Erik Moberg:

Models of International Conflicts and Arms Races

Originally published in Cooperation and Conflict, Nordic Studies in International Politics, 1966, nr 2, pp 80-93.

1. Introduction

In this essay a series of models will be presented which deal with international conflicts and arms races. The essay is entirely theoretical; not even operational definitions of essential variables, such as "arms level", are given. The purpose of the models is twofold. First it is thought that models of this kind might be useful for clarification of concepts. In the essay this is illustrated by a discussion of some strategic concepts, such as "balance of power" and "deterrence", with reference to the models. Second, there is a heuristic purpose, i. e. it is thought that models may aid fantasy and that therefore an exploration of models of this type might result in a better understanding of possible relations between nations.

In all the models there are only two nations and the focus of interest is on these nations' decisional situations.

2. A simple two-nation model

Two nations A and B have the arms levels x and y respectively. Here "arms level" is thought of as a comprehensive measure of a nation's total amount of armaments. Each nation can vary its own arms level. Concerning the course of an eventual war between these nations the following assumptions are made. First, the nation with the highest arms level wins the war in the sense that when the war is finished the victor has complete control over A and B.[1] This means that the kind of war taking place within the model is a total war leading to total victory for the one party and unconditional surrender for the other. As a result of the war the victorious nation gets, or rather takes, the other nation. However, in spite of the victory it is not certain that the war is profitable for the winner. This is so because, in order for the war to be profitable, the winner must also value the war trophy, i. e. the other nation, higher than the various losses incurred in the war. So in order to determine whether or not the war is profitable for the victor an estimation of the war losses is needed.

It is assumed that the total losses in the war, i. e. the sum of the losses of the two nations, is a function of the arms levels x and y. The losses can therefore be written f(x, y). Regarding this loss function, the following is assumed. It is defined at all points (x, y) such that $x, y \geq 0$. It is differentiable. It increases monotonously along the line $x = y$. Furthermore it is assumed that

(2.1) $f(x, y) = f(y, x)$,
(2.2) $f(x, 0) = 0$, and
(2.3) $\partial f / \partial x > 0$, when $x < y$.

Formula (2.1) is an assumption about symmetry. (2.2) in combination with (2.1) says that if the arms level of one of the nations is 0 there are no losses in a war between them. Formula (2.3) states in combination with (2.1) that if the nation with the lowest arms level increases its level while the arms level of the other nation is kept constant the losses in an eventual war will increase.

These assumptions seem fairly reasonable. It seems more problematic, however, to determine what happens to the total losses, when the nation with the highest arms level increases its level while that of the other nation is kept constant. Here one could argue in two ways. Either one could say that because the total amount of weapons used in the war increases, the losses should also increase. Or, to argue the other way, because the nation with the highest arms level becomes still more superior the war will be of shorter duration. The losses should consequently decrease. The two arguments give the following alternative assumptions concerning the loss function.

(2.4 a) $\partial f / \partial x > 0$, when $x > y$, or
(2.4 b) $\partial f / \partial x < 0$, when $x > y$.

Now, the loss function f(x, y) gives certain information about the size of the losses in a war between the two nations. But to determine whether a war is profitable for the winner or not, a comparison of its evaluation of the losses and of the war trophy is required.

Consider a point in the sector $x > y$, i. e. the sector where A wins a war. Assume that A is considering an attack on B and therefore tries to evaluate the war. It is assumed that in making this evaluation A compares the status quo, i. e. the state of the two-nation world before an attack by A on B, to the postwar state resulting from A's attack on and war with B. This post-war state is characterized by the fact that A can consider B a revenue and f(x, y) determined by the arms levels x and y a cost.[2] A's comparison of the two states will give an order of preference of these states. If the post-war state is preferred to status quo the war is profitable, otherwise not. Assume that when the comparison is made B's arms level is kept constant at, say, y_0. A might then find that for a certain arms level x_1 it prefers status quo to the post-war state, while for another arms level, say x_2, it prefers the post-war state to the status quo. Here x_1 and x_2 are both greater than y_0, but nothing is assumed about their relative order.

The following two assumptions, which are relevant when the nations consider whether to prefer the status quo or the post-war state, also are made.

(i) If a nation's preference-order of the two states is reversed at two different points in the x-y-plane, then there is a point between these two points at which the nation is indifferent toward the two states.

(ii) If a nation at a certain point in the x-y-plane is indifferent toward the two states, then the nation is also indifferent toward the states at all other points at which the losses are of the same size (that is, at all points at which the value of $f(x, y)$ is the same).

In the situation described, where A compares the two states, the first assumption implies that there exists a x_0 such that $x_1 > x_0 > x_2$, or such that $x_2 > x_0 > x_1$, and furthermore such that A is indifferent toward the two states at the point (x_0, y_0). The second assumption implies that A is indifferent toward these states at all points on the line $f(x, y) = f(x_0, y_0) = K_B$, where K_B is constant. This line is called A's indifference line. Because the constant K_B tells how great the losses are that A is prepared to suffer in order to conquer B, it is possible to consider K_B a measure of the value of B to A.

Now it is also convenient to define two possible military policy goals for a nation.

Military policy goal 1. A nation has this goal if it tries to get into such a position that it is profitable for it to attack the other nation. This goal could be considered aggressive, and a nation with such a goal is also called aggressive.

Military policy goal 2. A nation has this goal if it tries to get into such a position that it is not profitable for the other nation to attack it. This goal is defensive and a nation with such a goal is called defensive.

In order to get a clear view of the conditions under which the different goals are realized for the two nations it is convenient to draw the indifference lines of A and B in the x-y-plane (fig. 1). Only the first quadrant where $f(x, y)$ is defined is considered. In the sector $x > y$ the nation A wins a war. Within this sector the line

(2.5) $f(x, y) = K_B$

is relevant, because on the one side of this line it is profitable for A to attack, while on the other side the opposite is true. Correspondingly, in the sector $y > x$, the relevant line is

(2.6) $f(x, y) = K_A$

The slope of these two lines is obtained by derivation. Derivation of (2.5) gives

$dy/dx = -(\partial f/\partial x)/(\partial f/\partial y)$

According to (2.3) the derivative $\partial f/\partial y > 0$ when $x > y$. The sign of $\partial f/\partial x$ depends on which of the alternative assumptions (2.4 a) and (2.4 b) is considered valid. Thus dy/dx is negative when (2.4 a) holds while (2.4 b) would make dy/dx positive. The same reasoning holds for the line (2.6). In fig. 1 therefore the two lines have the slope which follows from assumption (2.4 b). If it is assumed that K_A and K_B are both positive, i. e. that each nation represents a certain positive value to the other, it is also clear that the two indifference lines will cut the line $x = y$ outside the origin, as in fig. 1.[3]

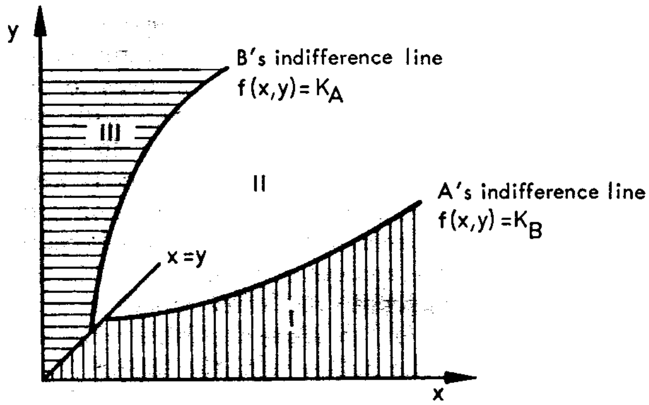

Fig. 1

The three areas I, II and III in the figure now have the following meaning with respect to the previously defined military policy goals.

Area I: Within this area military policy goal 1 is realized for A and no goal is realized for B.

Area II: In this area goal 2 is realized for both nations.

Area III: In this area goal 1 is realized for B and no goal is realized for A.

In connection with this model one can now define the concept of *balance of power*. Balance of power is said to exist between two nations when it is not profitable for either of them to attack the other. This means that in fig. 1 there is a balance of power between A and B in area II.[4]

It is now easily seen how the model illustrated in fig. I can be looked upon as a model for arms races. If both nations are aggressive A will try to reach area I by increasing the arms level and B will try to reach area III in the same way. Thus an arms race will result. On the other hand, if one of the nations, say A, is aggressive while the other is defensive there will be an arms race along the line (2.5) because A is striving to reach area I and B area II. It is also possible to elaborate the model in such a way that there will be an arms race even when both nations are defensive. This particular case, which corresponds to a situation where both nations have uncertain information concerning the opponent's arms level, will be discussed later.

If it is assumed that (2.4 a) is valid instead of (2.4 b) the derivatives of the indifference lines (2.5) and (2.6) will be negative. This situation, illustrated in fig. 2, obviously is different with respect to the nations' propensities to arm. For instance, a nation will not by an increase of arms level reach a point where an attack on the opponent is profitable. On the contrary such a point would be reached by disarming. Perhaps this somewhat peculiar fact indicates that (2.4 b) is a more reasonable assumption than (2.4 a). In any event the different consequences of (2.4 a) and (2.4 b) make further investigation of the loss function important. In this essay, however, (2.4 b) will henceforth be considered valid.

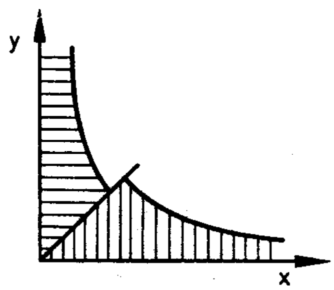

Fig 2

A problem related to what has been said so far affects the distinction between aggressive and defensive nations. Hitherto it has only been said that the distinction should be made but nothing much has been said about how an aggressive and a defensive nation should be defined respectively. Two possible definitions are suggested here.

Definition 1. A defensive nation is characterized by the fact that there are no arms levels (i. e. no points in the x-y-plane) at which it prefers the postwar state to the status quo. This means that the defensive nation contrary to the aggressive one, lacks an indifference line.

Definition 2. The defensive nation, like the aggressive one, has an indifference line. They are distinguished, however, by the fact that the defensive nation prefers the points in the x-y-plane where it prefers the status quo to the post-war state, whereas the opposite is true for the aggressive nation. From the point of view of behavior this definition implies that a defensive nation, if it finds itself at a point where it prefers the post-war state to status quo, it will disarm rather than attack.

For the main part of this essay it is not necessary to decide between these two definitions, since the discussion is not carried so far as to be affected by such a decision.[5] Besides, it is easily seen that the model might be developed in such a way that the two definitions are used simultaneously. Apart from the possibility of distinguishing between aggressive and defensive nations there is also the possibility of using the two definitions for distinguishing between three groups of nations: nations without an indifference line and among those with an indifference line, nations which prefer to be on the one side of it (area II) and those which prefer to be on the other side (area I or III). It might be interesting to investigate whether this classification of nations with respect to goals could be fruitfully applied to real nations.

It should be of great interest to the nations in the model to be able to decide whether the opponent is aggressive or defensive. Then, the simplest thing for the nations to do would be to apply the definitions of "aggressive" and "defensive" utilized in the model. However, one can easily imagine that the model is constructed in such a way that this simple application of the definitions is not possible. For instance this might be the case if the nations have limited information about each other. In this situation the nations might be forced to determine the opponent's goal by a more indirect method. For example, changes in the opponent's arms level could be looked upon as indicating an aggressive goal or a defensive goal. Conversely, one could imagine that the nations deliberately change their arms levels in order to communicate to the opponent their military policy goal. These problems will be discussed later.

When compared to reality it is immediately seen that the model presented so far suffers from serious oversimplifications. Some of these are listed here. Later in the essay it is shown how some of them could be eliminated by introducing further assumptions.

(i) It is a simplification that attack and defense should require the same amount of armaments, i. e. that the nation with the higher arms level wins a war independent of whether it is the attacker or the defender.

(ii) The important problem of what sort of information the nations have about each other, and how this affects their behavior, is not explicitly treated.

(iii) A further simplification is that the only type of war conceivable is war leading to unconditional surrender. Therefore the possibility of limited war ought to be introduced.

These three simplifications are discussed in the next pages. There are, however, additional simplifications, unfortunately not dealt with in this essay, among which are:

(iv) There are only two nations. A more realistic model should allow for several nations.

(v) Costs of armaments are not considered although it is obvious that in reality these costs are of great importance to a nation when determining its arms level.

(vi) No distinction is made between different kinds of weapons. It appears feasible, however, to give the nations the capacity to change the very shape of the loss function by introducing a distinction between offensive and defensive weapons.

3. Different weapons' effect in attack and defense

It is often the case that the amount of weapons required for defending a certain position differs considerably from the amount required for attacking and taking the same position. In military operational planning it is, for example, often estimated that in ground combat the attacker needs a superiority of 3:1 in relation to the defender. If this idea is applied at the national level, the model presented so far could be criticized on the ground that it is assumed that the nation with the highest arms level wins a war regardless of whether or not it is attacking. By changing this assumption the model could be made more realistic.

It is therefore assumed that the ratio between the attacker's and the defender's arms levels determines which nation wins a war. If this ratio exceeds a certain minimum value, which is a parameter in the model, the attacking nation wins, otherwise the defender wins. This parameter is called k. If $k > 1$ attack is said to be more difficult than defense and if $k < 1$ attack is easier than defense. Thus, the model already presented is the special case where $k = 1$. In the following the balance situation in the model is examined when $k > 1$ and when $k < 1$.

Attack more difficult than defense (k > 1)

In the x-y-plane is drawn the line $y = kx$ and its reflection in the line $x = y$ (fig. 3).These two lines divide the first quadrant in three sectors S_1, S_2 and S_3. In S_1 A wins a war independent of who attacks, and the same is true for B in S_3. In S_2 neither nation wins a war in which it is the attacker.

In order to know the positions of the indifference lines it is also necessary to describe the loss functions. As a matter of fact, two loss functions are utilized here: one which applies when A attacks B and another when B attacks A. Each of these functions has properties analogous to the properties of f(x, y). This means that for the function which applies when A attacks B and which is called $f_A(x, y)$ the derivative $\partial f_A/\partial x$ is less than 0 at points where $x/y > k$ (i. e. in the sector S_1) and greater than 0 where $x/y < k$ (i. e. in S_2 and S_3). Except for the assumption about symmetry (2.1) all other assumptions are the same as for f(x, y). The correspondence to the assumption of symmetry is that the function $f_B(x, y)$, i. e. the function which applies when B attacks A, is defined $f_B(x, y) = f_A(y, x)$.

Now it is possible to draw the indifference lines as shown in fig. 3. In area I it is profitable for A to attack and in area III it is profitable for B.

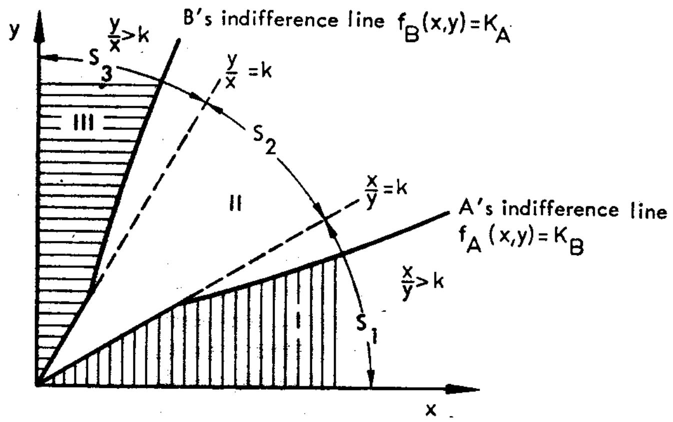

Fig. 3

The most striking feature of this model as compared to the earlier model is that the area in which there is a balance of power, that is area II, reaches the origin. Hence in this model very low arms levels do not necessarily mean imbalance.

Attack easier than defense (k < 1)

In this case the line y = kx and its reflection in y = x divides the quadrant in the sectors S_1, S_2 and S_3 with the following properties (fig. 4). In S_1 nation A wins irrespective of who attacks whereas in S_3, this is true for B. In S_2 the attacking nation, which might be either A or B, wins.

Too, in this case the indifference lines give three areas of interest with respect to profitable attacks. In area I it is profitable for A to attack, in area III it is profitable for B and in area II, where there is balance, it is profitable for no nation to attack. It is of interest, however, that in this model the areas I and III partially overlap. Consequently, there is in the neighborhood of origin an area in the figure (the lines cross) in which it is profitable for A as well as for B to attack. Therefore there is in this area another type of imbalance, different from the imbalance in the remainder of areas I and III. This new type of imbalance is quite similar to the imbalance which exists between two nations, when both have a good first strike capability and lack second strike capability.

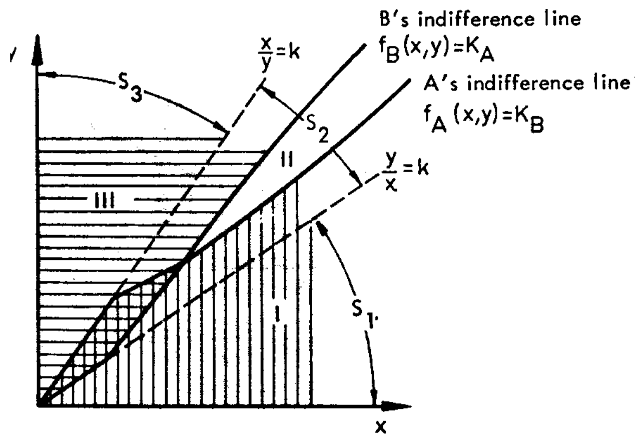

Fig. 4

4. The nations' information about their situation

The problem concerning what sort of information the nations have about their situation, and how their behavior is affected thereby, has not been treated so far. Here are listed some ways in which the nations may lack such information.

(i) they may be uncertain about the shape of the loss function.

(ii) they may be uncertain about the position of the opponent's indifference line.

(iii) they may be uncertain about the opponent's arms level.

(iv) they may be uncertain about whether the opponent is aggressive or defensive.

Here will be discussed the situation where both nations have uncertainties of types (iii) and (iv) but not of types (i) and (ii). Furthermore it is assumed that the construction of the model is such that "defensive" and "aggressive nation" are defined by definition 2 above. Because both nations are uncertain about the opponent's goal they have to infer this by finding out whether the opponent, when he changes his arms level, strives to get into area II or the one of areas I and III which is relevant to him. But such observations regarding changes in the opponent's arms level are difficult to make because it is also assumed that the nations do not have complete information about this very arms level.

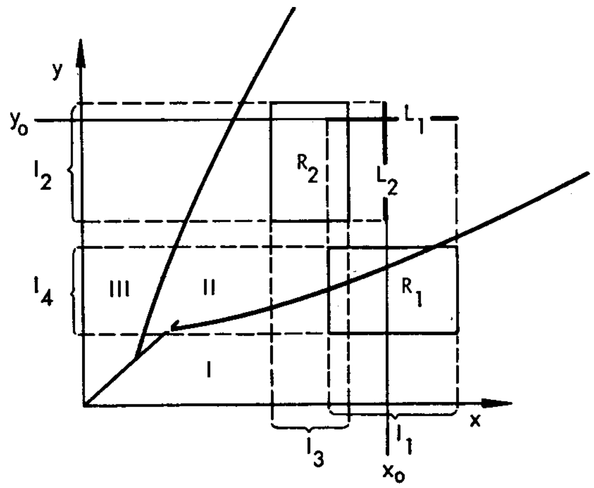

Fig. 5

x_0 = A's actual arms level
y_0 = B's actual arms level
I_1 = the interval within which B thinks A's arms level lies
I_2 = the interval within which A thinks B's arms level lies
L_1 = the line on which B thinks (x_0, y_0) lies
L_2 = the line on which A thinks (x_0, y_0) lies
I_3 = the interval within which A thinks that B thinks x_0 lies
I_4 = the interval within which B thinks that A thinks y_0 lies
R_1 = the rectangle within which B thinks that A thinks (x_0, y_0) lies
R_2 = the rectangle within which A thinks that B thinks (x_0, y_0) lies

The situation described is illustrated in fig. 5, where x_0 and y_0 signify A's and B's actual arms levels respectively. It is assumed that the nations' uncertainty about the opponent's arms level is of the character that they themselves are sure they know in which interval the arms level lies but admit uncertainty as to where in the interval. In the figure, I_1 and I_2 signify these intervals, i. e. the intervals in which B and A respectively think the opponent's arms level lies. Because both nations know their own arms level, A thinks (x_0, y_0) lies somewhere on L_2 and B thinks that that point lies somewhere on L_1. In order to infer the opponent's goal both nations now have to find out towards which area (I, II or III) the opponent is striving. From A's point of view the problem looks like this. In order to decide towards which area B tries to move (x_0, y_0) nation A must first determine which opinion B has about the position of the point. At first, since B knows its own arms level, it is evident that A's opinion about B's opinion about y_0 is the same as A's opinion about y_0. Then, when A's opinion about B's opinion about x_0 is considered, it is assumed that A's uncertainty is of the character that he thinks that B thinks that x_0 lies within a certain interval, which here is called I_3. To summarize, this means that A thinks that B thinks that (x_0, y_0) is somewhere in the rectangle R_2. Correspondingly B thinks that A thinks that the same point lies somewhere in the rectangle R_l.

The following example illustrates how the nations' behavior is affected by the uncertainties described. Assume that R_l completely, or mainly, lies within area I (or that R_l is moving towards this area if a certain interval of time is considered). B will then reasonably infer that A has aggressive intentions and will, on the basis of the information B has, expect an attack from A at any time. The result of such a war depends on where (x_0, y_0) lies. If the situation is the one illustrated in the figure it is clear that B cannot, regardless of whether B is aggressive or defensive, welcome an attack from A because (x_0, y_0) is in the area where war is non-profitable for both nations. B therefore will try to move R_l into area II. One conceivable way for B to do that is by increasing his own arms level since that could move the interval I_4 upwards. But such an arms level increase by B could also move R_2 towards III, which in turn would stimulate A to increase his arms level. It is interesting to note that the described process in its entirety is compatible with both nations having defensive military policy goals.

The conclusions in the last paragraph are vague and tentative. This is due to the fact that no explicit assumptions are made about which factors determine the two nations' opinions about each other in different respects. However, it seems feasible to build a more complex version of the model in which such explicit assumptions are included.

5. Limited wars

In the previous models all wars that could take place were wars leading to unconditional surrender; this surrender occurred when all weapons of the losing nation were lost or destroyed. The purpose of the following model is to eliminate this simplification. Thus in this model limited wars are permitted. Here a war is called limited when the adversaries neither try to force each other to surrender unconditionally nor consume all their weapons in the war. This last fact means that both parties are perfectly capable of continuing the war, but for some reason they choose not to do so. The purpose of this model, among other things, is to find an explanation for why this is so. That is, why should a nation consider it profitable to start a war, but consider it unprofitable to continue the war until the opponent surrenders unconditionally? Could such behavior be the result of rational calculations as was the behavior described in the previous models? It is possible that such behavior could result from nations not having correct information about each other's arms level. In this case a war could be started, but then when, as a result of combat experience, the nations acquire better information about each other's arms level, they might both prefer to end the war. In this model, however, it will be shown how a limited war can result from rational behavior without assuming false information.

Here it is assumed that each nation, A and B, has two kinds of weapons at its disposal: weapons designed for total war and weapons designed for limited war. This distinction is inspired by Kissinger, who writes: "... two basic organizations would be created: the Strategic Force and the Tactical Force. The Strategic Force would be the units required for all-out war. ... The Tactical Force would be the Army, Air Force and Navy units required for limited war."[6] The distinction between these two kinds of weapons may be such that nuclear weapons are present only among the weapons for total war. That is the case in Kissinger's later book *The Necessity for Choice* but not in the earlier *Nuclear Weapons and Foreign Policy.* The arms levels of the two nations with respect to armaments for limited war are denoted x and y respectively. Correspondingly the arms levels with respect to armaments for total war are denoted X and Y.

Between the two nations of the model limited wars as well as total wars can be fought. The assumptions concerning the course of both types of war are similar to the corresponding assumptions in the previous models.

In a limited war the amount of weapons introduced by A and B are x and y respectively. If $x > y$ nation A wins the limited war, whereas B wins if $y > x$. Victory in a limited war means that those forces of the one nation designed for limited war beat or destroy all of the opponent's corresponding forces. In this situation the winning nation would have complete control over both nations had there not been any weapons designed for total war.

Now it is assumed that it is possible for the victor in a limited war to occupy or take only part of the losing nation's territory. How great a part he will take the victor decides after the following consideration. If it is assumed that A wins and that the part of B that A takes is called b, then b has to be great enough compared to A's war losses and at the same time so small that it is not profitable for B to start a total war. It is therefore only in cases where such a b exists that A can profitably conduct a limited war for a limited purpose. The conditions for the existence of such a b will be discussed later.

The losses that A suffers in the limited war will be denoted $h_A(x, y)$. This function is defined in the sector $x > y$ of the first quadrant and has properties corresponding to (2.2), (2.3) and (2.4 b) in the original model. If B wins a limited war its losses amount to $h_B(x, y)$. This function is defined in the sector $y > x$ and has properties equivalent to those of $h_A(x, y)$. The assumptions concerning $h_A(x, y)$ and $h_B(x, y)$ consequently imply that the nation which wins a limited war can decrease its losses in such a war by increasing its arms level while the losing nation, by increasing its arms level, will increase the victor's losses.

In a total war the weapons introduced are X and Y respectively. Here the course of the war is completely analogous to that of a war in the original model. Who wins will depend on which is the greatest of X and Y. The total losses are determined by the function $g(X, Y)$ which, with respect to derivatives and zero-values, has properties corresponding to those of $f(x, y)$. The correspondence to (2.4 b) is assumed to be valid.

With respect to the relations between the arms levels of A and B two cases may be distinguished, which will be discussed in the following. In case 1 $x > y$ and $Y > X$. In case 2 $x > y$ and $X > Y$. (The two remaining cases where $y > x$ and $X > Y$ and where $y > x$ and $Y > X$ imply, as will be evident, nothing new compared to the cases 1 and 2.)

Case 1

It is first seen that in this case there are three possible final states in the model.

The first of these is the state that will result if there is no war at all, i. e. status quo.

The second final state will be the result of a limited war. In this state A has won the war and suffered some losses, namely $h_A(x, y)$. Furthermore A has taken the territory b from B. In this state therefore B will continue to exist as a nation, although it has lost b. Naturally also B has suffered some other war losses, but because of the nature of the conclusions drawn from the model it is possible to disregard these losses.

The third final state will result from a total war. In this state B has won the war and taken A but also suffered the losses $g(X, Y)$. A has ceased to exist as a nation.

These three final states are called T_l, T_2 and T_3. Now it is of interest to know the order of preference of these states for A and B respectively. Suppose, for example, that A prefers T_2 to T_l and that B prefers T_2 to T_3. In that situation A would start a limited war in order to take b. But if instead B prefers T_3 to T_2 it would no longer be profitable for A to start a limited war because B would then respond by initiating total war. Consequently, it is important to see what determines and affects the nation's orders of preference.

Concerning A it is obvious that T_3 is ranked last in every order of preference. The relative order between T_l and T_2 will depend partly on the size of $h_A(x, y)$ and partly on the size of b.

Now assume that in a certain case b has the value b_l. Assume furthermore that there are two points (x_1, y_0) and (x_2, y_0) such that $x_2 > x_1 > y_0$ and such that in the first point A prefers T_1 to T_2 whereas in the other it prefers T_2 to T_1. Then there is between x_1 and x_2 an x-value x_i such that A is indifferent to the two states in (x_i, y_0). Therefore A's indifference line is the line $h_A(x, y) = h_A(x_i, y_0) = C$, which is drawn in fig. 6.

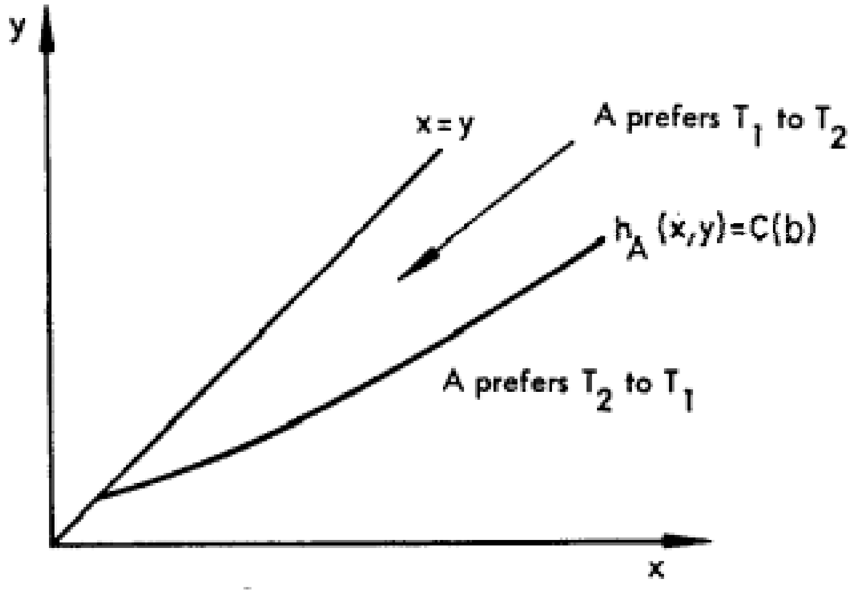

Fig. 6

Assume that instead of b_1 a value b_2 such that $b_2 > b_1$ had been considered. The indifference line in the figure would then move upwards, because the larger territory b_2 would be worth greater losses than the territory b_1. Therefore the constant C can be written as the monotonously increasing function C(b). Hence the complete equation of the indifference line will be

(5.1) $h_A(x, y) = C(b)$

Thus if x and y are given (5.1) will give a value of b for which A is indifferent between T_1 and T_2. In order to avoid confusion, the variable of (5.1) is therefore written b_A. And the equation is

(5.2) $h_A(x, y) = C(b_A)$ where $dC/db_A > 0$.

In considering B's orders preference of it is obvious that T_1 is always preferred to T_2. What is interesting therefore is to find the conditions under which T_3 is preferred to T_1 and, when this is not the case, the conditions under which T_3 is preferred to T_2.

If the relative order between T_1 and T_3 is considered first an indifference line can be drawn exactly analogous to the line in the original model (fig. 7). The equation of the line is $g(X, Y) = L$, where L is a constant.

In the area where B prefers T_3 to T_1 the state T_3 is also preferred to T_2. But in the area where T_1 is preferred to T_3 the order of T_3 and T_2 depends on the arms levels and the size of the area b, which in state T_2 is taken by A.

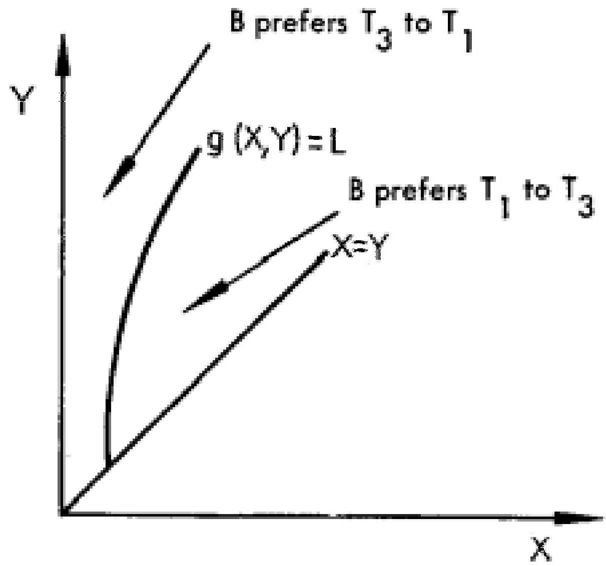

Fig. 7

Assume, in the same way as when A's orders of preference were considered, that in a certain case b has the value b_l. Assume furthermore that there are two points (X_0, Y_1) and (X_0, Y_2) such that $Y_2 > Y_1 > X_0$ and such that in the former B prefers T_2 to T_3 and in the latter it prefers T_3 to T_2. Then there is also a point (X_0, Y_i) where $Y_1 < Y_i < Y_2$ and at which B is indifferent to T_2 and T_3. The equation of the indifference line will be $g(X, Y) = g(X_0, Y_i) = K(b)$. Now, for a territory b_2 greater than b_1 it is clear that B is prepared to suffer heavier losses in order not to lose the territory. Thus the function K(b) is monotonously increasing. With arms levels given, the equation $g(X, Y) = K(b)$ will then determine the value of b which makes B indifferent to T_2 and T_3. The variable b is therefore written b_B and the equation of the indifference line is

(5.3) $g(X, Y) = K(b_B)$, where $dK/db_B > 0$.

Now assume that the four arms levels, that is x, y, X and Y, are given. (5.2) and (5.3) then determine b_A and b_B respectively. If, for these, the inequality

(5.4) $b_A < b_B$

holds, it is profitable for A to start a limited war against B. This is so because, for an arbitrary b in the interval between b_A and b_B, the state T_2 is preferred to T_1 by A and simultaneously T_2 is preferred to T_3 by B. For B it is therefore important to realize the inequality

(5.5) $b_A > b_B$

If A is defensive it is also satisfied with this. But if A is aggressive it will try to realize (5.4). How A and B can try to affect the relative order of b_A and b_B is seen from the signs of the following derivatives, which are obtained from deriving (5.2) and (5.3).

(i) $\partial b_A / \partial x = (\partial h_A / \partial x) / (dC / db_A) < 0$,

(ii) $\partial b_A / \partial y = (\partial h_A / \partial y) / (dC / db_A) > 0$,

(iii) $\partial b_B / \partial X = (\partial g / \partial X) / (dK / db_B) > 0$,

(iv) $\partial b_B / \partial Y = (\partial g / \partial Y) / (dK / db_B) < 0$.

Thus, A can try to achieve (5.4) by increasing x, which will decrease b_A, and/or by increasing X, which increases b_B. Nation B, on the other hand, can try to achieve (5.5) by increasing y, which will increase b_A, and/or by increasing Y, which decreases b_B. Increasing the amount of weapons designed for limited war and increasing those designed for total war could therefore be equivalent measures when a nation aims at realizing a certain military policy.

Here some strategic concepts may conveniently be discussed. Assume that (5.5) holds and that consequently it is not profitable for A to start a limited war. This means that B, although it is inferior with respect to weapons for limited war, is able to deter A from starting such a war. This kind of deterrence will be called *type II deterrence* with reference to Herman Kahn's similar concept with the same name.[7]

Because there will be no war, neither limited nor total, in the situation described, it represents a kind of balance. Now, one can imagine that one or both nations are badly informed about their situation. For instance, assume that A, due to ignorance concerning B's arms level or concerning B's evaluations, falsely believes that (5.4) holds. A will then start a limited war against B. But because (5.5) in fact holds, B will respond by starting a total war against A. Here A's starting of a limited war can be looked upon as a disturbance in the original balance situation and because this disturbance results in a total war, one can say that in the original situation there is an *unstable balance of power*. The course of the war, starting with a disturbance and developing into total war, can be considered a case of *escalation* or *eruption.*

In this model the policy known as "massive retaliation" can be seen as a special case of type II deterrence. It is already assumed that x > y and Y > X. If (5.5) is true and consequently A is deterred from starting a limited war by type II deterrence, one can say that B has conducted a massive retaliation policy, if it is also true that y and x are both very small. If A represents the Soviet Union and B the United States, the model could be compared to the world situation in the fifties and the early sixties. In the middle of the fifties, when X was very small, the policy of the United States was massive retaliation. Then X gradually increased. One result has been the claim for increased forces for limited war in the United States, i. e. a claim for greater y. In the model such an increase in y might keep (5.5) true despite the increase in X. This comparison between the model and reality must not, of course, be taken too literally.

Case 2

In this case x > y and X > Y. Obviously no war is profitable for B. For A a limited war or a total war, or both, may be profitable. It is clear, however, that A wins a limited as well as a total war. Consequently B is deterred from starting a limited war as A was in the previous discussion about type II deterrence. The "structure" of the deterrence is different in the two cases, however. This is so because in the previous case the deterred nation, i. e. A, would win a limited war, whereas in this case the deterred nation, i. e. B, will lose an eventual limited war. In this case it is convenient, again with reference to Herman Kahn, to speak about *type III deterrence*.[8]

This section will now be concluded with a discussion of two problems related to the last model. With reference to case 1 one can say that these problems concern the conditions under which B will increase the level of conflict to a total war. Hitherto it has been assumed that this will happen only if B prefers T_3 to T_2. One difficulty with this assumption, and this is the first problem, becomes obvious if one imagines that several consecutive attacks are possible. In this case A can start by taking a territory from B which is so small, that according to the assumptions made, it is not profitable for B to start a total war. Then A could take a new piece of territory of a similar size from B, and so on. The territories taken by A in this manner could finally add up to a size which would be more than enough to make it profitable for B to start a total war, had all the pieces of territory been taken simultaneously. From that point of view it might have been rational behavior on the part of B to respond to A's first small attack by initiating total war. Thus, the assumption that B starts a total war only when it prefers T_3 to T_2, is by no means self-evident.

In spite of this objection it seems, however, that in reality nations often do hesitate to increase the war level in the same way as B, which according to the assumption raises the level of conflict to a total war only when it prefers T_3 to T_2. Therefore it could also be advantageous to make a great conquest in several consecutive steps and thereby succeed without provoking the opponent to retaliate. This has happened in reality.

The second problem is this one. It is conceivable that B, even if it does not prefer T_3 to T_2, starts a total war in order to take its revenge for A's limited attack. If B does this, it could be seen as the fulfillment of the following kind of deterrence policy. B tries to deter A from limited attack by making A believe that B's response will be total war and by utilizing the fact that A prefers T_1 to T_3. Here B hopes to be successful in deterring A without taking into account the relative order of T_2 and T_3 in B's own order of preference. With reference to the US strategic debate, this policy can be considered a version of the *minimum deterrence* policy.[9]

Notes

[1]Nothing is assumed about what happens if the two nations have equal arms levels. Such an assumption would most likely only complicate the model without adding anything of interest.

[2]The reason for considering the total losses in the war, i.e. f(x, y), a cost for A alone is the following: as a result of the war B becomes A's war-trophy, and all damage done to B can therefore be considered a cost to A.

[3]This implies that disarming below a certain level is not possible if both nations want to realize the defensive military policy goal.

[4]This definition of balance of power has the reasonable consequence that, other things being equal, the more valuable a nation is to the opponent the higher arms level it needs.

[5]This is not true for the section "The nations' information about their situation".

[6]Henry A. Kissinger, *Nuclear Weapons and Foreign Policy,* 1957, p. 419. The quotation is made only to suggest that the distinction itself is reasonable. The

assumptions made about the effects of the two types of weapons are not from Kissinger.

[7]See the definition of "Type II Deterrence" in Herman Kahn, *On Thermonuclear War,* Princeton University Press, 1961, p. 126. Also see Herman Kahn, *Thinking about the Unthinkable,* Weidenfeld and Nicolson, 1962. On page 108 is written: "U.S. military policy currently seeks to achieve at least six broad strategic objectives:
1 Type I Deterrence - to deter a large attack on the military forces, population, or wealth of the United States, by threatening a high level of damage to the attacker in retaliation;
2 Type II Deterrence - to deter extremely provocative actions short of large attack on the U.S. (for example, a nuclear or even all-out conventional strike against Western Europe) by the threat of an all-out U.S. nuclear reprisal against the Soviet Union;" and so on.

On page 112 Kahn continues: "At some future date, the non-nuclear capability of NATO may be sufficient to repel a large but conventional attack on Western Europe. Until this time Western Europe will probably depend, at least in part, on Type II Deterrence (or Controlled Reprisal) to deter such attacks."

[8]See On *Thermonuclear War*, p. 126. This deterrence is also called, by *Kahn and others,* "graduated deterrence". See, for example, Sir Anthony W. Buzzard, "Massive Retaliation and Graduated Deterrence", *World Politics,* 1956, p. 229. Sir Buzzard writes when pleading for a graduated deterrence: "We should not cause, or threaten to cause, more destruction than is necessary. By this criterion, all our fighting should be limited (in weapons, targets, area and time) to the minimum force necessary to deter and *repel* aggression, prevent any unnecessary extension of the conflict, and permit a return to negotiation at the earliest opportunity - *without seeking total victory or unconditional surrender."* (italics mine).

[9]See, for example, *On Thermonuclear War,* p. 7 ff.