities must not be exceeded by the actual probabilities associated with the military forces of the respective pairs, which may be estimated based on political and military judgement. The latter may be assisted by appropriate systems analyses using models such as, for example, those discussed in this volume by Huber and Schindler (1993) and by Hofmann, Tolk and Schnurer (1993).

In fact, the principal approach pursued here, of separating the non-military from the military aspects of stability, is along the same lines of what Huber and Schindler (1993) discuss in their paper. However, they assume the stability threshold values for the probabilities of a successful attack or defense as given, depending on the quality of the non-military relations between the parties, and they estimate the force and operational parameters satisfying the threshold values.

2. Cooperative Games and the Shapley Value

Consider n *players* that represent nations or states in a multipolar situation. Their non-military interaction will be modeled by a *cooperative game*, which is a concept technically known as an n-person game in characteristic function form with transferable utility. This concept, and other concepts relevant to the proposed approach, such as simple games, weighted majority games, and the Shapley value, will be explained in this section.

Let the players be identified by their numbers $1, \ldots, n$ and let $N = \{1, \ldots, n\}$ be the player set. A *coalition* is any subset S of N, that is, any set of players. A *cooperative game* is a function v that assigns to any coalition S a real number $v(S)$, with the standardization $v(\emptyset) = 0$. (The function v defined on the set of subsets of N is also called the "characteristic function" of the game, in particular when it is derived from a more detailed description as done by von Neumann and Morgenstern (1944), who introduced this concept and many others in their fundamental work on game theory.) The interpretation of $v(S)$, which may be called the *worth* of the coalition S, is a measure of what the players in S can achieve together on their own when they form a coalition independent of the other players. In particular, $v(\{i\})$ is what a single player i can achieve by himself, his individual worth. For a coalition of two or more players, nothing is said about how this coalition is formed, in what way the players come to an agreement, or how likely it is that they join to form a coalition at all. The key point is that $v(S)$ is realized *only* by the agreement and cooperation of *all* members of S. The worth $v(S)$ of the coalition S is an aggregate figure, that is, a single number, and not a tuple of numbers giving to every player from the coalition an individual "payoff". If the coalition S forms, the total worth $v(S)$ can be split in any way among the members of S (therefore, v is sometimes also called a transferable utility game). The manner in which the members of S should split $v(S)$ among themselves (if the coalition S forms) is not part of the data of the game but rather of the solution concept.

A *simple* game is a cooperative game where the worth $v(S)$ of any coalition S is either 0 or 1. A simple game can be thought of as a decision making system; a coalition S may be either sufficient to pass a resolution (if $v(S) = 1$, in which case it is a *winning* coalition) or not (if $v(S) = 0$, in which case it is a *losing* coalition). A particular instance is a *weighted majority* game where each player i (for example, a party in a parliament) has a certain nonnegative weight w_i (for example, number of seats), and a certain *quota* $q \in [0,1]$, for example $q = 1/2$, of the total vote is needed to pass a resolution. That is, if the total weight $w(S)$ of a coalition S is defined by $w(S) = \sum_{i \in S} w_i$, then S is a winning coalition ($v(S) = 1$) if it has a majority with $w(S) > q \cdot w(N)$, otherwise a losing coalition ($v(S) = 0$).

There are a number of "solution concepts" that have been suggested for cooperative n-person games. One of them is the *Shapley value* introduced by Shapley (1953). It is a way to distribute the worth $v(N)$ of the whole coalition to the players, resulting into a *value* $\phi(i)$ for every player i, with $\sum_{i \in N} \phi(i) = v(N)$. Starting from certain plausible properties which are to be satisfied by ϕ, it has been proved by Shapley (1953) that the value $\phi(i)$ of a player is the *expected marginal contribution* of player i to the worth of a coalition, assuming a certain probability distribution on the possible ways in which a coalition can be formed. More precisely, suppose there is a random order of the players given by a permutation π of the set N, where $\pi(i)$ is the position of player i (for example, if $n = 5$ and the order of the players is 5, 2, 1, 3, 4, then $\pi(5) = 1$, $\pi(2) = 2$, $\pi(1) = 3$, $\pi(3) = 4$ and $\pi(4) = 5$), and assume that the players enter a room in this order. At a certain moment, the players in the room form a coalition S which has the worth $v(S)$. Each player i who enters the room encounters the coalition $b_\pi(i)$ of the players who came before him, given by $b_\pi(i) = \{j \in N \mid \pi(j) < \pi(i)\}$. The entering player, joining the coalition, is then given the increase in worth due to his contribution, that is,

$$v(b_\pi(i) \cup \{i\}) - v(b_\pi(i)).$$

In this process, the permutation π is taken at random, with the uniform distribution on the set S_n of all permutations of N. The Shapley value $\phi(i)$ for player i is then the expectation of this number,

$$\phi(i) = \frac{1}{n!} \sum_{\pi \in S_n} \Big(v(b_\pi(i) \cup \{i\}) - v(b_\pi(i)) \Big).$$

The Shapley value is uniquely defined for any cooperative game.

For example, in the weighted majority game with three players of the weights $(w_1, w_2, w_3) = (2, 1, 1)$, and with the majority quota $q = 2/3$, player 1 is a "veto" player that is required for any winning coalition, but can not form a winning coalition on his own. The Shapley values $(\phi(1), \phi(2), \phi(3))$ are given as $(2/3, 1/6, 1/6)$ since the marginal contribution of player 1 is 1 (that his, he is "pivotal" in getting the majority) in any permutation except those where he is in the first position. In simple games, the Shapley value counts the number of times a player is pivotal (normalized so that the sum of the Shapley values of the players is $v(N) = 1$). In weighted majority games, this may substantially differ from the proportion of the weights, reflecting closely the "power" of the respective player (or party in a parliament). As a measure of power, the Shapley value is a convincing and quite applicable concept; see, for example, Roth (1988).

3. The Model

As mentioned in the introduction, a cooperative game with n players shall model the following situation: Each player represents a nation or state as the smallest acting entity; "stability" shall be studied only among players. The game represents a non-military interaction of the players, economically and politically, and it should reflect the "importance" of a player in some sense, in particular in acting with other players. This game is changed by a war, where the conceivable gain of a successful attack is greater economic and political influence, and a loss of it in case the attack fails.

For simplicity, the game is assumed to be a weighted majority game, with the weight w_i of player i reflecting his importance (for example, gross national product, population, number of seats in an international institution or a similar figure). The concept of "voting" suggests a politically rather than economically determined interaction, captured by the decision making mechanism in the status quo situation. As mentioned, the Shapley value in a voting game measures the relative power of a player in the sense of "how often" he casts the crucial vote. This is of course a simplistic way of thinking of international interaction, but it is at least a first approach since the Shapley values of voting games can be quite interesting (see Roth (1988)).

If player i attacks another player j, then this war results in a change of the weights: If the attacker wins the war, one may think of his weight being increased from w_i to $w_i + w_j$ since the defender's weight is conquered, or it could be $w_i + 0.6w_j$ (or some other combination of the weights) to account for war losses and lack of cooperation with the conquered nation. The defender, who lost, may be assumed to have zero weight afterwards. This may, conceivably, be thought of as the only "gain" of the attacker (in a very destructive war), i.e., a change of w_j to zero and w_i staying the same. An unsuccessful attack (a lost war) may leave the defender's weight w_j unchanged (or reduce it to $0.7w_j$, for example), and the attacker's weight may be changed to zero (by his losing all political influence, for example). A less harmful outcome to the attacker may be modeled by his weight being decreased by the defender's, from w_i to $w_i - w_j$ (or to zero, respectively, if this number is negative), so in this case only the weight of the defender is at stake.

The original pre-war game has a certain Shapley value ϕ. Denote by ϕ_w the Shapley value of the new game resulting from a successful attack of i against j and by ϕ_l the Shapley value of the game resulting from the failure of such an attack. Then, if player i attacks player j, the Shapley value for player i after he wins is $\phi_w(i)$ and after he loses is $\phi_l(i)$. If the probability of winning is p, then the attacker's expected Shapley value is $p\phi_w(i) + (1-p)\phi_l(i)$. It is assumed that player i has an incentive of attack (so there is a lack of stability) if this expectation is greater than his original Shapley value $\phi(i)$, which is equivalent to

$$p > \frac{\phi(i) - \phi_l(i)}{\phi_w(i) - \phi_l(i)} =: c_{ij}$$

Plausible assumptions may be required to guarantee that c_{ij} is well defined and is between 0 and 1, such as "winning" is better than "losing" or that the status quo is not worse than "losing".

One may ask whether it is reasonable to take mathematical expectations of the Shapley values as the decision criterion for a "promising attack". Alternatively, one could use an expected utility attached to the Shapley values, or in addition to the expectations, the variances. As a basic model, however, it is reasonable to assume that the Shapley value is linear in the probabilities since it is itself an expectation (the expected marginal contribution to the worth of the random coalition that the player may join) of which an expectation can be taken again.

Clearly, this model is not confined to only two outcomes of a war. In fact, a probability distribution with any number of outcomes can be assumed to model the risk underlying an attack. However, the restriction to victory or defeat only characterizes the distribution by a single probability p of winning for any pair (i, j)

Table 1: Maximal probabilities for a successful attack tolerable in a stable situation based on the Shapley value for a voting game with weights 8, 5, 1, 1, 1, 1 for players 1 to 6 and majority quota $q = 1/2$.

defender j / attacker i	1	2	3, 4, 5, 6
1	—	0.615	0.285
2	0.067	—	0.500
3, 4, 5, 6	0.067	0.667	0.667

of attacker and defender, for which a critical threshold is given by c_{ij}, which is easy to grasp and computed from the games alone.

4. Examples

In this section, a simplified example of a multipolar situation will be discussed with some variations, although not with all of those mentioned in the previous section. There are six players, one large, one intermediate and four small. These are represented, in a simple weighted majority game, by the weights 8, 5 and 1, respectively, so one can think of a parliament with 17 seats with the seats for the six players given by 8, 5, 1, 1, 1, 1. If the majority quota is $q = 1/2$, then at least 9 votes are required to reach a majority. The Shapley values in this game are given by 2/3 for the large player and by 1/15 for *every* other player, including the one with 5 seats. The normalized weights of the players are 0.470, 0.294, 0.059, 0.059, 0.059, 0.059, whereas their Shapley values are 0.667, 0.067, 0.067, 0.067, 0.067, 0.067. That is, the large player has a fraction of 0.470 of the votes but a Shapley value of 0.667 since he needs only one additional player to form a winning coalition. On the other hand, the player with 5 votes, that is, a fraction of 0.294, has the same Shapley value of 0.067 as any one of the single-vote players with a vote fraction of 0.059, since he has no more power than a small player in being pivotal for a winning coalition.

If player i attacks player j, the change of weights for the game is assumed as follows: If i wins, his weight will be increased from w_i to $w_i + w_j$ and that of the defender reduced from w_j to 0. If the attacker i loses the war, then his weight is reduced to $w_i - w_j$ (or to zero if this number is negative), and the weight w_j of the defender will be unchanged — he will not be particularly rewarded for having defeated the attacker. The resulting critical probabilities c_{ij} (rounded to three digits) for a successful attack are given in Table 1.

If player 1 successfully attacks player 2 or one of the players 3 to 6, his Shapley value will rise to 1 (which is not shown in Table 1) since he then has the absolute majority of votes. If the attack is unsuccessful, however, his weight will be reduced

Table 2: Maximal probabilities for a successful attack for the game of Table 1, but with majority quota $q = 2/3$.

defender j / attacker i	1	2	3, 4, 5, 6
1	—	0.333	0.000
2	0.333	—	0.444
3, 4, 5, 6	0.061	0.067	0.667

to 3 respectively to 7, with new (lower) Shapley values of 0.133 and 0.533 (again not shown), which leads to the displayed critical probabilities of $c_{12} = 0.615$ and $c_{13} = 0.285$. The second figure is much lower than the first, since the risk to the attacker is smaller because of the rule that an attacker only loses the weight of the defender if the attack fails. So for the situation to be stable, player 1 should not have too many forces to be able to attack one of the small players successfully. On the other hand, any of the small players $i = 2, \dots, 6$ gets the absolute majority if he succeeds in an attack on player 1, and only risks that his own Shapley value goes to zero if the attack fails. This results in the small critical probabilities of $c_{i1} = 0.067$ in the first column.

With the same weights, the game changes drastically if the majority quota is raised to $q = 2/3$, where at least 12 or more of the 17 seats are required for a majority. The Shapley values for the first three players are then given by 8/15, 1/3 and 1/30, respectively, so the Shapley values for all players are approximately 0.533, 0.333, 0.033, 0.033, 0.033, 0.033. In particular, the Shapley value for the second player in comparison to one of the small players is ten times larger here, which also makes him an interesting target for an attack by one of the small players, represented by the low critical probability of $c_{32} = 0.067$, which is shown, among the other figures, in Table 2.

An interesting entry in this table is certainly $c_{13} = 0$ which makes stability impossible. This is due to the fact that for $q = 2/3$, the games with weights 8, 5, 1, 1, 1, 1 and 7, 5, 1, 1, 1, 1, where the second represents the game after player 1 lost an attack, have the same Shapley value for all players, so player 1 does not risk anything by an attack.

In order to avoid such an effect, a "stability-increasing measure" would be to punish an attacker i who failed more heavily, by not only subtracting the weight w_j of the defender j from his weight w_i but also reducing that weight to, say, half its previous value, i.e., w_i is replaced by $w_i/2 - w_j$ after i loses his attack against j. The new game with this assumption produces Table 3.

Whereas this new rule does not change the critical probabilities for a small player, who in any case would lose all his weight after a lost attack, it drastically enhances the stability between a large player as a potential attacker and a small defender.

Table 3: Maximal probabilities for a successful attack for the game of Table 2 (with $q = 2/3$) where an attacker who lost has a reduction in weight irrespective of the defender's weight.

defender j / attacker i	1	2	3, 4, 5, 6
1	—	0.533	0.933
2	0.333	—	0.583
3, 4, 5, 6	0.061	0.067	0.667

The desirably high critical probability of $c_{13} = 0.933$ would be increased further if a losing attacker were to lose all his weight. Examples 2 and 3 demonstrate the critical assumptions about the effects of a won or lost war on the weights in the game.

Table 4: Shapley values of the countries represented in the European Parliament, regarded as a weighted majority game with majority quota $q = 1/2$ and $q = 2/3$, respectively.

country	weights		Shapley value	
	seats	fraction	$q = 1/2$	$q = 2/3$
Germany	99	0.175	0.181	0.178
Great Britain, France, Italy	87	0.153	0.162	0.162
Spain	64	0.113	0.117	0.106
The Netherlands	31	0.055	0.045	0.044
Belgium, Greece, Portugal	25	0.044	0.037	0.039
Denmark	16	0.028	0.025	0.029
Ireland	15	0.026	0.025	0.029
Luxembourg	6	0.011	0.009	0.010

Finally, Table 4 lists the Shapley values for the countries in the European parliament, which is regarded as a weighted majority game. Each country has a certain number of seats (shown in decreasing order) which are taken as weights. Two games are considered, for the majority quota $q = 1/2$ and $q = 2/3$, respectively. Note that in both cases, the Shapley values are similar to the relative proportion of the weights. This can be expected from the probabilistic interpretation of the Shapley value in a weighted majority game if there is a large number of players, none of which is by himself close to a majority. However, the Shapley value is sensitive to slight variations of the data since they may change the winning coalitions. For example, if $q = 2/3$, then two thirds of the total number (567) of votes are 378 votes, but *more* than these, that is, 379 votes, are required to reach a majority according to the definition of a weighted majority game. If the quota is changed so that 378 votes suffice, the Shapley values change very little except for Denmark and Ireland, which no longer have the same Shapley value of 0.029, but 0.032 and 0.023 instead. The reason is that then Denmark is pivotal in joining a coalition with players of weights 99, 87, 64, 31, 25, 25, 25 and 6 but Ireland is not. This illustrates that any quantitative output of the suggested model, like the Shapley values or the critical probabilities (which are not shown for this example), should always be subjected to a sensitivity analysis.

References

Avenhaus, Rudolf, Buyang Cao, Bernhard von Stengel and Shmuel Zamir (1992). Discussion paper on game-theoretic models of multipolar stability.

Hofmann, Hans W., Andreas Tolk and Robert Schnurer (1993). On the impact of combat support systems on military stability: design and results of simulation experiments with the closed division/corps combat simulation model KOSMOS. A chapter in this volume.

Huber, Reiner K. and Otto Schindler (1993). Military stability of multipolar power systems. An analytical concept for its assessment, exemplified for the case of Poland, Byelarus, the Ukraine and Russia. A chapter in this volume.

Roth, Alvin E., ed. (1988). The Shapley Value. Essays in honor of Lloyd S. Shapley. Cambridge University Press, Cambridge.

Shapley, Lloyd S. (1953). A value for n-person games, Ann. Math. Studies 28, pages 307–318.

von Neumann, John and Oskar Morgenstern (1944). Theory of Games and Economic Behavior. Princeton University Press, Princeton, NJ.