V.I.Danilov, A.I.Sotskov

Social Choice Mechanisms

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo
Hong Kong Barcelona
Budapest

Table of Contents

Introduction

The theory of social choice deals with both the processes and results of collective decision making. In this book, we explore some issues in the theory of social choice and mechanism design. We examine the premises of this theory, the axiomatic approach, and the mechanism design approach.

The main questions are what is collective interest, how is it related to individuals' interests, how should one design social interactions, laws, and institutions? These questions are not new. Philosophers, social scientists have indeed pondered upon them for years. And, in fact, the organizational structures of many social institutions - courts, parliaments, committees and regulatory boards - often lack a sound theoretical base. This is not surprising, as it is, indeed, difficult to provide for a comprehensive formalization of the activities of such organizations. Nevertheless, there has been a definite trend towards providing clear and unambiguous rules for collective decision making. These very rules constitute the body of social choice theory and its main object.

The basic problem of social choice

We explain here more precisely what a problem of social choice is, what approaches might be used to tackle it, and what kind of solutions it leads to. We introduce a few basic notions in preliminarily fashion and, in doing so, we stress both motivations and explanations.

To start with, we consider a group N of persons (people or organizations), called agents or participants. A social choice problem arises when this group is summoned to make a choice among one or several objects (projects, plans, candidates, etc.). We shall call these objects 'alternatives' and shall assume they form a set A. In practical set-ups, it is often difficult to describe the set of alternatives precisely, a priori, but in formal set-ups we shall always consider that the set A is both given and clearly outlined. Sometimes it might turn out that it is more interesting and crucial to analyze how the set A forms than to investigate the choice itself. In particular, the addition of new compromise alternatives to the set A could be a crucial issue, as it might be the only way to yield a satisfactory solution to a social choice issue. However, we put this issue aside (since we are as yet unable to describe appropriately the creative process by which alternatives form) and consider, in the sequel, the set A to be a given, fixed, abstract set. We therefore focus on the formal choice from a given set A.

In what follows, we restrict ourselves to choices of a single element from the set A. This is not as major a restriction as it may appear. First, it is perfectly possible without any special modification to consider the more general problem of choosing a subset X from the set A. Second, the set A could be substituted by the set 2^A of all its subsets (or some feasible part of 2^A). And we deal within this set-up with a single-valued choice again. Of course, there will be subtleties here. However we wish to keep matters simple for the time being. We will consider from now on that a problem of choice is solved whenever a single element from A is determined.

Clearly before any decision making takes place, one should make sure that all participants be familiar with all the alternatives and understand them in order to evaluate them. This stage is crucial, but we shall not examine it here. Instead we shall assume that this stage is over and that every participant has a clear preferences on the alternatives at stake. It is immaterial to the issue whether a participant bases his evaluations on subjective considerations or objective characteristics, whether the agent behaves egoistically or altruistically. The preferences elaborated by agents are modelled as binary relations, in fact as linear orders (see details in 1.1).

Now we come to the issue of choice. Our group comes up with the choice of an alternative, based on both the individuals' preferences and on some rule. In this whole process, the participants' preferences are the inputs to the choice problem, while the chosen alternative is the output; the choice itself is a kind of black box, which we are going to examine closely.

Simply put, one can imagine social choice as follows. Participants state their preferences on voting papers and then send these papers to a processing center. The center then processes the information. A bundle of voting papers of the agents is called a *preference profile* and is denoted by R_N. An alternative a, selected at this profile R_N, is denoted by $f(R_N)$. The point is that we do not know the profile R_N beforehand. Therefore, the processing center should select $f(R_N)$ for a sufficiently wide domain of preference profiles. The center's activity is described by social choice function (SCF), that is a mapping

$$f : \{\text{preference profiles}\} \to A.$$

The aim of the theory of social choice is both to investigate concrete SCFs and to construct new SCFs, which exhibit specific and desirable properties. Going a step further, we might address the issue of choice between different SCFs. We might wonder whether the central authority has the power to enforce a selected alternative. Will the participants agree on the alternative proposed by the center? This will depend on the center's authority, which in turn depends on parameters such as power, honor, intelligence or skill.

We now make two points. First, choice depends on a preference profile. However, it may depend on many other circumstances and factors, such as the experience of agents, their knowledge, their ability to agree, to cheat, and so on. These factors are not always explicitly fixed. Therefore a rule may lead to different outcomes for the same preference profile. And in fact the most important source of multi-value outcomes is the multiplicity of equilibria in the game generated by the social choice mechanism (see 1.3 and 1.4). Therefore, social choice theory deals with multi-valued social choice correspondences (SCCs)

$$F : \{\text{preference profiles}\} \to 2^A,$$

as well as with SCFs.

An SCC, F, assigns a subset $F(R_N)$ in A to a preference profile R_N, where $F(R_N)$ is usually non-empty. The interpretation of an SCC is already less clear than that of an SCF. One can understand that every element from $F(R_N)$ can be chosen for the profile R_N. Of course, only those SCCs, F, whose associated sets $F(R_N)$ are not too large will be of interest.

Second, a social choice function (or correspondence) may turn out to be undefined for certain "not too nice" profiles R_N. This might definitely turn out to be a drawback of a particular rule. We might not be too happy with a rule that fails to do its job for certain profiles. However a rule that works well on a restricted set of preferences might do. We discuss this further in 1.4 and Chapter 3.

Axiomatic approach

The main object of social choice theory is thus a SCF (or at least a SCC whose set of outcomes is small enough) determined on the set of all (or maybe sufficiently wide a subset) preference profiles. However, there is a great number of such SCFs, although most of them will turn out to be rather absurd or of little or no interest from the viewpoint of social choice. The theory of social choice began with the study of concrete SCFs, based on "reasonable" assumptions and desiderata (majority rule, Borda's rule, Copeland's rule, and so on). As it turns out, every rule has its qualities and drawbacks. And in fact, as soon as a particular drawback is dealt with, others emerge. The thorough comparison of various SCFs has helped define and select an array of desirable properties for SCFs. We discuss in detail in 1.2 and 1.3 some of these properties: Pareto optimality, neutrality with respect to alternatives, monotonicity, some form of democracy (anonymity, absence of dictatorship, veto principle), non-manipulability.

The axiomatic approach to social choice has evolved gradually from the following process. First desirable properties of SCFs are formulated, then the class of SCFs possessing these properties undergoes complete description and thorough study. Arrow's impossibility theorem is a classic example of this approach. It states that there exists no preference aggregating rule (see 1.A1), which satisfies simultaneously several "natural" conditions. Arrow's somewhat negative result stimulated a plethora of studies whose conclusions turned out to be just as depressing: an apparently "innocuous" set of requirements on SCFs yielded either inconsistency or dictatorship, a rather unpleasant feature.

The gloomy character of these studies stimulated the search for an alternative approach to social choice issues. This alternative approach is called mechanism design. The body of this literature was constituted by and large in 1973, although some ideas formed before. Our book is devoted to the study of mechanism design.

The revelation issue

The axiomatic approach explicitly or implicitly assumes that the center knows the preference profile of agents R_N at the time of decision making. In this case, the social choice problem boils down to a multi-criteria optimization problem. However there is a difference between these two problems and in order to clarify this we bring in the issue of revelation. How does the center learn about the profile? In principle, only from the participants' messages. Indeed, participants'preferences are usually personal and intimate. The participant is the only one to know them. Given this, a question arises: will the participants likely communicate their preferences truthfully? This is a crucial question which brings the preference revelation issue to the fore. (The numerous attempts to reconstitute utility functions for a society, a managing group, a collective enterprise, or individual worker, be it through computation methods or direct interrogation, will convince the reader of the relevance of this issue.)

Now, if a participant is really interested in the outcome of a social choice issue and has the power to influence the outcome (if not, why bother to ask), then why should we expect this participant to truthfully reveal his genuine preferences? Whether a participant might want to reveal his true preferences or to misrepresent them by manipulating either his preferences or any other personal information, will essentially depend on the decision rules adopted for the social choice problem at stake. And, in fact, it is quite possible that an a priori "reasonable" Pareto optimal rule (assuming truth-revealing behaviour) yield non-optimal social outcomes, because agents will have distorted their preferences.

In short, agents can, and most probably will, behave strategically. This is precisely what distinguishes social choice theory from multi-criteria optimization theory. And, in fact, a realistic description of the behaviour of participants is an intricate problem which draws both on social choice theory and on game theory. We return to this issue later. Meanwhile, we merely state here that we must add to a social choice rule a certain number of behavioral assumptions. These assumptions specify and formalize our thoughts about the participants' actions and information. The assumption about truthful revelation of preference is the simplest, albeit not the most realistic.

Social choice mechanisms

The strategic approach to social choice, while bringing to the fore quite a few delicate points, opens in turn new opportunities. In the axiomatic approach, preferences perform a double task. On the one hand, they indicate the aims of participants and allow to evaluate the outcome. On the other hand, viewed as messages, they provide the means for attaining these aims. One task task might impede the other. However, agents' messages might not consist only of preferences.

Messages can be anything as long as they yield good outcomes. In fact, many real-life choice procedures are not given as SCFs. Take for example, the two stage election procedure. This electoral procedure works as follows: the two most popular candidates remaining on the electoral register after the first round, are then selected according to a simple majority rule which constitutes the second round. The point is the following: even when preferences serve as messages, these preferences may differ at different stages of the procedure. The social choice mechanism approach decisively separates aims and means.

With this we pinpoint the core notion of this book, namely the notion of a social choice mechanism or, more simply, the notion of mechanism. A mechanism consists in two things: a bundle of strategy sets S_i (one for every agent $i \in N$) and an outcome function, $\pi : \Pi_{i \in N} S_i \to A$. Every agent starts by selecting a message $s_i \in S_i$. Then the outcome function π determines an alternative $a = \pi(s_N)$, where $s_N = (s_i)_{i \in N}$.

A priori, no restrictions are imposed on the structure or size of the sets S_i. However, in practical applications, it is essential that these messages be kept simple, and that the mechanism (i.e., the mapping π) be clearly defined and stated. All participants are supposed to know the mapping π, which means that participants know the effect of their actions on the outcome. Ideally, any formalized procedure of choice should be organized in this fashion. Of course, this ideal set-up seldom obtains. Indeed, real-life procedures generally incorporate both unclear messages and inconclusive outcomes, moreover they may depend on external factors, etc.

The difference between the SCFs, or axiomatic approach to social choice, and the mechanism design approach can be expressed as follows. In the first approach, we know the preference profile R_N and we impose an outcome $f(R_N)$. In the second approach, we are not interested in the preferences of agents; instead we hand them an instrument, namely a mechanism, and tell them: "And now sort it out for yourselves." The SCF approach, as we said, contains an in-built difficulty: the issue of preference revelation. The mechanism design approach, while dispensing with it, formulates it explicitly as the research of an appropriate solution concept of some game.Thus the mechanism design approach circumvents Arrow's seemingly dead-end road by opening new ways to the study of social choice.

The solution concept issue

Now let us discuss the agent's behaviour problem in more detail. Fix a mechanism $\pi : S_N = \Pi_{i\in N} S_i \to A$ and a profile of true preferences $R_N = (R_i)_{i\in N}$. This enables us to build up a game $G(\pi, R_N)$ in which strategies are elements of S_i and situations $s_N \in S_N$ are evaluated with the help of the outcome function π and the preferences R_i. At first glance, it seems that we just replaced a problem of choice from A by a problem of choice from S_N. However, this is not as straightforward as it seems. We now stress two points. This new choice problem is simpler than the former: every agent i receives an unrestricted right to dispose of his set S_i, and to choose individually from it. However every agent's problem becomes more intricate. Indeed, an agent will have a more difficult time evaluating his strategies s_i in this set-up. Of course, he can easily evaluate a bundle s_N with the help of π and R_i. Nevertheless, now s_N depends not only on the agent's component s_i, but also on all other agents'choices $(s_j)_{j\neq i}$.

This, in turn, brings about a difficulty familiar to any game theoretic set-up: the issue of solution concepts. What is the solution of a game? Experience of game theory teaches us that there is no satisfactory single answer to this question. The point is that preferences alone do not allow the outcome of a game to be predicted, although they constitute by and large its most important input. Indeed, another important input is the agents' information on the behaviour of the others, viz. the information each participant has about his and others' strategies and reactions. This, in turn, may depend on some knowledge about the preferences of others, on their ability to communicate with each other, to coordinate their actions and so forth. Game theory has not yet elaborated or provided a fully consistent and complete picture of all this. Instead, various equilibrium concepts have been proposed. Which one of them proves the most adequate for a specific purpose depends usually on the particular situation and taste of the researcher.

The Nash equilibrium concept is by large the most popular solution concept in game theory. It consists in a bundle of strategies $s_N^* = (s_i^*)_{i\in N}$ such that for any agent $i \in N$ the strategy s_i^* is the best possible strategy for i, given that other agents use strategies s_j^*. The content of the Nash solution concept is explained in detail in any game theory textbook (see also Chapter 2 here), thus we do not describe it here. However, note that it is difficult to find discussions about the way by which agents attain equilibrium in textbooks. In fact, often enough we do not have any clear understanding of the agents' behaviour out of equilibrium, which in turn makes it difficult to say which equilibrium they might attain, and whether they generally attain it. One can only hope that when they do reach an equilibrium, they remain there.

There is one interesting and important case in which it is relatively easy to find a Nash equilibrium. This is when there are dominant strategies, that is, when agent i with preference R_i has a strategy s_i^* which is best (with respect to R_i) for any strategy choice of other participants. In this case, indeed, agent i does not have to make sophisticated guesses about what the others will do; he may stick to his dominant strategy s_i^*. If all agents have a dominant strategy, one might confidently expect that the outcome of the game be $\pi(s_N^*)$.

The notion of dominant strategies is not very popular in game theory because game theory focuses on the analysis of concrete games. Such games usually do not support the existence of dominant strategies. The issue is quite different in the theory of mechanism design, where one of the main issues consists in constructing mechanisms which exhibit dominant strategies (see Chapter 3 for more detail). Incidentally, a standard question in mechanism design is whether we can exhibit a SCF for which truthful revelation of preferences would be a dominant strategy for every agent. As we shall see later, there are very few cases in which such a function can be exhibited. In general, the requirement of existence of dominant strategies for all profiles of preferences imposes strong restrictions on mechanisms and often allows a complete description of such mechanisms.

There are other solution concepts for which the agent strategy formation takes place on an individual basis, i.e., without any latent cooperation between agents as is the case in Nash equilibrium set-ups. We mention only the maximin and sophisticated equilibrium concepts, which reflect opposite degrees of agents' knowledge about the interests of the other members of the group. Note also that dominant strategies guarantee neither the optimality of outcomes nor coalitional non-manipulability. This leads us to the class of solution concepts which explicitly take into account the willingness of agents to cooperate and in practice to form coalitions.

Assume agents may communicate freely with one another and coordinate their actions. A new possible difficulty arises since, in effect, any such agreement is itself a social choice problem. For example, participants of some subgroup may find out that, notwithstanding their having agreed upon some actions (and assuming the others keep their strategies fixed), they can improve the outcome for all members of the coalition. It is natural to think that they would swiftly impair the formed Nash equilibrium. An equilibrium is called coalitional or strong if no coalition can improve the outcome of its members. If the reactions (threats) of others are taken into account, then other solution concepts emerge (such as the core). We elaborate upon these issues in more detail in Chapters 4 and 5.

The relationship between the theory of social choice mechanisms and game theory can be summed up as follows. A mechanism π can be viewed as a rule of a game, that is, viewed as its material part. And to set-up a game, one should fix the aims of the agents, that is, a preference profile R_N. Therefore we might associate to any mechanism a series of games which have the same rules but variable preferences. Finally, in order to evaluate how the mechanism functions, we must add a solution concept.

Clearly a game $G(\pi, R_N)$ can have several equilibria. The important thing is to ensure that at least one equilibrium exists, because if not, this means that the solution concept does not describe appropriately the agents' behaviour within the game $G(\pi, R_N)$. We say that the mechanism is consistent with respect to a given solution concept if there exists an equilibrium, for any feasible profile R_N. The SCC, F, is implemented by the mechanism π if $F(R_N)$ is the set of equilibrium outcomes in the game $G(\pi, R_N)$. The implementability of an SCC can be viewed as a desirable property, and indeed, as a sign of its viability. We can think of SCCs as "wishes". Their implementability just means that our wishes can be enforced. Note that other important properties, like monotonicity or Pareto optimality are closely related to the issue of implementability of an SCC. This means in practice that "bad" (perverted, unrealistic or absurd, etc.) SCCs are difficult to implement. Conversely, implementable SCCs turn out to be automatically "good". Thus, in some sense, implementability warrants some desirable properties of social choice.

In this introduction, we discussed mechanisms, and solution concepts. We elaborated upon consistency, and stressed the fact that the theory of social choice mechanisms is a synthesis of abstract social choice theory and game theory. The theory of social choice mechanisms, unlike game theory, does not fix preferences of agents. Rather, it focuses on a wide spectrum of aims. Mechanism design does not only investigate real-life mechanisms. It also, and this justifies its name, devotes its efforts towards the designing of rules exhibiting nice properties. The hope here is to provide new tools for concrete real-life collective decisions. This pragmatic feature of mechanism design theory distinguishes it from both the theory of abstract social choice and game theory.

Acknowledgments

The authors are grateful to Professor Sertel for his gentle suggesting to publish this book in the series of Economic Design and for his great patience. We also warmly thank Ms. C. Lang for her selfless offer to undertake the huge task of improving our English.

1. Basic Concepts

In this chapter we introduce the main concepts used in the sequel, that is preferences and preference profiles, social choice functions and correspondences (Section 1.1). We also recall various properties of these correspondences (Section 1.2), and discuss some important properties like monotonicity (Section 1.3). We define the notion of social choice mechanism (Section 1.4) and its rough form: the effectivity function or blocking (Section 1.5). In Appendices 1.A1 and 1.A2, we discuss two seminal theorems of social choice theory, that of Arrow and that of Gibbard-Satterthwaite. Both are closely related with Mueller-Satterthwaite theorem from Section 1.3. In Appendix 1.A3, we investigate the notion of minimal monotone SCC.

1.1 Preferences

(1.1.1) As we mentioned in the introduction, a social choice problem arises when a group of agents is asked to chose one alternative within a set A of *alternatives*. This set might be a set of dishes in a menu, of books in a library, of candidates in a presidential election, of production plans and so on. In the sequel, however, we consider A as an abstract set. Yet, what we need to know are the preferences of the agents on this set.

Preferences can be described in different ways. Often enough, preferences are given through *utility functions*. Let $u : A \to \mathbf{R}$ be a function on the set A with values in the set $\mathbf{R}$ of real numbers. Intuitively, the number $u(a)$ might be interpreted as the "utility" of alternative $a \in A$. If $u(a) > u(b)$, then alternative a is preferred to alternative b; and if $u(a) \geq u(b)$ then a is no less preferable than b. Thus every utility function u induces a binary relation R on the set $A : aRb$ if and only if $u(a) \geq u(b)$. Nowadays, it is more usual to start from a binary relation R.

The underlying of such an ordinal approach is that, after acquiring sufficient knowledge about the proposed alternatives, an agent is always able for any two different alternatives to state whether he[1] prefers one to the other or whether she considers them equivalent. Such an approach seems to be both

[1] Throughout this book the pronouns "he" and "his" are used in a generic sense encompassing both sexes.

more acceptable and simpler compared to the cardinal approach, because it frees us from the issue of utility units and their comparability between agents. The ordinal approach does not prevent an agent to have a utility function, but it regards somehow the specification of her utility function as being her own business. Of course, in a number of cases, there may be an objective ground to working with an valuation in utility units. For instance, it may occur that alternatives be evaluated in monetary units. However, a strict follower of the ordinal approach might say that enlarging the set A, namely by adding money, does the trick. In effect, in this case, the preferences are stated on pairs of (alternative, money).

Thus it is natural enough to state preferences as binary relations in social choice theory. Transitivity of preferences (or rationality of individual choice) is less innocuous an assumption on preferences. We discuss this issue more formally.

A binary relation on a set A is a subset $R \subset A \times A$. We write aRb or even $a \succeq b$, instead of $(a, b) \in R$, and read it : a is not worse than b. *A weak order* on a set A is a binary relation R satisfying the following two conditions:

1. (Completeness) For any $a, b \in A$, aRb or bRa holds.
2. (Transitivity) If aRb and bRc then aRc.

A weak order is usually called a preference. Preference relations induced by utility functions are weak orders. The converse is true if the set A is finite and also practically in all interesting cases.

Since we shall deal further with weak orders and their properties, it is convenient to introduce here a few concepts and notations. We shall consider predominantely cases in which the set A is finite ($| A |$ denotes the number of elements in A.). However, most concepts introduced here are easily extended to the case of infinite sets.

In the sequel, we consider preferences (or weak orders) of a special type, namely *linear orders*. They satisfy the additional following condition:

3. (Antisymmetry) If aRb and bRa then $a = b$.

Linear orders define finer preferences than weak orders, in that they enable to distinguish any two different alternatives from each other. Restricting ourselves to linear orders, while not altering the generality of our analysis, simplifies somewhat both discussions and further conceptual constructions. Actually, from a theoretical point of view, weak orders are preferable due to the functorial nature of weak orders.

Linear orders are also called rankings, because they rank all elements of $A : x_1 \succ x_2 \succ ... \succ x_m$, $m = |A|$.

One sees readily that the set $\mathbf{L} = \mathbf{L}(A)$ of all linear orders on A contains $m!$ elements. The best element of A with respect to R (in this case x_1) is denoted further $\max R$ and the worst x_m , $\min R$.

Let $X \subset A$ be a subset. $R \mid X$ denotes the restriction of R to X, i.e. a binary relation $R \mid X = R \cap (X \times X)$ on the set X. It is a linear order, therefore the notations $\max R \mid X$ and $\min R \mid X$ are meaningful.

(1.1.2) Shuffling Orders. Let R and Q be two linear orders on A, and $A = X \sqcup \overline{X}$ be a subdivision of the set A into two disjoint subsets X and $\overline{X}$. ($\overline{X}$ will denote the complement to X in A, i.e. $A \setminus X$; the symbol $\sqcup$ denotes a disjoint union). One forms a new order $S = (R \mid X,\ Q \mid \overline{X})$ as follows: any element of X is preferred to any element from $\overline{X}$; within X (correspondingly within $\overline{X}$) the rankings are as in R (correspondingly in Q). For example, let R be the following ranking of alternatives

$$a \succ b \succ c \succ d \succ e \succ f \succ g,$$

and let $X = \{b, e, f\}$, and $\overline{X} = \{a, c, d, g\}$, then the ranking $(R \mid X,\ R \mid \overline{X})$ consists in :

$$b \succ e \succ f \succ a \succ c \succ d \succ g.$$

Expressing this somewhat loosely, we say that in $(R \mid X,\ R \mid \overline{X})$, the subset X is "propped up" and the subset $\overline{X}$ is "propped down". Symbols $(* \mid X,\ * \mid \overline{X})$ or $X \succ \overline{X}$ will denote an arbitrary ranking where X is propped up. The meanings of $(R \mid X,\ *)$, $(*,\ R \mid X)$ are straightforward.

(1.1.3) Lower Contour. In this paragraph, we introduce yet another important and useful concept, that of lower contour. Let R be a linear order, $a \in A$. The set

$$L(a, R) = \{x \in A, aRx\}$$

is called a *lower contour* of R with respect to a. It is the set of elements which are not preferred to a. Evidently,

$$aRb \Leftrightarrow L(a, R) \supset L(b, R).$$

Symmetrically, we define a binary relation $\succeq_a$ on **L,**

$$R \succeq_a Q \Leftrightarrow L(a, R) \supset L(a, Q)$$

where R and Q are preferences on A. Intuitively, $R \succeq_a Q$ means that the alternative a ranks higher (not lower) in R than in Q. Consider the following example:

x	x	a	z
a	a	y	a
y	z	x	x
z	y	z	y
R_1	R_2	R_3	R_4

(The higher the alternative the better.) Here $R_1 \succeq_a R_2$, $R_2 \succeq_a R_1$, $R_1 \succeq_a R_2$ (so R_1 and R_2 are $a-$equivalent), $R_3 \succeq_a R_1$, $R_3 \succeq_a R_2$, $R_3 \succeq_a R_4$; but R_1 and R_4 are incomparable with respect to $\succeq_a$. The relation $\succeq_a$ on the set **L** is reflexive and transitive, but is not complete. Here a certain duality peeps out between alternatives and linear (or weak) orders: orders rank alternatives and alternatives "rank (albeit not completely)" orders. In general, it is useful to treat alternatives and agents in a symmetric way (viz. Danilov and Sotskov (1987b)).

(1.1.4) Preferences and Choice. The choice from A is obvious if there is a preference R on A : it is $\max R$. Sometimes one is restricted in choosing only from a subset $X \subset A$. The natural candidate is then $c_R(X) = \max R \mid X$. Thus every linear order R defines *a choice function*

$$c = c_R : 2^A \setminus \{\emptyset\} \to A$$

such that $c(X) \in X$ for any nonempty $X \subset A$. We denote by 2^A the set of all subsets of A. The single-valuedness of c_R follows from the fact that R is not an arbitrary weak order, but a linear order.

Not all choice functions $c : 2^A \setminus \{\emptyset\} \to A$ can be obtained by this manner. A choice function induced by a linear order, satisfies the following rationality condition:

let $Y \subset X$ and $c(X) \in Y$, then $c(Y) = c(X)$.

It is easy to check that the converse is also true. For every choice function c, we define a binary relation R as follows: xRy if $x \in c(\{x,y\})$. R is called the *revealed preference* and is transitive if c satisfies the rationality condition.

As we mentioned above one could work with weak orders instead of linear orders. On one hand, linear orders are convenient because the maximum and the minimum are unique. On the other hand, restricting to linear orders does not involve a loss in generality. In fact, there are "sufficiently many" linear orders in the sense that for any weak order Q there exists a linear order R s.t. $Q \supset R$ (i.e. order R is finer than Q).

(1.1.5) Preference Profiles. A linear order describes the preferences of an individual agent. To specify the preferences of a group N, we have to give the preferences R_i of every agent $i \in N$. *A preference profile* for the group N or simply *a profile* is an array $(R_i)_{i \in N}$, where $R_i \in \mathbf{L}$. It is denoted ordinarily by R_N. In other words, a profile R_N is an element of Cartesian degree $\mathbf{L}^N$ of the set $\mathbf{L} = \mathbf{L}(A)$ or a function $R_N : N \to \mathbf{L}(A)$. Usually, we represent profiles in tabular form:

x	u	z	x
y	z	y	z
z	x	u	y
u	y	x	u
1	2	3	4

Here the group N consists of four agents 1, 2, 3, 4 and $A = \{x, y, z, u\}$. In the columns above, the best alternatives figure at the top.

It is straightforward to extend the previously introduced concepts to profiles. For example, given a set $X \subset A$, $R_N \mid X = (R_i \mid X,\ i \in N)$ denotes the restriction of the profile to the subset X. If $K \subset N$ is a *coalition* of agents, then R_K denotes the restriction of a profile R_N to K. Given the profiles R_K and $R_{K'}$ of disjoint coalitions K and K', $(R_K, R_{K'})$ denotes a profile of preferences of the coalition $K \sqcup K^{'}$ coinciding with R_K on K and with $R_{K''}$ on K''.

These are natural extensions. But here comes an important concept. Let K, R_K, and a be respectively a coalition, a coalitional profile and an alternative. We define

$$L(a, R_K) = \cup_{i \in K} L(a, R_i) = \{x \in A,\ \exists\, i \in K \ \text{ s.t. } \ aR_ix\}.$$

If $K = \emptyset$, we define $L(a, R_\emptyset) = \{a\}$ by convention.

For example, in the case of the profile given above, $L(y, R_{\{3,4\}}) = \{y, u, x\}$.

Now a fundamental concept.

(1.1.6) Definition. An alternative a is *Pareto optimal* (or *efficient)* for a profile R_N if $L(a, R_N) = A$.

We grasp the full power of this definition by considering its negation. If an alternative a is not efficient, then there exists another alternative x which is strictly better than a for *all* agents.

We denote the set of all Pareto optimal alternatives by $Par(R_N)$.

At last, for any two profiles R_N and R'_N and an alternative a, we write $R_N \succeq_a R'_N$ if $R_i \succeq_a R'_i$ for any $i \in N$.

1.2 Social Choice Correspondences

(1.2.1) Let N be a set of *agents* and A be a set of *alternatives* . In the sequel, we always work with N finite and identify it to the set $\{1, 2, ..., n\}, n = \mid N \mid$ when necessary. To simplify matters, we also assume that the set A is finite, although this is not essential to the analysis. We suppose that $\mid A \mid \geq 2$; the case $\mid A \mid = 2$ being somewhat special.

A social choice function (SCF) is a mapping

$$f : \mathbf{L}(A)^N \longrightarrow A.$$

It is a formalized solution of a social choice problem. It recommends to choose an alternative $f(R_N) \in A$ for any preference profile $R_N \in \mathbf{L}^N$. A SCF can also sometimes be called a constitution. Only a few mappings from $\mathbf{L}(A)^N \to A$ qualify as "reasonable" solutions and in order to be admissible, they should satisfy certain conditions. We discuss some of these conditions below, reviewing the others in Section 3.

Let us start with a few concrete examples of social choice functions.

(1.2.2) Example: *a constant SCF*. Fix an alternative $a \in A$ and for any profile $R_N \in \mathbf{L}^N$, let $f(R_N) = a$.

The *constant* SCF clearly does not account for agent's preferences. It is therefore not overly exciting with respect to the aims of social choice.

(1.2.3) Example: *a dictatorial SCF*. Pick an agent $i_0 \in N$ and pose $f(R_N) = \max R_{i_0}$, where $\max R_{i_0}$ stands for the most preferred alternative for agent i_0.

A *dictatorial* SCF is already more interesting a SCF than the constant SCF. However, it is certainly not the most desirable among all conceivable SCFs. In effect, it is difficult to see where there is here an issue for social choice, since everything rests on the rankings of one agent i_0 - the dictator. Therefore dictatorial SCFs are also not overly exciting as they seem too drastically unilateral.

(1.2.4) Example. We now consider a social choice function with stronger democratic foundations - a *simple majority rule*. It selects the alternative which ranks highest for the largest number of agents. It may occur that two or more alternatives collect the same number of votes. The rule has to provide then for a tie-breaking procedure. Usually it does not matter much what kind of tie-breaking procedure one chooses, but more about this later. For the time being, we would like the reader to notice that this rule calls no less for criticism than the others reviewed above. Consider the following profile R_N, where $\mid N \mid= 5, \mid A \mid= 3$:

x	x	y	y	z
z	z	z	z	x
y	y	x	x	y
1	2	3	4	5

According to this rule we have to choose either x or y. However, three participants out of five (the majority) prefer z to x, and three (different) participants prefer z to y. And this should alert us.

In practical cases, one often uses the following modified version of this rule. Here is a non-formal description of the latter. In a first stage, we consider the two alternatives, which earned the largest amount of votes among the participants. Then, one of them is selected by simple majority. In the case described above, the two first alternatives to be selected are x and y. Then finally x. However, even with this rule the criticism we just raised remains valid, since three participants prefer z to x.

Anticipating a little, any SCF can call for some sound criticism as soon as $|A| \geq 3$. In order to organize these criticisms properly, we need first to list carefully our desiderata about SCFs, then either to control whether these are fulfilled by a given SCF or not, or else try to come up with an SCF adequately suited to the latter. We shall return to this issue later on, but for the time being we discuss another yet unsuccessful attempt at constructing a satisfactory social rule.

(1.2.5) Example. Let R_N be a preference profile. An alternative $a = c(R_N)$ is a *Condorcet winner* if it wins (i.e. collects more than a half of all votes) in pairwise comparisons with any other alternative. For instance, alternative z is a Condorcet winner in Example (1.2.4). Moreover, a Condorcet winner possesses the nice following property: it is uniquely determined. Unfortunately enough, in many cases, there is no Condorcet winner alternative. Consider the following example,

x	x	y	y	z
y	y	z	z	x
z	z	x	x	y
1	2	3	4	5

Here three agents prefer x to y, four agents prefer y to z, and again three agents prefer z to x. Thus there is no Condorcet winner alternative in this case. Condorcet (1785) uncovered this phenomenon, for the first time, while studying a voting issue. In fact, 1785 marks the very beginning of the theoretical study of social choice problems, while the practice of social choice is immemorial.

One might choose to be unworried by the non-existence of a Condorcet winner. In fact, there are some ways to deal with the issue. A first path consists in restricting the rule to the subset of profiles for which such winners exist. A second path consists in adding some other alternative to the set of alternatives, when the issue of non-existence arises. Although neither these paths is a good solution to the issue, none is unfounded.

For instance, in the first alternative path, we could follow the unanimity "rule", which would mean to choose the best preferred alternative for every agent (if such an alternative exists) and refuse to choose otherwise. However the unanimity "rule" yields a choice outcome only in absence of conflict, a rare and exceptional situation. The natural and interesting question - how often do Condorcet winners exist?- is out of the scope of this book.

We already sensed, through our analysis of Example (1.2.4), that many natural procedures whose choice-outcome is single-valued for "general" profiles, may however yield multi-valued choice in some special profile-instances. Thus the specification of a SCF requires almost always the additional specification of a "tie-breaking" procedure. This leads to the following definitions.

(1.2.6) Definition. A *social choice correspondence (SCC)* is a multi-valued mapping (a correspondence)

$$F : \mathbf{L}(A)^N \to 2^A \text{ or } F : \mathbf{L}(A)^N \Rightarrow A.$$

We assume usually that $F(R_N) \neq \emptyset$, for any profile $R_N \in L^N$. The notion of SCC can be interpreted as follows: given a profile R_N, a choice-outcome is an element of the set $F(R_N)$. With this definition, a SCF is nothing else than a single-valued SCC.

Of course, ceteris paribus the finer the SCC, the better. What would be the utility of the following "constant" SCC, which maps every profile R_N into $F(R_N) = A$? However, a SCC comes in quite handy since it is easily transformed into a SCF by any choice function $c : 2^A \setminus \emptyset \to A$. It suffices to pose $f(R_N) = c(F(R_N))$. Often, the choice function c is defined through an auxiliary linear order R_0 on A (see 1.1.4).

Let us examine two examples of social choice correspondences.

(1.2.7) Example: *the Borda rule.* The idea underlying the Borda rule is simple. It generalizes, in fact, the procedure presented in Example (1.2.4). However, there is an important distinction with respect to the simple majority rule. It does not focus only on the "best" alternatives, rather it considers a full-range ranking of alternatives: those ranking second, third etc. Namely the score or points of every alternative is computed as a function of the position it occupies in agents' rankings. The Borda rule then selects the alternatives which collect the greatest number of points. More precisely, it associates with every alternative $x \in A$ a number

$$u(x, R_N) = \sum_{i \in N} \mid L(x, R_i) \mid .$$

Recall that the integer $\mid L(x, R_i) \mid$ indicates the place occupied by x in ranking R_i as measured with respect to the bottom. Then we set

$$F(R_N) = \operatorname{Argmax} u(., R_N).$$

Thus $F(R_N)$ is the set of alternatives yielding the largest values of the function $u(., R_N)$.

The construction of the Borda rule rests on the number of elements in $L(x, R_i)$, but in the same spirit, one could construct other rules by considering any set function depending monotonously on $L(x, R_i)$. In fact, the simple majority rule in Example (1.2.4), is just a particular case. Incidentally, the Borda rule is used in determining the best soccer player in soccer clubs. Soccer experts are asked to rank the three best players in decreasing order; then a computation operation is performed on each individual ranking: a player in first position collects three points, two points if in second position, one point if in third position; eventually, each player's final number of points over all experts's ranking is computed and the player collecting the highest number of points is declared the best player. The advantage of these types of rules is that they can be applied to any N and A.

(1.2.8) Example: *the Pareto rule.* The Pareto rule is based on a simpler idea than the Borda rule. One compares and selects alternatives with respect to the size of the set $L(x, R_N)$ (the largest possible set $L(x, R_N)$ being clearly equal to A). The set $Par(R_N)$ is never empty (when A is finite), since it clearly contains $\max R_i$ for every agent i. Therefore $Par : L^N \Longrightarrow A$ is fully-fledged SCC. However, $Par(R_N)$ may end up being quite large: somewhat of a drawback.

(1.2.9) Properties of Social Choice Correspondences. In our review of a few examples of SCCs, we discovered that each of these rules had some desirable and undesirable features. It is now time to discuss the properties of social choice correspondences more thoroughly. We cover only the most standard properties here.

We suggested above that fine SCCs, i.e. for which the sets $F(R_N)$ are of small size, are desirable. Thus a desirable property of SCCs is "fineness", in the sense of associated size of the sets $F(R_N)$ (the finest SCC being a SCF). However "fineness" is neither the most important nor the sole requirement on SCCs (see the conditions listed below). Therefore it is often necessary to sacrifice single-valuedness in order to satisfy other important requirements. It is difficult to evaluate either when it is legitimate to sacrifice single-valuedness with respect to other requirements or when on the contrary to consider that the set $F(R_N)$ is really too "big". For example, one usually agrees that the Pareto rule is a good rule, albeit not overly refined: the set $F(R_N)$ is very big (see Makarov et al. (1982) about evaluations of number of Pareto points).

It is not necessary to force all Pareto optimal alternatives to figure in the SCC. A more important requirement is that $F(R_N)$ consists of Pareto optimal alternatives only. In fact, it would be rather absurd to offer in the choice menu an alternative, when there exists another alternative strictly preferred by all. However, it may sometimes happen that some socially or otherwise meaningful, albeit not Pareto optimal, alternative ought be included in the choice menu of a SCC.

(1.2.10) Definition. An SCC F is *Pareto-optimal* (or *efficient*) if, for any profile R_N,

$$F(R_N) \subset Par(R_N).$$

The efficiency property is satisfied for a majority of "reasonable" social choice rules. In particular, the Pareto rule, the Borda rule and almost all its variations, the dictatorial rule in Example (1.2.3) are efficient. On the contrary, the constant rule (in Example (1.2.2)) is not efficient. Moreover the reader will notice that efficiency of a SCC implies another property: for any alternative $a \in A$ there exists a profile R_N such that $F(R_N) = \{a\}$. This property is usually called *sovereignty.*

The following two properties - anonymity and neutrality - reflect respectively some requirement about equality of agents and/or alternatives. Anonymity expresses the idea of equality of agents. Let σ be a permutation of N, i.e. a one to one mapping of N into N. Given a profile R_N, we denote by R_N^σ the permuted profile in which agent i's preference is given by $R_{\sigma(i)}$. An SCC F is called *anonymous* if

$$F(R_N) = F(R_N^\sigma)$$

for any profile R_N and any permutation $\sigma : N \to N$.

We can visualize anonymity as follows: suppose a center collects bulletins bearing agents' signatures, however its decisions doesn't depend upon them.

The rules in Examples (1.2.2), (1.2.4), (1.2.5), (1.2.7) and (1.2.8) are anonymous, but the rule in Example (1.2.3) is not. This is understandably because " dictatorship" and "anonymity" stand completely in opposition on the scale of democracy. Anonymity is a desirable property in many cases, but we shouldn't erect is as an absolute principle. In fact, one may be drawn, to comply with other desirable properties, into allowing for "small deviations" from the anonymity principle. (We shall elaborate on this issue after introducing the notion of blocking.)

The neutrality condition performs an analogous role with respect to equality of alternatives. Let $\rho : A \to A$ be a permutation on the set of alternatives. Given a preference R, R^ρ denotes the following permuted order:

$$xR^\rho y \Leftrightarrow \rho(x)R\rho(y).$$

Analogously R_N^ρ denotes the permutation of a profile R_N. *Neutrality* of an SCC means that :

$$F(R_N) = \rho(F(R_N^\rho)).$$

In other words, if we permute the alternatives through ρ, the new profile R_N^ρ selects those alternatives $(F(R_N^\rho))$ which rank in the position which were occupied by the alternatives in $F(R_N)$. An example will clarify the point. Consider the following profile R_N :

x	y	z
y	x	x
z	z	y
1	2	3

and assume x is chosen for that profile. Then if we consider the new profile R'_N :

z	x	y
x	z	z
y	y	x
1	2	3

the neutrality property implies that alternative z be selected. The social choice correspondences in Examples (1.2.3-8) are neutral, while the constant

SCF, being precisely constructed to select one particular alternative whatever the profile, is obviously not neutral. As what regards the desirability of neutrality for SCCs, the remarks made above about anonymity apply.

Returning to SCFs, we have to add the following caveat: it may be impossible for a SCF to satisfy both anonymity and neutrality at once. Consider the following example with three agents and three alternatives. Assume the profile R_N looks like this,

x	y	z
y	z	x
z	x	y
1	2	3

then the three alternatives x, y, z are absolutely symmetric and picking any one of them violates symmetry.

In this context, we mention a result by Moulin (1983):

(1.2.11) Proposition. *If* $\mid N \mid = n$ *is divided by a number* $\leq \mid A \mid$ *then there does not exist a SCF which is simultaneously efficient, anonymous and neutral.*

It is important to remark here that the proposition applies to SCFs only. In contrast, a SCC may satisfy all three properties (viz. Examples (1.2.7-8)) at once. In one word, we shall regard anonymity and neutrality as desirable, albeit non-obligatory, properties of SCFs.

We devote the following section to a last property, monotonicity, which we believe to be fundamental for SCFs and SCCs.

1.3 Monotone Social Choice Correspondences

(1.3.1) Interesting structures, deep and nontrivial assertions emerge in social choice when the choice outcomes for different, but somewhat similar, profiles are related to each other by some special features. In such cases, SCFs (or for the matter, SCCs) become more than a simple array of choice outcomes for varying profiles. In fact, it has a consistent structure.

Neutrality property somehow addresses this issue with respect to alternatives. The idea underlying monotonicity is similar. It goes as follows : suppose that in one instance, we behaved in a certain way, then we should behave in a similar way if presented with a somewhat close instance.

We return to the case exposed in Example (1.2.4) in order to improve our understanding.The profile R_N is,

x	x	y	y	z
z	z	z	z	x
y	y	x	x	y
1	2	3	4	5

Recall that the alternatives x and y were chosen, in the first round, since they collected the largest number of votes. In the second stage, alternative x won over alternative y in the confrontation between x and y. Thus, in this case, the outcome is: $f(R_N) = x$.

Assume now, that the preferences R_4 of agent 4 are modified as follows, i.e. $R'_4 = (z \succ x \succ y)$,

x	x	y	z	z
z	z	z	x	x
y	y	x	y	y
1	2	3	4	5

In this new setting, the alternatives x and z are chosen in the first round, and alternative z wins over x by three votes against two, in the second round. Thus the outcome of choice is $f(R'_N) = z$. The reader might note the paradoxical nature of this result. In the change from profile R_N to R'_N the position of alternative x in all agents rankings remained identical or increased (it moved from the last place to the second place in agent 4's ranking). This increase in "attractiveness" of x is not reflected in the social choice outcome, since oddly so, x fails to win the final round!

Moulin (1983) in Ch. 3, Sec. 2 documents even more striking an example, in which a given alternative fails to win the election, albeit being propped from the second place to the first one.

Of course, it is an undesirable feature of a social rule that the improving of a candidate's position hinder his being elected. Monotonicity (or strong positive association and so forth) helps prevent such drawbacks.

Recall that $R_N \preceq_a R'_N$ means that for any agent $aR_ix \Rightarrow aR'_Nx$ (i.e. $L(x, R_i) \subset L(x, R'_i)$ for any $i \in N$). Intuitively, the position of a in profile R'_N improves with respect to profile R_N.

(1.3.2) Definition. An SCC $F : \mathbf{L}^N \Longrightarrow A$ is *monotone* if $a \in F(R_N)$ and $R_N \preceq_a R'_N$ imply $a \in F(R'_N)$.

This property is also called sometimes in the literature - the strong monotonicity property (Moulin (1983) and Peleg (1984)) or Maskin monotonicity. Monotonicity seems very desirable a property for an SCC. The Pareto rule and the dictator rule are monotone as is readily seen from the behaviour of $L(a, .)$. In general, the Borda rule and its modifications are non-monotone. Related to this, we ask : how many monotone SCCs or SCFs exists there? How can they be constructed? Is it possible to describe them exhaustively? These issues are neither simple nor are they yet solved. In the meanwhile, we give two simple examples.

(1.3.3) Example. Fix an alternative $a \in A$ ("status quo") and pose

$$U(R_N) = U(a, R_N) = \{x \in A,\ xR_i a \ \forall\ i \in N\}$$

for a profile R_N.

It is easy to understand that this SCC is monotone.

(1.3.4) Example. *Maskin correspondence.* As above, fix $a \in A$. One readily checks that :

$$Par(R_N \mid U(R_N)) = Par(R_N) \cap U(R_N).$$

The correspondence $M : \mathbf{L}^N \Longrightarrow A$, defined by this formula, is also monotone. This follows from monotonicity of both Par and U and from the following simple assertion :

(1.3.5) Lemma. *The intersection and the union of monotone correspondences yields monotone correspondences.* ■

Notice that the correspondence $U(a, .)$ is obtained as the intersection of the simpler correspondences $U_i(a)$:

$$U_i(a, R_N) = \{x \in A,\ xR_i a\}.$$

Multi-valuedness is the principal drawback of the Pareto rule and even of the Maskin rule. Therefore the following important question arises naturally: how large is the class of monotone SCFs? We will assume that a SCF is surjective, i.e. sovereign. It turns out the answer differs widely depending on whether $\mid A \mid$ is equal to two or larger than two. When $\mid A \mid = 2$, the set of monotone SCFs is large. Mueller and Satterthwaite (1977) establish that this set consists only of dictatorial SCFs when $\mid A \mid > 2$.

(1.3.6) Theorem (Mueller-Satterthwaite). *Let $f : \mathbf{L}^N \Longrightarrow A$ be a sovereign SCF and $\mid A \mid \geq 3$. If f is monotone, then f is dictatorial.*

The remainder of this Section is devoted to proving this Theorem, incidentally we examine the case $\mid A \mid = 2$.

First, we establish that a sovereign monotone SCF is efficient. More precisely, we prove the following lemma.

(1.3.7) Lemma. *Let F be a monotone SCC possessing the following property: $F(R_N) = \{a\}$ when an alternative a is in top position for all rankings R_i. Then F is efficient.*

Proof of the lemma. Assume that F is not efficient, i.e. assume a profile R_N and two distinct alternatives x and y, such that xR_iy, for all $i \in N$, and that $y \in F(R_N)$. Construct now, the profile $R'_N : R'_N = (R_i \mid \{x, y\}, *)$ where x and y have been propped up with respect to R_N. Then $R_N \preceq_y R'_N$ and by monotonicity $y \in F(R'_N)$. On the other hand, $x = \max R'_i$, for all $i \in N$, therefore $F(R'_N) = \{x\}$. Contradiction.■

If an SCF f is sovereign then, for any a, there exists a profile Q_N with $f(Q_N) = a$. If profile R_N is such that the alternative a is in top position for all rankings R_i, then $Q_N \preceq_a R_N$ and by monotonicity $f(R_N) = a$. Thus, by Lemma (1.3.7), f is efficient. Therefore, without loss of generality, we can assume f efficient.

(1.3.8) Now we start with a monotone efficient SCF f. Take an alternative x and a coalition K. Then all profiles of the form

$$\begin{array}{|c|c|} x & * \\ * & x \\ \hline K & \overline{K} \end{array}$$

are x-equivalent. Thus, by monotonicity, if the property $f(R_N) = x$ obtains for one of the profiles above, it does for all such profiles and in this case, we say that the coalition K *forces* x. Denote by $W(x)$ the set of all coalitions forcing x. $W(x)$ possesses the following formal properties:

1. $N \in W(x)$ (this is a consequence of efficiency).
2. If $K \in W(x)$ and $K \subset K'$ then $K' \subset W(x)$. This follows simply from monotonicity.

The third important property is phrased in the following Lemma.

(1.3.9) Lemma. *Let x and y be two distinct alternatives and $K \subset N$. Then $K \in W(x)$ if and only if $\overline{K} = N \setminus K \notin W(y)$.*

Proof of the lemma. Suppose first that $K \in W(x), \overline{K} \in W(y)$ and consider a profile R_N of the form :

$$\begin{array}{|c|c|} x & x \\ * & * \\ \hline K & \overline{K} \end{array}.$$

Then $x = f(R_N) = y$, contradiction.

Suppose now that $K \notin W(x), \overline{K} \notin W(y)$ and consider a profile Q_N

$$\begin{array}{|c|c|} x & y \\ y & x \\ * & * \\ \hline K & \overline{K} \end{array}.$$

By Pareto optimality, $f(Q_N) \in \{x, y\}$. By symmetry, one can assume that $f(Q_N) = x$. Take a new profile

$$\begin{array}{|c|c|} x & y \\ y & * \\ * & x \\ \hline K & \overline{K} \end{array}$$

where x has been propped down in the preferences of all agents belonging to coalition $\overline{K}$. Since $Q_N \succeq_y Q'_N$, then $y \neq f(Q'_N)$. And since by Pareto optimality $f(Q'_N) \in Par(Q'_N) = \{x, y\}$, then $f(Q'_N) = x$, which contradicts the assumption $K \notin W(x)$. ■

The number of elements in A did not intervene up to now. It is time to separate the cases $| A |= 2$ and $| A |> 2$.

(1.3.10) A Two Alternatives Set-up. In this case, we can say nothing more about W than properties 1, 2 and Lemma (1.3.9). In fact, take any set $W \subset 2^N$ satisfying properties 1 and 2 and define the SCF

$$f : \mathbf{L}(\{x, y\})^N \to \{x, y\}$$

by the explicit formula: $f(R_N) = x$ if a coalition $\{i \in N, x = \max R_i\}$ belongs to W, and $f(R_N) = y$ otherwise.

Property 2 implies that the rule will be monotone.

If we are interested in neutral SCFs, the collection of "winning" coalitions W has to satisfy the following condition:

3. $K \in W$ if and only if $\bar{K} \notin W$.

The set of coalitions $W \subset 2^N$ such that properties 1,2 obtain is called sometimes a *simple game.* If in addition property 3 obtains, then this set is called a *maximal simple game* or *a majority family.* This is a generalization of the notion of simple majority, where we say that a coalition wins if it includes more than half of all agents ($| N |$ is supposed to be odd).

These structures appeared long ago both in game theory and social choice theory. Von Neumann and Morgenstern (1945) already undertook the describing of such structures; they noted that not all majority families are of the weighted majority type. Monjardet (1978) (and before him, the logician E.Post (1941)) was able to give the following nice description of all majority families. He remarked that there is a ternary operation, which he calls 'median',

$$m(W, W', W'') = (W \cap W') \cup (W' \cap W'') \cup (W'' \cap W),$$

such that for any three majority families W, W', W'', the median yields a new majority family $m(W, W', W'')$. *Monjardet theorem* states that any majority family can be constructed by starting from a set of dictatorial families (when W consists of all coalitions containing dictator i) and applying to it several times the median operation. In the sequel, we use the following formal assertion about majority families. A family $W \subset 2^N$, satisfying properties 1 and 2, is called a *filter* if the following condition holds:

4. If $K, K' \in W$ then $K \cap K' \in W$.

(1.3.11) Lemma. *A family $W \subset 2^N$ satisfying properties 1-4 is dictatorial.*

A family W is dictatorial if W is a maximal filter (or a ultrafilter); that is $W = \{K \subset N,\, i_0 \in K\}$, where i_0 is a "dictator". Of course, it is crucial that N be finite here. The proof of Lemma (1.3.11) goes as follows. Consider the intersection C of all coalitions in W. The coalition C is an element of W by virtue of property 4 and therefore is the least winning coalition. By the properties 1 and 3, C is non-empty. Let $i \in C$. Since coalition $C \backslash \{i\}$ does not belong to W, then by property 3 its complement $\overline{C} \cup \{i\}$ belongs to W and must contain C, i.e. $\{i\} \in C$. Thus W consists of all coalitions containing "dictator" i. ■

(1.3.12) The $\mid A \mid > 2$ Set-up. Recall that above, given a monotone SCF f, we associated to every alternative $x \in A$, the set $W(x)$ of coalitions forcing x and proved properties 1-3. When $\mid A \mid \geq 3$, we end up with two additional properties. The first one is almost obvious : $W(x)$ does not depend on x, for all $x \in A$ (and we write simply W).

Indeed, we show that $W(x) = W(y)$ for any two alternatives x and y. To this end, we take a third alternative z different from x and y. Let $K \in W(x)$; then by Lemma 3 $\overline{K} \notin W(z)$ and by the same Lemma, $K \in W(y)$. We get $W(x) \subset W(y)$ and by symmetry, the converse inclusion obtains.

In particular, we note that the rule f is neutral. The second property generalizes Lemma (3.9).

(1.3.13) Lemma. *Let $N = K_1 \sqcup K_2 \sqcup K_3$ be a partition of N into three disjoint coalitions $K_i, i = 1,2,3$. Then one of the coalitions K_i belongs to W.*

Proof. We take three different alternatives x, y, z and a profile R_N

$$\begin{array}{|c|c|c|} x & y & z \\ y & z & x \\ z & x & y \\ * & * & * \\ \hline K_1 & K_2 & K_3 \end{array}\,.$$

Efficiency implies $f(R_N) \in \{x, y, z\}$. Assume $f(R_N) = x$. We claim then that $K_1 \in W(x) = W$.

Exchange the positions of y and z in the preferences of all members belonging to coalition K_2, this yields a profile R'_N

$$\begin{array}{|c|c|c|} x & z & z \\ y & y & x \\ z & x & y \\ * & * & * \\ \hline K_1 & K_2 & K_3 \end{array}\,.$$

Since the position of x remaining unchanged (i.e. $R'_N \approx_x R_N$), monotonicity of f implies $f(R'_N) = x$. From the other side, if $K_1 \notin W$ then by lemma

(1.3.9) $K_2 \cup K_3 \in W$ and thus (by monotonicity) $f(R'_N) = z$. Contradiction. Therefore $K_1 \in W$. ■

The Mueller-Satterthwaite theorem comes out now quite easily. Let C be a minimal (in the sense of inclusion) coalition from W and $i \in C$. Suppose that $C \neq \{i\}$. Then, C being minimal in W, coalitions $\{i\}$ and $C \setminus \{i\}$ do not belong to W nor does the coalition $\overline{C}$ (Lemma (1.3.9)). This contradicts Lemma (1.3.13). Thus $C = \{i\}$. If $K \ni i$ then $K \in W$ by monotonicity of W. ■

The main lesson from this theorem is that one cannot avoid multi-valued SCCs. Moreover, assume we want all together to work with monotone SCCs, universal environments and consider set-ups with more than two alternatives, then it states that we have to allow for multiplicity of choices for some strongly conflicting profiles. The proviso about universal environment is crucial; in effect, we shall see later on that there exists monotone SCCs for certain restricted classes of profiles (roughly speaking when a Condorcet winner obtains).

1.4 Social Choice Mechanisms

(1.4.1) Until 1973, social choice theory, implicitly or explicitly, adopted a non strategic viewpoint. Either participants revealed truthfully their preferences, or it was simply assumed that the preferences were known. The previously mentioned properties of SCCs make sense only in this set-up. Consider for instance efficiency, then it can be shown that an efficient rule may yield a non Pareto optimal outcome if agents behave strategically. Take the following example : two agents (or for the matter two distinct groups of homogeneous agents, more or less equal in size) and their preference profile $R_{\{1,2\}}$:

x	y
y	x
z	z
t	t
1	2

We use the Borda rule defined in Example (1.2.5). For this profile, either x or y has to be a winning alternative, because x and y collect the largest number (equal to 7) of points. In order to decrease the winning chances of y and increase the winning chances of x, the first agent may claim her preference to be $R_1^* = (x \succ z \succ t \succ y)$ instead of R_1. Agent 2 similarly might declare $R_2^* = (y \succ z \succ t \succ x)$. Therefore, for this new profile R_N^*, alternatives x and y collect 5 points each and open the way to alternative z, which collects 6 points. But alternative z is not efficient with respect to the true profile of preferences R_N!

This example was meant only to emphasize the difficulties arising when agents behave strategically. It should convince the reader of the necessity of both other approaches and methods of solving the problem. This is the aim of the theory of social choice mechanisms; it is motivated by the possibility of strategic behaviour.

(1.4.2) Definitions. A social choice mechanism (or mechanism) is simply and only a game rule; the aims of players are not specified. More formally, given N and A, *a mechanism (*or *a game form)* consists in a family of subsets $(S_i, i \in N)$ and a mapping,

$$\pi : \prod_{i \in N} S_i \longrightarrow A.$$

The elements of S_i are called the *messages* or *strategies* of agent $i \in N$. The mapping π is called an *outcome function.*

The mechanism is articulated as follows: first agents prompt messages $s_i \in S_i$, then these messages are gathered in a *strategy profile* $s_N = (s_1, ..., s_n) \in S_1 \times ... \times S_n = S_N$, and finally the mechanism yields an outcome $\pi(s_N) \in A$.

Any social choice function $f : \mathbf{L}^N \to A$ can be viewed as a mechanism whose strategies are preferences, or here linear orders, on A. We call such mechanisms *direct mechanisms.* Preferences play a double role in these mechanisms: they reflect the aims of agents and serve as means to full-fil them. If we admit that aims and means might differ, then messages shouldn't be necessarily specified as linear orders. In fact, we can use anything as a message as long as both the list of possible messages is clearly outlined and the outcome function is known to all participants. Therefore each participant knows where this or that message profile might lead to.

Our not concentrating on direct mechanisms solely provides for extended and new possibilities to conceive mechanisms, as for example in the *king-maker* mechanism. Consider a group of agents. The first agent - the king-maker - chooses a "king" among the remaining participants: this " king" announces some alternative, which will be outcome. Of course, this is not yet strictly speaking a mechanism. We would really then need to specify both the strategies of agents and the outcome function. Even then, there are many possible available specifications. Here is one. Let the set $S_1 = N \backslash \{1\}$ and for the remaining agents $S_i = A$. The mapping π is given by :

$$\pi(s_1, ..., s_n) = s_{s_1}.$$

A mechanism really can be viewed somehow as an instrument which participants give life to by their actions. The mysterious part of how the mechanism operates lies in the actions of agents, in their behaviour, in the choice of strategies s_i. Of course, the selecting of a mechanism step is also an important aspect of the whole issue. An imperfect and coarse mechanism can bring to nought the efforts of the most skillful agent, whereas a good one will serve any kind of agent's purpose, even the less experienced. The primary focus of this book is the examine the impact of a mechanism's structure on social choice.

As we mentioned in the introduction, a mechanism $\pi : \prod_{i \in N} S_i \longrightarrow A$ transforms the problem of choosing an element from A into the problem of choosing an element from S_N. Where do we improve here? Well, now agent i is totally free to choose any element in the set S_i and this is just the main idea of a mechanism - to untie choices of agents, to reduce the social choice problem to individual choices. This simplification is however somehow illusory. In effect, the outcome still depends upon joint actions, upon the choices of all agents.

What strategies s_i^* will the agents choose? This really depends on their preferences R_i in the first place. In effect, as soon as we supplement a mechanism $\pi : \prod_{i \in N} S_i \longrightarrow A$ with a preference profile R_N, we obtain a game $G(\pi, R_N)$ in the usual sense. Note that "gains" are neither evaluated in "monetary units" nor in real numbers here. But this is un-essential. Thus, having fixed a profile R_N, we enter in the realm of game theory and thereby we expect to be instructed about the behaviour of agents. But we wait in vain for game theory provides us with many too many an answer. In short, this is because agents' behaviour is not only determined by their preferences R_i but also by many additional things such as : information about the mechanism, about the preferences and strategies of the other agents, ability to conduct negotiations, to form coalitions, to agree upon strategies and to threaten, ability to collect themselves, to put themselves in the place of other agents, to figure out the knowledge of other agents and their strategic abilities, and so on to infinity.

Of course, even if these things aren't always stipulated explicitly, one should try to take them into account. However, we shall not explore the wild territories of human and social psychology here, rather we make the following compromission. We view preference profiles R_N as the main factor influencing message formation s_i and assume that all the rest is clad (hidden) within the "behaviour concept", which is formalized by an equilibrium concept (or solution of a game).

(1.4.3) Solution Concept. The equilibrium approach (in games or in more general settings) originates in a refusal to state or to predict what agents would do, which strategies (s_i^*) they might choose and what the outcome $\pi(s_N^*)$ of the game would be. Instead it formulates some conditions for which a strategic profile s_N^* might be considered stable and might be realized as an equilibrium. These conditions usually imply that agents have exhausted their opportunities for a better outcome. The solution concept neither tells us what happens outside equilibrium, nor does it inform us on which equilibrium will be selected or whether any equilibrium would be selected at all. In many cases, however, the emphasis bears on the issue of existence of equilibrium, maybe simply because existence provides us with a slight hope to actually do some theory.

Game theory has elaborated many solution concepts, see Moulin (1985) or Myerson (1991), Harsanyi and Selten (1988). The most famous and the most broadly used concept is of course that of Nash equilibrium. But there are many other concepts, for instance, that of sophisticated equilibria, dominant-strategy equilibria, coalitional equilibria, core etc. We shall discuss some of these concepts later on, but meanwhile we pursue the course of our general considerations. *An equilibrium concept* E picks out a set of "equilibrium" situations $E(\pi, R_N) \subset S_N = \prod_{i\in N} S_i$ for a profile R_N. This yields a correspondence

$$E(\pi) : \mathbf{L}^N \Longrightarrow S_N.$$

The set of equilibrium outcomes, is the image of the set $E(\pi, R_N)$ by the mapping $\pi : S_N \to A$. This builds up a new correspondence

$$F = \pi(E(\pi, .)) : \mathbf{L}^N \Longrightarrow A.$$

We represent it with a commutative diagram, see Fig. 1.

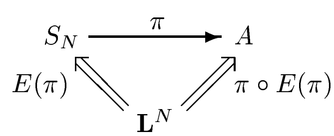

Fig. 1

The correspondence $F = \pi(E(\pi, .))$ is said to be $E-$*implemented* by the mechanism π. Moreover, to ensure that F be a SCC, we should check that the set $E(\pi, R_N)$ is not empty for every preference profile $R_N \in \mathbf{L}^N$. In this case, the mechanism π is called $E-$*consistent*.

(1.4.4) Environment. A consistent mechanism yields an "equilibrium" outcome for every preference profile R_N. However, even if the set $E(\pi, R_N)$ might happen to be empty for some profiles, there may be a subset of the set of profiles for which it is not empty. An *environment* is a subset $D \subset \mathbf{L}(A)^N$ (the environment $\mathbf{L}(A)^N$ is called *universal).* A priori, D is any arbitrary subset of $\mathbf{L}(A)^N$, but usually it is a Cartesian product $\times_i D_i$, where D_i is a set of *admissible* individual preferences. Restricted environments may be of interest either when agents' preferences are not arbitrary and belong to a set given a priori or when the preferences of distinct agents are interdependent. This arises usually when the set of alternatives A (or the group N) possesses an additional structure, with which the preferences of agents should be compatible. The following examples show how additional structure and compatibility interact.

1. In economics, the set of alternatives A usually figures bundles of goods and the classroom set-up assumes " that the more goods the better". In this

case, there is an a priori structure on the set A, namely a partial order $\geq$. Remark then that individual preferences R_i have to be compatible with this order in the sense that $\geq \subset R_i$. Therefore, not all preferences are admissible.

Consider a somewhat close situation. Assume there is a worst element $a \in A$. Then we say that preferences R_i are admissible here if $\min R_i = a$. The existence of such an element might help to agents to reach an agreement.

2. The set A may have a natural topological structure (in such case A is usually infinite). Then a natural compatibility condition might be continuity of preferences R_i (for example, assuming R_i is closed as a subset of $A \times A$).

3. The set A may have a convex structure. A compatibility conditions might be that preferences be convex in the sense that a mixture of two alternatives is not worse than one of them. This defines a convex environment.

There is an interesting variation of this idea. Assume that we wish to extend a preference order, given on an initial set A, to the set 2^A of all subsets of alternatives. Then a natural requirement for extension is that any subset $X \subset A$ is not worse than the worst element of X, i.e. $\min R \mid X$ (Kannai and Peleg (1984); Danilov and Sotskov (1987)).

4. It may occur that the set A has a natural metric ρ. Then one might associate to every $a \in A$ the preference on A, defined by the utility function $u_a(.) = -\rho(., a)$, where the closer an alternative x to the "bliss point" a the better it is valued. This set-up is related somehow to that of single-peakedness.

5. Single-peaked or unimodal preferences are very popular in the social choice literature. They are indeed particular cases of 3 and 4. Assume the set A has a tree structure, i.e. A is a connected graph without cycles. For any two $x, y \in A$, we denote by $[x, y]$ the minimal connected subgraph containing x and y. A preference R is said to be *single-peaked* if $aR \min R \mid \{x, y\}$ for any $a \in [x, y]$. For more details, refer to Demange (1982) or Danilov (1994).

6. Many economic settings do not to fit in the frameworks set-above and therefore it is not possible to devise the previously devised restrictions on preferences . Of course in these settings, preferences are rarely arbitrary. However it is often quite difficult to describe appropriately the class of admissible preferences. This is why we shall focus on general social choice mechanisms here. A few economic applications of social choice theory appear in Sections 4 and 5 of Chapter 3.

The set of consistent mechanisms is larger the narrower the environment. But of course, mechanisms which are consistent for broad environments are more interesting. We shall focus in our book with universal environments, at least most of the times. And when possible, we shall describe the structure of universally consistent mechanisms. In Chapter 3 however, we work with a restricted environment while investigating dominant equilibrium concepts.

Once we have a consistent mechanism, we can try to mesure how well it performs. Do the implemented equilibrium outcomes have "nice" properties? It turns out that, for many interesting equilibrium concepts, the implemented

SCC has automatically nice properties such as monotonicity. Thus we may now add implementability to the list of desirable properties of a SCC for we feel that probably only "reasonable" SCCs may be implemented by consistent mechanisms.

The following four chapters examine separately four most important game solution concepts: Nash equilibrium, equilibrium in dominant strategies , core or "threat " equilibrium, and strong (coalitional) equilibrium. For each concept, we investigate the associated consistent mechanisms and implemented SCCs. We do not elaborate on the sophisticated equilibrium concept, because we think it is both too refined and unrealistic. The interested reader should consult Moulin (1983), Golberg and Gurvich (1986). We neither account for the Maxmin concept because it has not been investigated widely.

1.5 Effectivity Functions and Blockings

(1.5.1) When we analyze a social choice mechanism, we are interested in evaluating (at least coarsely) the power of both agents and coalitions, i.e. their ability to impose outcomes through a mechanism. Of course, this "power" is not necessarily expressed as a real number (though sometimes it can be reduced to a number). It expresses the capacity of an agent or a coalition to force some outcomes, to monitor them into some subset(s) of alternatives or equivalently to block them from the complementary subset(s). Take, for example, a dictatorial SCF. Obviously, the division of power is such that all the power lies in the hands of the dictator (with respect to the choice of an alternative in A), whereas the remaining agents are powerless (they are unable to influence outcome). Now take the simple majority rule (Example (1.2.4)); in contrast here, any coalition including more than a half of all agents has practically full-powers. Therefore such a coalition can enforce any decision provided that its members have reached an internal agreement. For other mechanisms, a coalition's power might take any intermediate value between "full-power" and "no-power".

A coalition *forces* a subset $X \subset A$ (or is *effective* on X), if it is able to drive outcomes into the subset X. It *blocks* a subset Y, if it is able to avoid outcome from "hitting" into the subset Y. Thus, to enforce X is equivalent to blocking $\overline{X}$. Depending on circumstances, we shall use either terms. However to avoid cumbersome repetitions, definitions will be stated in most cases in terms of blocking; the stating in terms of effectivity functions is straightforward.

(1.5.2) Definition. Let $\pi : \prod_{i \in N} S_i \longrightarrow A$ be a mechanism. A coalition $K \subset N$ *blocks* a subset $X \subset A$ (or *forces* a supplementary subset $\overline{X} = A \setminus X$) through mechanism π, if there exists a coalitional strategy $s_K \in S_K$ such that, for any strategy $s_{\overline{K}} \in S_{\overline{K}}$ of the complementary coalition $\overline{K}$, $\pi(s_K, s_{\overline{K}}) \notin X$. We write for simplicity $K B_\pi X$.

It is easy to check that the blocking relation[2] B_π (between 2^N and 2^A), generated by the mechanism π, has three fundamental properties:

B1. If K blocks $X, K' \supset K$ and $X' \subset X$, then K' blocks X'.

B2. If K blocks X, K' blocks X' and $K \cap K' = \emptyset$, then coalition $K \cup K'$ blocks $X \cup X'$.

B3. Every coalition blocks the empty set; no coalition blocks A.

Property B1 is called *monotonicity* with respect to participants and alternatives. Property B2 is called *superadditivity*: two disjoint coalitions which decide to unite, may achieve together no less than separately. Property B3 provides some sort of 'boundary' condition; it states that some outcome has to occur and that the set A exhausts all possible outcomes. Note that if the mapping $\pi : S_N \to A$ is surjective, then B_π is sovereign in the following sense :

B4. The total coalition N blocks any subset $X \subset A$, $X \neq A$.

(1.5.3) Definition. A binary relation B between sets 2^N and 2^A is called *a blocking* if it satisfies axioms B1-B3.

The following *basic property of blockings* derives from axioms B1-B3:

assume that $K_1, ..., K_m$ be pairwise disjoint coalitions, blocking respectively the sets $X_1, ..., X_m$; then $X_1 \cup ... \cup X_m \neq A$.

In particular, note that if K blocks X then $\overline{K}$ does not block $\overline{X}$.

Theoretically a blocking need not to derive only from mechanisms as in definition 1. In fact, if some agents have the necessary resources to the realization of some projects or some a priori rights then the opportunity to "block" some alternatives may arise naturally. And in these more general cases, the emerging "blocking relation" might not satisfy conditions B1-B3. In particular, the "blocking" may be self-conflicting and allow for no outcome. Consider the following broad way of constructing a blocking.

(1.5.4) Definition. Let $\pi : S_N \to A$ be a mechanism. A coalition K $\beta-$*blocks* a subset X through the mechanism π if for any strategy $s_{\overline{K}}$ of the complementary coalition $\overline{K}$, there exists $s_K \in S_K$ such that $\pi(s_K, s_{\overline{K}}) \notin X$.

In general, β-blockings B_π^β are not superadditive.

(1.5.5) Example. *Roulette.* An agent's i message is a pair (k_i, a_i), where k_i is a integer number from 1 to $\mid N \mid = n$ and a_i is an alternative. An outcome is determined by the formula

$$\pi((k_1, a_1), ..., (k_n, a_n)) = a_{\sum k_i \bmod n}.$$

[2] One can say that a blocking is to a mechanism what a game in characteristic form is to a game in normal form. In blockings, we focus as well on the opportunities of coalitions rather than on the concrete means of their realization.

The agents select the agent k (the "king") whose alternative will ultimately wins according to the rule $k = \Sigma k_i$ modulo n. In this mechanism, one sees readily that every agent $\beta-$forces any alternative. In effect, any agent i can practically elicit himself as the "king" provided he selects suitable a k_i, given the k_j of other agents j, $j \neq i$. Of course, this "blocking" is not superadditive.

We present now another possible construction of a "blocking".

(1.5.6) Definition. Let $F : \mathbf{L}^N \rightarrow A$ be a correspondence. A coalition K *blocks* a subset $X \subset A$ through F (denoted by KB_FX) if there exists a profile $R_K \in \mathbf{L}^K$ such that

$$F(R_K, *) \cap X = \emptyset.$$

Axiom B3 is violated here when F can have empty values. However, if F is a SCC (i.e. nonempty-valued) the relation B_F is a fully-fledged blocking and satisfies axioms B1-B3.

We give two more examples of blockings.

(1.5.7) Example. We remind that a *simple game* is a pair (N, W), where $W \subset 2^N$ is a set of "winning" coalitions, satisfying two conditions:

a) if $K \in W$ and $K \subset K'$ then $K' \in W$;

b) if $K \in W$ then $\overline{K} \notin W$. We denote by B_W a blocking associated to the simple game by the formula

$$KB_WX \Leftrightarrow \text{either } [K \in W \text{ and } X \neq A] \text{ or } [X = \emptyset].$$

In other words, winning coalitions are almighty whereas loosing coalitions are helpless. The blocking B_W satisfies axioms B1-B3 and B4 if $N \in W$.

(1.5.8) Example. Consider the following "veto" function

$$v : \{0, 1, ..., n =| N |\} \rightarrow \mathbf{Z}_+.$$

Define B_v as follows: any coalition consisting of k members blocks any set of alternatives whose size is inferior or equal to $v(k)$. B_v is a fully-fledged blocking if the two conditions are fulfilled:

1) $v(n) < |A|$,

2) if $k + k' \leq n$ then $v(k + k') \geq v(k) + v(k')$.

Veto-blockings are *anonymous* (equal with respect to agents) and *neutral* (equal with respect to alternatives). Moreover, any anonymous and neutral blocking is a veto-blocking.

We saw that blockings are interesting because they allow us to evaluate mechanisms and force of agents. Blocking are also interesting because starting from any blocking, one might construct many distinct correspondences and social choice mechanisms. Therefore in some sense blockings can be viewed as "premechanisms". We now give a few constructive examples. Let B be a blocking. An alternative a is called *individually rational* for the profile R_N if no agent i blocks the set $L(a, R_i)$. We denote by $IC(B, R_N)$ the set of all individually rational alternatives for the profile R_N.

(1.5.9) Lemma. *The set $IC(B, R_N)$ is never empty.*

This is readily seen if we consider a more explicit description of $IC(B, R_N)$. Let Z_i be the largest "lower" set, blocked by agent i, in linear order R_i. Then $IC(B, R_N)$ is the complement of $\cup_{i \in N} Z_i$ (which is not equal to A by the basic property of blockings).

The correspondence IC plays an important role in Chapter 2 when we study Nash equilibria. In Chapter 4, we investigate in details the core correspondence $C(B, .)$ associated to a blocking B.

(1.5.10) Composite Mechanisms. We show now how a mechanism can be constructed from a given blocking. However first, we have to describe a simple and useful method of constructing complex or refined mechanisms from simpler or coarser mechanisms. Assume we have the following 'coarse' mechanism

$$\sigma : S_N \to J$$

with values in the auxiliary set J. Moreover assume that for every $j \in J$, we define a terminal mechanism with values in the set A

$$\rho_j : \prod_i T_i^{(j)} \to A.$$

Then one forms *the composite mechanism*

$$\sigma * \rho : \prod_{i \in N} (S_i \times T_i) \to A$$

where $T_i = \prod_{j \in J} T_i^{(j)}$ and for $s_N \in S_N, t_N \in T_N$

$$(\sigma * \rho)(s_N, t_N) = \rho_{\sigma(s_N)}(t_N^{\sigma(s_N)}).$$

In other words, given strategies s_N, the "coarse" mechanism σ associates a number $j = \sigma(s_N)$, which selects the terminal mechanism ρ_j, which in turn selects a final outcome $\rho_j(t_N^{(j)})$.

The construction is in fact quite straightforward, even if the formal description appears cumbersome. It is essentially similar to that found in extensive game forms settings or in two-stage games settings. Often things are as follows. We first delimit some natural subset to which the desired outcome should belong. Then a final choice is made from this subset by means of an additional mechanism. This additional mechanism in turn takes often enough the form of the roulette mechanism described in Example (5.5). Incidentally, note that the roulette mechanism can be decomposed into simpler parts. First one chooses a "king"; then this king acting as a dictator chooses a final alternative. In this sense, the roulette mechanism is similar in spirit to the kingmaker mechanism.

We go back to the issue of constructing an associated mechanism from blocking B. We start with a "coarse" mechanism. Every agent i announces a subset $X_i \subset A$. Therefore the set of strategies of agent i, $S_i = 2^A$. Agents are then partitioned into groups of "similar" agents : two agents belong to the same group if they happened to send the same message. Denote these groups by respectively $K_1, ..., K_s$. The agents in K_r's message is denoted X_r. A coalition K_r is said to be *acting* if it blocks X_r. Let X be the union of X_r for all acting coalitions K_r. Finally denote by $X_\emptyset$ the maximal set blocked by the empty coalition. We pose

$$F(X_N) = A \backslash (X_* \cup X_\emptyset).$$

The basic property of blockings implies that $F(X_N)$ is a nonempty set. Clearly any terminal mechanism should take its values in the set $F(X_N)$. The blocking B_π which emerges from this process is stronger than B (in the sense $B \subset B_\pi$) whatever the terminal mechanisms considered. Indeed, assume that KBX. Now if all members of the coalition K announce X, then the reader will check that final outcome does not belong to X and therefore $KB_\pi X$.

We consider here the two following types of terminal mechanisms. The first type of terminal mechanism we look at, given a situation X_N, is the roulette mechanism with values in the set $F(X_N)$. We denote the resulting composite *canonical* mechanism by π_B. We claim that the blocking B_{π_B} then essentially coincides with the initial blocking B. More exactly, if a coalition K is different from N and $K\overline{B}X$ then $K\overline{B_{\pi_B}}X$. In fact, given any action of coalition K, all members of the complementary coalition $\overline{K}$ may send the message: "empty set" and then choose an outcome from X, provided they selected one of them as a "king". As regarding the opportunities open to coalition N, it is easy to see that it can select any element from $A^* = A \backslash X_\emptyset$ as an outcome. We shape this short proof further.

We shall say that a mechanism π *generates* a blocking B if $B = B_\pi$. In the sequel, one of the typical issues we shall address is whether a given blocking is generated by a mechanism possessing prescribed additional properties such as consistency, coalitional stability and so on. We shall elaborate on this mainly in Chapters 2 and 5. For the time being, we only ask whether there exists any mechanism such that it generates a given blocking? Obviously any blocking of the type B_π satisfies the following property:

B3*. *Coalition N forces any set which is not blocked by the empty coalition.*

Note that B3* is stronger than B3. It turns out the inverse is also true.

(1.5.11) Proposition. *A blocking B is generated by a mechanism if and only if it satisfies property* B3*.

Indeed, as proved above, if B satisfies B3* then it is generated by its canonical mechanism π_B.■

We turn now to the second type of terminal mechanisms. We pick an agent i_0 and we define the terminal mechanism, given a situation X_N, as dictatorship of the agent i_0 on the set $F(X_N)$. We denote by B' the blocking generated by this mechanism. We previously remarked that $B \subset B'$. B' may be described explicitly:

$$KB'X \text{ if either } KBX, \text{ or } \overline{KBX} \text{ and } i_0 \in K.$$

The blocking B' satisfies a maximality property which we shall encounter quite often in the sequel.

(1.5.12) Definition. A blocking is called *maximal* if it is a maximal element in the set of all blockings ordered by inclusion.

In other words, if a blocking is maximal then it is not possible to strengthen any coalition without violating any one of the properties B1-B3. This definition explains the word 'maximal', however it is not very convenient to use and to work with. The following simple maximality criterion can be used in place of it.

(1.5.13) Proposition. *A blocking B is maximal if and only if the following property holds: $K\overline{B}X \Rightarrow \overline{K}B\overline{X}$.*

Proof. We show that this property implies maximality. Let $B \subset B'$ and $KB'X$. Then $\overline{KB'X}$ hence $\overline{KBX}$ and therefore KBX. So $B = B'$.

Converse. Let B be maximal. Assume $K\overline{B}X$. We show that $\overline{K}B\overline{X}$. To do so, we pick an agent i_0 outside coalition K and define B' to be the blocking constructed above. By maximality of B, the equality $B = B'$ obtains. Moreover the explicit description of B' implies that $\overline{K}B'\overline{X}$, whence $\overline{K}B\overline{X}$.■

(1.5.14) Corollary. *Every maximal blocking is generated by a mechanism.*

1.A1 Arrow's Impossibility Theorem

(1.A1.1) Preference aggregation. Classical social choice theory focused upon the preference aggregation issue, which can be considered as an intermediary stage of social choice. We devote little space to this issue, since it is addressed everywhere in the literature (Arrow (1951), Sen (1970, 1986), Fishburn (1973), Mirkin (1974), Danilov (1983)). Aggregating preferences consists, given a set of alternatives A, in constructing the preference of a social group $R = \Phi(R_N)$ based on individual preferences (and more precisely on a preference profile R_N). With such a preference R it is then easy to solve any social choice problem in A.

We noted that the constructing of many rules of social choice involves the specifying of a group utility. For example, take the Borda rule. In this rule, we constructed the number $u(x, R_N)$, and this number can be regarded as a group utility.

Therefore aggregating preferences amounts to dealing with the following kind of mappings

$$\Phi : \mathbf{L}(A)^N \to \mathbf{L}(A).$$

Of course, this is only one way to do it. In fact, one can take, instead of $\mathbf{L}$, any other class of binary relations (weak orders, transitive relations, tournaments and so on). We shan't give here an exhaustive overview of this domain of social choice theory (the interested reader should refer to surveys). However, we shall consider a typical example, in order to give the reader a flavor of what are the issues involved, the possible answers and the reasonings. We assume here that $\mid A \mid \geq 3$.

In order to be a "reasonable" rule of preference aggregation, the mapping Φ should satisfy some requirements. Following Arrow, we impose two conditions. The first (Unanimity or Pareto Condition): if preferences R_i of all agents are the same and equal R then $\Phi(R_N) = R$, is practically indisputable. The second (Independence of Irrelevant Alternatives) goes as follows: a group preference $\Phi(R_N)$ on any pair of alternatives $\{x, y\}$ depends only on the preferences of agents R_i on the same pair of alternatives. One can either agree or disagree with this requirement (for example, the Borda rule does not satisfy it). However it sounds quite innocuous. Therefore it is surprising that only dictatorial rules of preference aggregation (i.e. $\Phi(R_N) = R_{i_0}$ for a given agent i_0) satisfy both these conditions.

This is the content of the well-known Impossibility (or Possibility) theorem by Arrow (1951), a milestone of modern social choice theory. Its proof is similar to that of the Mueller and Satterthwaite theorem (1.3.6). We sketch two alternatives ways of proving Arrow's theorem. The first way is direct.

(1.A1.2) We denote by $W(x, y)$ the set of coalitions K such that $x \succ_K y$ (i.e. $x \succ_i y \ \forall i \in K$) implies $x \succ y$ (where $\succ = \Phi((\succ_i)_{i \in N})$. For example, $N \in W(x, y)$ by unanimity and IIA. Let $W = \cup_{x,y} W(x, y)$ and C be a minimal coalition in W belonging to some $W(x, y)$. We claim that C consists of a unique agent. Indeed, let $i \in C$ and z be an alternative different from x and y. We consider the following profile R_N (compare with profile R_N in the proof of lemma (1.3.13):

x	z	y
y	x	z
z	y	x
$*$	$*$	$*$
i	$C \setminus i$	$\overline{C}$

Since $x \succ_C y$ and $C \in W(x,y)$ then $x \succ y$. Since $z \succ y$ only for the participants of coalition $C \backslash \{i\}$ and $C \setminus \{i\} \notin W(z,y)$ then relation $z \succ y$ is not fulfilled, hence $y \succ z$ is true. By transitivity, $x \succ z$. However, as $x \succ z$ holds only for agent i, $\{i\} \in W(x,z)$ for any alternative $z \neq x$.

We need to show that agent i is decisive on any pair. Take two alternatives y and z , both different from x, and consider a profile Q_N:

y	z
x	y
z	x
$*$	$*$
i	$N \setminus i$

Then for $\succ = \Phi(R_N)$ we have $x \succ z$ (since $\{i\}$ belongs to $W(x,z)$ and $y \succ x$ by unanimity). So $y \succ z, \{i\} \in W(y,z)$ and i is a dictator.■

(1.A1.3) The second proof rests on a stronger monotonicity condition than the IIA condition, albeit the former is derived from both IIA and unanimity (see for example Danilov (1983)). The proof goes as follows. Let R_N and R'_N be two profiles, x and y be two alternatives. The monotonicity of Φ means that if $x\Phi(R_N)y$ and if, for any $i \in N$, xR_iy implies xR'_iy, then $x\Phi(R'_N)y$. Under this condition the proof boils down to Theorem (3.6). For this purpose, we define the SCF f by the formula $f(R_N) = \max \Phi(R_N)$. By virtue of the unanimity condition, f is sovereign, and by monotonicity of Φ, the SCF f is monotone as well. Thus, there exists a dictator, i.e. an agent i_0 such that $f(R_N) = \max R_{i_0}$. It is easy to understand that $\Phi(R_N) = R_{i_0}$, i.e. i_0 is an arrovian dictator.

1.A2 Non-manipulable SCFs

(1.A2.1) We pointed to the issue of truthful preference revelation in the Introduction while we dwelled upon possible different desirable properties of SCFs. The question is now will agents reveal truthfully their preferences? It is difficult to say what agents will do. And therefore what we should ask is whether it is profitable for them to reveal truthfully their preferences, whether a SCF will cause them to lie, to distort their preferences, to manipulate them in an attempt to get a better outcome? To start with, we present a few concrete examples.

(1.A2.2) Example. Let there be a unique agent and a SCF determined as follows: $f(R) = \min R$ (anti-dictatorship).

Clearly if the agent knows the principle of choice, he will declare instead of his truthful preference R the opposite preference R^{-1}. The rule is then consciously distorted and perverted.

Let us consider a more "reasonable" rule.

(1.A2.3) Example. There are three agents and three alternatives, $A = \{x, y, z\}$. The rule acts as follows: we compare first alternatives x, y and determine a winning alternative which is then compared with z. We consider the following profile R_N :

x	y	z
y	z	x
z	x	y
1	2	3

In the first stage x wins over y, in the second stage z wins over x, thus $f(R_N) = z$. However agent 1 might notice that if he shoves an allegated preference $R_1' = (y \succ x \succ z)$ instead of his true $R_{1=}(x \succ y \succ z)$, then the outcome $f(R_1', R_2, R_3) = y$ will be better for him than z. Why would he then reveal R_1 and therefore harm himself? Moralists may tell him that it is not good to deceive in general. But then our agent might object to this (as in Example (1.A2.2) that the social outcome is determined through a bad SCF and that it would be better indeed to spend some more time and efforts to devise a more reasonable collective choice mechanism, which in turn would not cause him to lie.

(1.A2.4) Example. Again there are three agents and three alternatives. The winning alternative here is that which has been listed at the top of rankings by a majority of agents. Assume that there is no such winning alternative, then the outcome is determined by the first agent's preferred alternative. Take again the profile R_N figuring in Example (1.A2.3). Here $f(R_N) = x$. However agent 2 replacing R_2 by $R_2' = (z \succ y \succ x)$ may obtain outcome z, which he prefers to x.

Do these cheating attempts occur just by chance? Are we able to exhibit SCFs which do not have the drawbacks? Before discussing these issues, we give the following definition.

(1.A2.5) Definition. A SCF $f : \mathbf{L}^N \to A$ is *non-manipulable* (or *strategy-proof*, i.e. immune to strategic behavior) if for any profile $R_N \in \mathbf{L}^N$, any agent i, and any order $R_i' \in \mathbf{L}$ the following relations hold,

$$f(R_N) R_i f(R_i', R_{N \setminus i}).$$

In other words, these relations state that truth is the most profitable message for every agent. More exactly, no agent i ever profits from any deviation from his truthful preference R_i. At least two reasons explain the desirability of non-manipulability. First non-manipulability simplifies the life of agents, they need not torture themselves in order to come up with a best message.

The best message is obvious (it is the truthful message) and moreover independent of the messages of other agents. Second non-manipulability ensures the social planner that the chosen SCF yields the outcome $f(R_N)$ when the profile is R_N.

However the mere formulating of a desirable requirement, albeit important, just does half of the business. The main issue remains: do non-manipulable SCFs exist, how many of them are there and can we describe them? The answer is unfortunately not quite satisfactory- in fact, there are very few of them. We establish the following fact.

(1.A2.6) Lemma. *A SCF is non-manipulable if and only if it is monotone.*

Proof. Let f be non-manipulable and R_N, R'_N be two profiles. Assume that $f(R_N) = a$ and $R_N \preceq_a R'_N$. We have to show that $f(R'_N) = a$. Without loss of generality, we can assume that $R_j = R'_j$ for $j \neq i$ so that only agent i modified his preference. Pose $b = f(R'_N)$. By non-manipulability, aR_ib and bR'_ia. Since $R_i \preceq_a R'_i$, aR_ib implies aR'_ib. Together with bR'_ia this yields $a = b$.

Conversely, let f be a monotone SCF, we have to check that

$$f(R_i, R_{N\setminus i})R_i f(R'_i, R_{N\setminus i}).$$

Since profile $R_{N\setminus i}$ will not play any role here, we can assume it is fixed and not write it as an argument. Let $f(R_i) = a$, $f(R'_i) = b$. Assume that $a \neq b$ and bR_ia. We construct an auxiliary order $R = (b \succ a \succ ...)$. Since $R_i \preceq_a R$ then $f(R_N) = a$, since $R'_i \preceq_b R$ then $(R_N) = b$, which is in contradiction with $a \neq b$. ■

The following result proved by Gibbard (1973) and Satterthwaite (1975) comes out as a consequence of both the previous Lemma and the Mueller-Satterthwaite theorem. Barbera (1983) devised another proof of this theorem; we sketch it in Chapter 3.

(1.A2.7) Theorem. *Let $f : \mathbf{L}^N \to A$ be a sovereign non-manipulable SCF and $|A| \geq 3$. Then f is dictatorial.*

Note that Gibbard proved a somewhat more general result, which follows as well from similar considerations. Let $f : \mathbf{L}^N \to A$ be a non-manipulable SCF. Then either f is duple (i.e. the image of f contains at most two alternatives) or f is unilateral (i.e. $f(R_N)$ is determined solely by the preference R_i of dictator i).

This is one of the principal results of social choice theory. If we use a non-dictatorial SCF and if the real choice involves more than two alternatives, then we inevitably come across a preference profile R_N and an agent i which profits from distorting his preference R_i. However one shouldn't dramatize: few such profiles may occur (though we do not know exact assertions of such type) and the allegedly profiting agent i might not necessarily distort his preferences. Thus this somehow negative result does hamper the practice of social choice and voting procedures. However from a theoretical point of view, the matter is slightly more serious. This result implies that the requirements on SCF turned out to be too strict and somewhat contradictory. Non-dictatorial and non-manipulable SCFs exist no more than perpetuum mobile. Therefore, we should relax somewhat our requirements and look for means of retreat. The interest of this theorem is that its conditions indicate the directions of retreat. We shall discuss three of these directions.

A first simple direction consist in restricting the set of alternatives such as $| A |= 2$. In this case there is quite sufficient a number of "nice" (for example, almost anonymous) non-manipulable SCFs (see the case of two alternatives (1.3.10)).

A second direction consists in restricting the environment. Indeed, for instance in the single-peaked environment, one can find many interesting non-manipulable SCFs. Section 3.2 will elaborate more thoroughly on these issues.

The third direction consists in changing the solution concept, in accommodating with manipulability provided we compensate manipulability by some other 'nice' properties. The fact that an agent might lie is not that much of a problem. But the fact that he might not have a dominant strategy is more of a problem, because then we don't know what he would do, neither do we know what other agents would do, nor do we know the outcome. The only consistent decision in this case is to complement a SCF or, for the matter an arbitrary social choice mechanism (it isn't necessary to restrict ourselves solely to direct mechanisms), with a suitable behavioral or equilibrium concept. For instance, we might take the Nash equilibrium and we explore this in the next chapter. An interesting and prospective direction is that proposed by B.Peleg (1978) which considers the stronger concept of coalitional equilibrium. We explore it in Chapter 5.

1.A3 Minimal Monotone SCCs

We saw in Section 1.3 that monotone SCCs might happen to be multi-valued for some profiles. To what extent can one minimize such a multi-valuedness? What is an "almost single-valued" SCC? These questions lead us to the following notion.

(1.A3.1) Definition. A monotone SCC F is called *minimal* (by inclusion), or simply minimal, or a MMSCC, if every monotone nonempty-valued sub-correspondence $G \subset F$ coincides with F.

In other words, a sub-correspondence $G \subset F$ is either non-monotone or $G(R_N) = \emptyset$ for some profile R_N. Of course, every monotone SCF is a MMSSC but there are other MMSCCs. We establish now a criterion of minimality due to Moulin (1983) and then we give a few examples of MMSCCs.

(1.A3.2) Proposition. *Let F be a monotone SCC. The following assertions are equivalent:*

1. F is a minimal monotone SCC.

2. For any $x \in F(R_N)$ there exists a profile $R'_N \preceq_x R_N$ such that $F(R'_N) = \{x\}$.

One may say that F is minimal if it is single-valued for sufficiently many profiles. We now prove the proposition.

2. $\Rightarrow$ 1. Let $G \subset F$ be a monotone sub-correspondence of F and $x \in F(R_N)$. Pick a profile R'_N satisfying 2. Then by monotonicity of G, $x \in G(R_N)$, thus $G = F$.

1. $\Rightarrow$ 2. Assume a profile R^*_N and an alternative $x \in F(R^*_N)$, satisfying

$$R_N \preceq_x R^*_N \Rightarrow F(R_N) \neq \{x\}\,.$$

We construct then a new correspondence G as follows :

$$G(R_N) = \begin{cases} F(R_N) \setminus \{x\} & \text{if } R_N \preceq_x R^*_N, \\ F(R_N) & \text{otherwise.} \end{cases}$$

One can see readily that $G \subset F$, is non-empty valued and monotone: this is in contradiction with the minimality property of F.■

Consider the following examples of MMSCCs.

(1.A3.3) Example. *Maskin correspondence,* see Example (1.3.4). The Maskin correspondence is monotone; it is straightforward with the above-given criterion. Let $x \in F(R_N)$, i.e. x is a Pareto optimal alternative and xR_ia, for all $i \in N$. If $x = a$, then $F(R_N) = \{x\}$. If $x \neq a$, then prop a up in order to bring it just below x and leave all other alternatives' positions unchanged. The new profile R'_N is x-equivalent to R_N and $F(R'_N) = \{x\}$.■

Moulin (1983), in his book, claims that for a given MMSCC F and $x \in F(R_N)$ there exists a profile R'_N which is $x-$equivalent to R_N and $F(R'_N) = \{x\}$. The example below shows that it is not quite true.

(1.A3.4) Example. *A modified Borda rule.* Let $N = \{1,2,3\}$, $A = \{a,b,c\}$. We consider the following monotone variant of Borda rule $\tilde{F}$. We value alternatives depending on their positions in the ranking of a definite agent and this for each agent. An alternative figuring at the bottom receives 0 points, 1 point if it stands in the middle and 2 points if it figures at the top of the ranking. Moreover agents are weighted : the first and the second agent are given weight 2, whereas the third's weight is 1. Thus for the profile R_N :

a	b	c
b	a	b
c	c	a
1	2	3

the value of alternative a is $2{\times}2 + 2 \times 1 + 1 \times 0 = 6$ points, of b is $2{\times}1 + 2 \times 2 + 1 \times 1 = 7$ points and of c is $2{\times}0 + 2 \times 0 + 1 \times 2 = 2$ points. We pose

$$\tilde{F}(R_N) = \{\text{the set of alternatives whose number of points is } \geq 6\}.$$

For instance, here $\tilde{F}(R_N) = \{a, b\}$. The correspondence $\tilde{F}$ is non-empty valued. Indeed,

1) all alternatives taken together collect 15 points;

2) one only of the alternatives collects an odd number of points: if it figures in second position in the third agent's ranking.

Obviously $\tilde{F}$ is monotone. We modify the correspondence $\tilde{F}$ as follows,

$$F(R_N) = \begin{cases} \tilde{F}(R_N) & \text{if 1st and 2nd agents' top-alternatives, are different} \\ \{x\} & \text{if 1st and 2nd agents' top-alternative is equal to } x.\ \# \end{cases}$$

In the # case, the profiles look as follow,

x	x	$*$
$*$	$*$	$*$
1	2	3

We check now F's monotonicity. Assume for instance that $a \in F(R_N)$ (thus a collects a number of points ≥ 6 in R_N). Let $R_N \preceq_a R'_N$ (thus a collects at least 6 points in R'_N as well) and assume that $a \notin F(R'_N)$. This happens only if R'_N has the form

x	x	a
a	a	$*$
$*$	$*$	$*$
1	2	3

In this case, if we consider the previous profile R_N (such that $R_N \preceq_a R'_N$), then we cannot bring a into $F(R_N)$, which contradicts the assumption. Thus F is monotone.

To check minimality of this SCC one uses the minimality criterion. The reader will sort out himself all the adequate profiles (there are essentially four different such profiles) and check the criterion. We now give an idea of the argument of the proof. Take the following profile R_N :

a	b	a
b	c	b
c	a	c
1	2	3

Here the set of outcomes $F(R_N) = \{a, b\}$ since a and b collect 6 points each. In order to reduce it to a, it is sufficient to select a $R'_N \approx_a R_N$ where we exchange b and c's positions in the first agent's ranking. Then $F(R'_N) = \{a\}$. Now to reduce that set to b, we take $R'_N \preceq_b R_N$ where the third agent's ranking is modified as follows $c \succ a \succ b$. Then alternatives a, b, c, collect respectively 5, 6 , and 4 points, thus $F(R'_N) = \{b\}$. We see here that without lowering alternative b, it is impossible to expel a, in the sense that there is no profile $R'_N \approx_b R_N$ for which $F(R'_N) = \{b\}$.

Propositions (4.2.7) and (4.2.10) of Sec. 4.2, Ch.4 present other examples of MMSCCs constructed from given blockings. We shall emphasize the interest of this concept of MMSCCs in next chapter when we consider implementation issues.

Bibliographic Comments

The material presented in Sections 1.1 and 1.2 is rather classical. Weak and linear orders are fairly common concepts which are presented in many standard textbooks. We recommend in particular Fishburn (1978), Kiruta et al. (1980), and Mirkin (1973), Kreps (1988). These books also discuss relations between preference and utility function. The approach to preferences through rational choice functions goes back to Arrow's seminal work (1951), further developments can be found in Aizerman, Malishevsky (1981) or Kreps (1988).

The origins of SCCs are difficult to trace, but they figure implicitly in the early works in social choice theory. However Arrow's book launched a series of investigations on preference aggregation rules (for historical details see Lesina (1987), Sen (1970)). Incidentally the preference aggregation issue and the SCC issue are closely related; Fishburn (1973) and Lesina (1987), for instance, provide numerous examples of classical SCCs. Gibbard in his 1973 article makes an explicit use of the notion of SCF. He also introduces in an explicit way the notion of a mechanism (albeit this concept had been around for some time), denoted a game form, and proves the theorem about manipulability. 1973 might therefore be marked as the year of birth of social choice mechanisms theory. We recommend also a survey by Groves (1979).

The notion of monotonicity (under this or another denomination or variations) appeared somewhat long ago (see, for example, Polterovich (1973)). Gradually it became one of the central notions of social choice theory (especially after Maskin's (1979) seminal article) and Moulin (1983) or Peleg (1984) devote to it an important space in their books. The Mueller-Satterthwaite theorem (3.6) derives historically from the Gibbard-Satterthwaite theorem (about manipulability, which in turn is based on Arrow theorem. Our presentation, in contrast, treats the Mueller-Satterthwaite theorem as the central piece and derive all other theorems from it. The results on Nash-implementability (Chapter 2) emphasize the importance of monotonicity with respect to SCCs.

The concept of forcing (or of effectivity function) was introduced in the fundamental paper by Moulin, Peleg (1982). Incidentally, both this notion and the notion of maximal blocking appear in Gurvich (1975) albeit under other names. We shall discuss blockings and their applications more extensively in Chapters 2, 4 and 5.

2. Nash-consistent Mechanisms

This chapter is devoted to Nash-consistent mechanisms, that is mechanisms possessing Nash equilibria at every preference profile. In Section 2.1, we examine a few examples, then proceed to investigate blockings generated by Nash-consistent mechanisms (Section 2.2). In Section 2.3, we show that the correspondence of equilibrium outcomes exhibit a somewhat stronger property than monotonicity, which is called strong monotonicity. In Section 2.4, we describe Nash-implementable SCCs. In the more-than-two-agents case, the class of Nash-implementable SCCs coincides with the class of strongly monotone SCCs. The case of two agents is considered in Section 2.5. In Section 2.6, we discuss acceptable mechanisms, that is consistent mechanisms whose outcomes are Pareto optimal.

In Appendix 2.A we give a simple mechanism implementing Walrasian equilibria.

2.1 Definitions and Examples

(2.1.1) This chapter is devoted to an application of the Nash equilibrium concept (the most used in game theory) to social choice mechanisms.

Let $\pi : \prod_{i \in N} S_i \to A$ be a mechanism and R_N be a preference profile. A strategy profile $s_N = (s_i, i \in N)$ is a *Nash equilibrium* for the game $G(\pi, R_N)$ if it satisfies the following individual optimality property: for each participant $i \in N$ and any strategy $s_i \in S_i$ of this participant,

$$\pi(s_N^*) R_i \pi(s_i, s_{N-i}^*),$$

where s_{N-i}^* denotes the strategy profile $(s_j^*, j \neq i)$. The preceding relation can be rephrased as

$$\pi(s_i, s_{N-i}^*) \in L(\pi(s_N^*), R_i).$$

The set of all Nash equilibria of a game $G(\pi, R_N)$ is denoted $NE(\pi, R_N)$. The mechanism π is *consistent* (more precisely, *Nash-consistent*) if $NE(\pi, R_N)$ is non-empty for any preference profile $R_N \in \mathbf{L}^N$. For brevity, we shall simply say equilibria when we mean Nash equilibria in this chapter.

We start by a short discussion of the notion of (Nash) equilibrium. It is based on two behavioural postulates. First it is assumed that if an avalaible strategy s_i to a participant i improves upon the outcome with respect to using a strategy s_i^* (for a fixed s_{N-i}^*), then the strategy s_i^* will not be used. This assumption is reasonable as long as participant i can count on the non-reaction of other participants to his deviating. However, in many cases it turns out that this kind of expectation is both myopic and thoughtless. Second it is assumed that if no deviation from the strategy s_i^* improves upon the outcome for a participant i, then he will use s_i^* even if that strategy might seem quite absurd. We shall frequently encounter such situations. Notwithstanding these drawbacks of the Nash equilibrium concept, we use this concept and investigate its consequences for mechanisms. We begin with a few concrete examples of mechanisms.

(2.1.2) Example. *Roulette.* This mechanism was defined in Example (1.5.5). The only important thing we are interested here is the following property of the roulette mechanism: for any $i \in N$, $a \in A$ and $s_{N-i} \in S_{N-i}$ there exists $s_i \in S_i$ such that $\pi(s_i, s_{N-i}) = a$. It is clear that equilibria do not exist for generic preference profiles. More precisely, equilibria exist only for preference profiles R_N, characterized by the fact that all participants have the same best alternative, i.e. $\max R_i = \max R_j$ for all i, j (in other words when there is unanimity).

Here the absence of equilibria is due to very large, albeit illusory, enforcing possibilities of participants. In any situation (that is for any s_{N-i}) a participant i can enforce his best alternative in the outcome. And he will undertake this knowing that some other participant might immediately ruin his attempt at improving the outcome. In the next section, we develop the idea that consistent mechanisms should rather strongly restrict the enforcing possibilities of participants.

In what follows, we use roulette mechanisms in order to suppress undesirable equilibria.

(2.1.3) Example. *Kingmaker.* This mechanism was mentioned in (1.4.2). Recall that $S_1 = N - \{1\}, S_j = A$ for $j \neq i$, and

$$\pi(s_1, ..., s_n) = s_{s_1}.$$

This simple mechanism possesses interesting properties. To begin with, it is consistent. In effect, let R_N be a preference profile. Note that each participant $j \neq 1$ has a dominant strategy, i.e., a strategy which is optimal for any strategies of other participants, namely $s_j^* = \max R_j$. The kingmaker does not have a dominant strategy; his best response s_1^* depends on alternatives chosen by the other participants. Namely, he selects the participant who chose the best (among $s_2, ..., s_n$) alternative with respect to R_1. Clearly, this strategy profile is an equilibrium, and the corresponding outcome is Pareto optimal.

Note that, apart from this "natural" equilibirum, there might be other equilibria. Nevertheless (and this is another positive property of this mechanism), any equilibrium outcome is Pareto optimal, because it is the best alternative for the "king".

Therefore every equilibrium outcome lies in the set $\{\max R_2, \ldots, \max R_n\}$. We might think that it is equal to $\max R_1 \mid \{\max R_2, \ldots, \max R_n\}$. However, this is not true; in effect, any element from the set $\{\max R_2, \ldots, \max R_n\}$ can be an equilibrium outcome at the preference profile R_N. For example, let us show how to enforce alternative $a = \max R_2$ as the equilibrium outcome. Suppose that all remaining participants name alternative a, and participant 1 crowns participant 2 as a "king". It is easy see that a is an equilibrium outcome. However, if $\max R_j \neq a$ then naming alternative a would be a strange behaviour for participant j. For example, if $(\max R_j) R_1 a$, then participant j, by choosing $s_j = \max R_j$, and participant 1, making j the king, would be better off. But this requires a certain coordination in the actions of participants 1 and j, for instance uniting them in a coalition. This line of reasoning shows, once more, that the Nash equilibrium concept requires some refining (see Chapters 3 and 5).

(2.1.4) Example. We present here a variant of the preceding mechanism. We consider only three participants for simplicity. Let $S_1 = \mathbf{L}$, $S_2 = S_3 = A$, and

$$\pi(R, a_2, a_3) = \max(R \mid \{a_2, a_3\}).$$

Participants 2 and 3, as in Example (2.1.3), name alternatives whereas participant 1, this time, chooses an alternative out of these two (in Example (2.1.3) he chose one of the participants). We might think that this is not very much of a difference, but this is wrong.

This new mechanism is also consistent. Participant 1 has now a dominant strategy - to call his true preference R_1. And if $s_2^* = s_3^* = \max R_1$, then this is an equilibrium. However it is quite difficult to believe that participants 2 and 3 will name precisely $\max R_1$. We can expect they will name their best alternatives $a_i = \max R_i$, $i = 2, 3$. But their behaving this way may not be an equilibrium. Assume the following preference profile R_N,

x	u	z
y	y	u
z	z	x
u	x	y
1	2	3

and $s_1^* = R_1$. If participant 2 calls u, and participant 3 calls z, then the outcome is equal to z. Participant 2 can improve upon the result if he names y instead of u. Thereafter, participant 3 might also switch from z to x. Now the strategy profile (R_1, y, x) is an equilibrium with outcome x. x is efficient, but for the coalition $\{2,3\}$ this equilibrium is worse than z.

If participant 1 uses his dominant strategy, ie., names his true preference, then any equilibrium outcome is Pareto optimal. But in general, an equilibrium outcome can turn out non-optimal. Let us consider the following preference profile,

x	x	x
y	y	y
z	z	z
1	2	3

,

and strategies $s_1^* = (z \succ y \succ x)$, $s_2^* = s_3^* = y$. This is an equilibrium with non-optimal outcome y. However, the behaviour of participants in this case seems implausible.

(2.1.5) Example. *The Simple Majority.* We consider only three participants to make matters simple. Each participant names an alternative; the alternative which collects the most votes is the winning alternative. If all three participants name different alternatives, then the outcome is the alternative named by participant 1.

Here participant 1 has the following dominant strategy which consists in naming his best alternative, i.e. $\max R_1$. Pose $a = \max R_1$; this alternative a serves as the starting-point for participants 2 and 3 (compare this with the "status quo" point appearing in examples (1.3.3) and (1.3.4) of Chapter 1). Given fixed preferences R_2 and R_3, we define the set

$$U = U(a) = \{x \in A,\ xR_2a \text{ and } xR_3a\}.$$

Then any element $x \in U$ can be an equilibrium outcome: it suffices that participants 2 and 3 name x.

However, these "reasonable" equilibria are not the only ones to exist. In the spirit of the preceding example, we should expect some "foolish" equilibria to appear. And there are some: any element $x \in A$ is a possible outcome in a "foolish" equilibrium, that is when all three participants name x. We shall consider the equilibrium outcome issue more in detail in Sections 2.3-2.5.

(2.1.6) Example. This is a variant of the preceding example. Again there are three participants. In the first step, they elect a "king". This king then (second step) names an outcome. The king is elected by majority rule. If each participant collects one voice, then the roulette mechanism is used (see Example (2.1.2)).

This mechanism is consistent and implements the following SCC:

$$R_N \longmapsto \{\max R_1, \max R_2, \max R_3\},$$

that is the union of three dictatorial SCFs. We now show how to implement $x = \max R_1$ as an equilibrium outcome. To this end all participants should first elect unanimously the king to be participant 1, and then participant 1 should name the outcome x. Obviously any alternative, different from either $\max R_i$, for $i = 1, 2, 3$, is not an equilibrium outcome.

Thus all equilibrium outcomes of this mechanism are efficient, albeit no single one of them seems very plausible.

The moral of this story is that often enough we attain consistent mechanisms through somewhat "foolish" or "deadlock" equilibria from which no participant may hope to escape by some individual action. As a rule, this kind of phenomenon occurs as soon as the number of participants is higher than two. When there are only two participants, consistent mechanisms also exist although they are more difficult to construct.

(2.1.7) Example. There are two participants and three (for simplicity) alternatives. Each participant is to reject one of the alternatives; the outcome consists of a non-rejected alternative. Formally, we should give a tie-breaking rule to account for the case when the both participants reject the same alternative. Usually this tie-breaking rule does not play too big a role. For instance here, we take the following auxiliary order $R_0 = (x \succ y \succ z)$ and pose that

$$\pi(a_1, a_2) = \max R_0 | (A \setminus \{a_1, a_2\}),$$

where a_i figures the alternative rejected by participant i. We assert that this mechanism is consistent. To prove it we consider two cases. First case: $\min R_1 \neq \min R_2$. Here an equilibrium consists of the "natural" strategies $a_i = \min R_i,\ i = 1, 2$. Second case: $\min R_1 = \min R_2 = a$. Here one equilibrium is: $a_1 = a,\ a_2 = \min R_2 \mid (A \setminus \{a\})$.

In this example again, we get some foolish equilibria. For example, take the following preference profile,

$$\left| \begin{array}{c|c} x & x \\ y & y \\ z & z \end{array} \right|,$$

and let both participants reject x. This is an equilibrium: its outcome is y.

This mechanism implements (see Section 3) the individual core correspondence (see Section 1.5 of Chapter 1),

$$F(R_1, R_2) = A \setminus \{\min R_1, \min R_2\}$$

with one exception, namely for the following preference profile

$$\left| \begin{array}{c|c} z & z \\ y & y \\ x & x \end{array} \right|.$$

Here the only equilibrium outcome is z.

(2.1.8) Example. This is again a variant of the preceding example. Participant 1 rejects one alternative, while participant 2 chooses within the remaining alternatives. More formally: $S_1 = A$, $S_2 = \mathbf{L}$ and

$$\pi(a, R) = \max R \mid (A \setminus \{a\}).$$

This mechanism is also consistent (by the same line of reasoning as in Example (2.1.7)). It implements exactly the individual core correspondence.

With this, we end our list of examples. And what conclusions do we draw? First, any reasonably constructed mechanism has equilibria. Second, often enough equilibria are either foolish or implausible. In particular, they might be non-optimal. Notable exceptions to this rule are the kingmaker mechanism and the mechanism presented in Example (2.1.6).

The fact that we find many "foolish" equilibria among Nash equilibria suggests that one should focus only on those which are "reasonable". Game theory elaborated a number of refinements and reinforcements of the Nash-equilibrium concept. In what follows, we shall dwell upon two particular reinforcements of Nash-equilibria. The first one is related to the use of dominant strategies, because as we saw in the preceding examples, many absurd equilibria result from some participant's rejecting his dominant strategy, though the latter exists. The second Nash-reinforcement permits coordinated actions of participants. These issues will be elaborated further in Chapters 3 and 5.

2.2 Blockings Generated by Consistent Mechanisms

(2.2.1) We have already noticed in Example (2.1.2), that the existence of Nash equilibria restricts participants' power (in the sense of β-blocking, see Section 1.5). We discuss now more in detail the issue of blockings generated by consistent mechanisms.

Let us begin with the following simple remark. Assume that the strategy profile s^*_N is an equilibrium for a game $G(\pi, R_N)$. Then for every $i \in N$, the coalition $N - \{i\}$ forces the set $L(\pi(s^*_N), R_i)$ through s^*_{N-i}. To obtain more specific assertions about the power of coalition $N - \{i\}$ one needs to apply this argument for some well chosen preference profiles. We examine here only one instance.

Let $A = X_1 \cup ... \cup X_n$ be a covering of the set A by sets (X_i), $i \in N$. Suppose that the preference R_i has the following form $(\overline{X_i} \succ X_i)$, where the "bar" designates the complementary subset. Assume the mechanism π is consistent, then there exists an equilibrium outcome $a \in A$ for the game $G(\pi, R_N)$. This outcome a belongs to some set X_i, therefore the coalition $N - \{i\}$ forces X_i. Thus we have proved the following proposition.

(2.2.2) Proposition. *Let π be a consistent mechanism. Then for any covering $A = X_1 \cup ... \cup X_n$, there exists a participant i such that the coalition $N - \{i\}$ blocks the set $\overline{X}_i = A - X_i$.*

In order to rephrase this result, we introduce the following definition. An alternative a is a *pre-equilibrium* for a game $G(\pi, R_N)$ if no participant i β-blocks the set $L(a, R_i)$ (that is $a \in IC(B_\pi^\beta, R_N)$, in the terms defined in Section 1.5 of Chapter 1). It is clear that an equilibrium outcome is a pre-equilibrium. Therefore, if a mechanism π is consistent, then the correspondence $IC(B_\pi^\beta, \cdot)$ is never empty-valued; this is another formulation of Proposition (2.2.2).

It is unlikely that the converse of Proposition (2.2.2) will be true. However, we can say something about some reverse implication. The matter is simplest in a two participant case, where the consistency of a mechanism is determined in terms of the blocking B_π.

(2.2.3) Theorem (Gurvich)**.** *A mechanism π, involving two participants, is consistent if and only if the blocking B_π is maximal.*

Proof. The maximality of the blocking B_π was proved in Proposition 2.2.1. The main difficulty consists in proving the opposite assertion: suppose the blocking B_π is maximal, then the mechanism π is consistent. To this end, we should be able to exhibit a Nash equilibrium for each game $G(\pi, R_N)$ constructed with a preference profile $R_N = (R_1, R_2)$. We divide the proof in two parts.

I. Suppose first, that we have an alternative a and two sets X_1,$X_2 \subset A$ satisfying the following three conditions:

a) $X_1 \cap X_2 = \{a\}$,

b) $X_i \subset L(a, R_i)$, $i = 1, 2$,

c) participant i does not block X_i, $i = 1, 2$.

Then there exists an equilibrium. Following c) we know that participant 1 does not block X_1. Due to the maximality of B_π, participant 2 forces X_1. Let s_2^* be a forcing strategy for participant 2; that is a strategy driving the outcome in X_1. Similarly, let s_1^* be a forcing strategy for participant 1, i.e., driving the outcome in X_2. We assert that the strategy profile (s_1^*,s_2^*) is a Nash-equilibrium. First of all, $\pi(s_1^*,s_2^*) \in X_1 \cap X_2$ and, in accordance with a), $\pi(s_1^*, s_2^*) = a$. Further, $\pi(\cdot, s_2^*) \in X_1$ and, in accordance with b) participant 1 can not improve on the outcome. This is also true for what concerns participant 2. Therefore (s_1^*,s_2^*) is an equilibrium.

II. We now need to show how to build the required a, X_1 and X_2. Note that it is quite natural to construct the sets X_i using lower parts of the linear orders R_i (somehow being careful in dealing with possible common parts). More exactly, suppose that we have two sets $Z_1, Z_2 \subset A$, and

$$a_i = \min R_i | (A - (Z_1 \cup Z_2)).$$

A pair (Z_1, Z_2) is called *admissible* if

a′) $Z_1 \cap Z_2 = \emptyset$,
b′) $Z_i \subset L(a_i, R_i),\ i = 1, 2$,
c′) participant i blocks $Z_i,\ i = 1, 2$.

For example, the pair $(\emptyset, \emptyset)$ is admissible. Let (Z_1, Z_2) be a maximal (with respect to inclusion) admissible pair. Then participant 1 does not block the set $Z_1 \cup \{a_1\}$ (otherwise we could enlarge Z_1 to $Z_1 \cup \{a_1\}$); similarly, participant 2 does not block the set $Z_2 \cup \{a_2\}$. By force of the maximality of the blocking and the property a′), this is possible only if $a_1 = a_2$. Pose now $a_1 = a_2 = a$ and $X_i = Z_i \cup \{a\}$. It is obvious, that the properties a), b), c) are satisfied for these a, X_1, X_2. ■

In the case of more than two participants, the blocking B_π, generated by a consistent mechanism π, might be not maximal (despite the claim of Golberg and Gurvich, 1986).

(2.2.4) Example. Suppose four participants name alternatives, so that $S_i = A$. If three or more happen to name the same alternative, then this alternative wins. Otherwise, the outcome is determined through the use of the roulette mechanism described in Example (2.1.2).

It is clear that the mechanism is consistent. Indeed, at any preference profile, there exists a “stalemate” equilibrium, i.e., in which all participants name the same alternative.

We assert that any coalition K of a size two is powerless. For any fixed strategy of the coalition K, the complementary coalition $\overline{K}$ can force any alternative as an outcome. Indeed, the coalition $\overline{K}$ just needs to strive in order to force the outcome to be determined by the roulette mechanism, and then manoeuvre in order to crown one of them as “king”. Therefore, here the blocking B_π is not maximal.■

In the case of three or more participants maximality of the blocking B_π does not imply the consistency of the mechanism π. The reader will find an illustrative example in Gurvich (1975). Nevertheless, a weaker version of this claim happens to be true.

(2.2.5) Theorem. *If a blocking B is maximal, then there exists a consistent mechanism π such that $B = B_\pi$.*

We obtain as a corollary that to any blocking B we can associate a consistent mechanism π, such that $B \subset B_\pi$.

The proof of this theorem is constructive. Given a blocking B, we devise a mechanism π. Then we prove that B is consistent and verify the equality $B = B_\pi$. Omitting the maximal subset blocked by empty coalition, we can (and shall) assume that the blocking B is sovereign (i.e. that it satisfies the axiom B4). Moreover, we assume that there are more then two participants.

Constructing a mechanism. Recall that in Chapter 1, (1.5.10-11), we constructed the mechanism π_B generating blocking B. However π_B might happen to be non-consistent. So we make use here of an “old trick” in the realm of Nash theory. Namely, before putting the mechanism π_B to work, we grant participants a chance to come to a unanimous agreement (consensus) about an outcome. More formally, we consider a composite (in the sense of (1.5.10)) mechanism. The coarse part of the mechanism uses the unanimity rule: each participant i names some alternative a_i and if all named a_i are identical, then the outcome is determined and equal to a_i. Otherwise, when disagreement occurs, the outcome is defined by the mechanism π_B.

Proving consistency. Let R_N be a preference profile, and a be some individually rational alternative at R_N. We construct an equilibrium whose outcome is a. Each participant sends a message $a_i = a$. However this is only the "coarse part" of his strategy. Additionally, he suggests to remaining participants to block the set $\overline{L(a, R_j)}$, if the message of a participant j is not equal to a. We assert that this profile of messages is a Nash equilibrium. Indeed, first, the outcome at this strategy profile is equal to a. Secondly, if some participant j tries to change the outcome and therefore sends a message $a_j \neq a$, then all remaining participants attack him and block the set $\overline{L(a, R_j)}$. Note that they actually can block $\overline{L(a, R_j)}$, since due to maximality of B participant j does not block $L(a, R_j)$ (individual rationality). Therefore an outcome remains in $L(a, R_j)$ for any j, which proves that we have reached an equilibrium.

Proving that $B = B_\pi$ is straightforward. Let us show that $B \subset B_\pi$. Suppose that a non-empty coalition K blocks a set X. Then this coalition can do the following: first, break a consensus (if there was one) and second, block X, using mechanism π. This proves the needed inclusion. The reverse inclusion follows from maximality of B. ■

(2.2.6) The Mixed Strategy Issue. If the agents resort to mixed strategies (denoted by $\sigma_i \in \Delta(S_i)$) then the mechanism's outcomes will be lotteries on A, i.e., elements of $\Delta(A)$:

$$\pi : \prod_i \Delta(S_i) \to \Delta(A).$$

It is quite natural to take the set of affine utility profiles $u_N \in U$ to be defined on $\Delta(A)$ as the appropriate environment. In chapter 3, we study direct strategy-proof mechanisms in affine environments. We discuss here the Nash-consistency issue only. Using mixed strategies might bring about new equilibria. For example, the "roulette mechanism", which seldom has any pure strategy Nash equilibria, is Nash-consistent in mixed strategies. For instance, the participants need only choose the uniform distribution on the set $\{1, \ldots, n\}$ and indicate their best alternatives. Then every pure and mixed strategy yields the same expected pay-off.

Jackson (1992) provides an interesting example of a mechanism with $|N| = 2, |A| = 4$ in which a mixed-strategy Nash equilibrium Pareto dominates a pure strategy Nash equilibrium.

We know almost nothing about the issue of mixed strategy Nash-consistency in universal environments. Maskin and Moore (1987) discuss the role of

Pareto efficiency in accounting for mixed-strategy equilibria in implementing mechanisms. We give here a simple and necessary condition for the mixed-strategy Nash-consistency of mechanisms, which is similar to Proposition (2.2.2).

(2.2.7) Proposition. *Let π be a mixed-strategy Nash-consistent mechanism. Then for any covering of $\Delta(A)$ by sets of the form $L_i = \{x \in \Delta(A) | u_i(x) \leq c_i\}$, $i = 1, \dots, n$, there exists an agent j such that coalition $N \setminus \{j\}$ enforces the set L_j.*

Proof. We take the affine utility profile $u_N \in U$ defined on $\Delta(A)$. Let $\pi(\sigma_N^*) = x^*$ be the mixed-strategy Nash equilibrium outcome. Then $x^* \in L_j$ for some j. Since σ_N^* is an equilibrium strategy profile, $u_j(\pi(\sigma_j, \sigma_{-j}^*)) \leq u_j(\sigma_N^*) \leq c_j$ for any $\sigma_j \in \Delta(S_j)$. Hence σ_{-j}^* enforces L_j. ■

2.3 Strongly Monotone Social Choice Correspondences

(2.3.1) In this section, we investigate a correspondence of equilibrium outcomes $F_\pi : \mathbf{L}^N \Longrightarrow A$ generated by a mechanism π and defined as:

$$F_\pi(R_N) = \pi(NE(\pi(R_N)).$$

In this case, the correspondence F_π is said to be *Nash-implemented* (or simply *implemented*) by the mechanism π. An SCC $F : \mathbf{L}^N \Longrightarrow A$ is called *implementable* if it takes the form F_π for some mechanism π. Here and further on, we shall not assume that the mechanism π is consistent.

What are the properties of an implementable correspondence? We have discussed previously (refer to the reformulation of Proposition (2.2.2)) one property of implementable correspondence, namely, $F_\pi(\cdot) \subset IC(B_\pi^\beta, \cdot)$. Another important property - monotonicity - was discovered by E. Maskin.

(2.3.2) Proposition. *An implementable correspondence is monotone.*

The proof is very simple. Assume that the outcome $a = \pi(s_N^*)$ is an equilibrium at a preference profile R_N, and let R'_N be another preference profile such that $R_N \preceq_a R'_N$. We assert that the strategy profile s_N^* is also an equilibrium at R'_N. Indeed, for any $i \in N$ we have $\pi(\cdot, s_{N-i}^*) \subset L(a, R_i) \subset L(a, R'_i)$.■

(2.3.3) Remark. This line of reasoning is valid to prove monotonicity in the cases where we use alternative equilibrium concepts (core, strong equilibrium), although it does not work for all kinds of equilibrium concepts since some "myopia" in agents' behaviour is required. For example, the sophisticated equilibrium (or subgame perfect equilibrium) concept generates non-monotone SCCs. One could introduce an equilibrium-with-expected-reply-reactions concept, generalizing both Nash (or strong) equilibria (in which one assumes that other participants do not react on deviation) and core (in which any reactions are admitted). And again the equilibrium outcome correspondence (in this generalized sense) would be monotone.

(2.3.4) Strong Monotonicity. A correspondence of Nash equilibrium outcomes possesses a property that reinforces monotonicity. We need one additional notion in order to formulate this property.

Let $F : \mathbf{L}^N \Longrightarrow A$ be an SCC. Fix a participant i and a set $X \subset A$. We say that an alternative x from X is *F-essential* for i if $x \in F(R_N)$ for some preference profile R_N with $L(x, R_i) \subset X$. We denote the set of F-essential (for a participant i) alternatives in X by $\mathrm{Ess}_i(F; X)$; it is a subset of X, an "essential" part of X. We shall often omit the letter F when the context is clear enough.

In order to understand this notion, note that a non-essential alternative x is blocked by the participant i as soon as $L(x, R_i) \subset X$, that is as soon as he sets x low enough in his preferences. For example, an alternative x that does not belong to the image of the correspondence F is essential for no participant. Another example: we exhibit the essential elements for the correspondence $U(a)$ from Example (1.3.3). As is easily seen,

$$\mathrm{Ess}_i(U(a), X) = \begin{cases} X, & \text{if } a \in X \\ \emptyset, & \text{if } a \notin X \end{cases} .$$

(2.3.5) Definition. A correspondence F is *strongly monotone* if it has the following property. Let $a \in F(R_N)$ and let R'_N be a preference profile such that $L(a, R'_i) \supset \mathrm{Ess}_i(F; L(a, R_i))$ for each participant i, then $a \in F(R'_N)$.

Note that strong monotonicity implies monotonicity. Indeed, if $R'_N \succeq_a R_N$, then $L(a, R'_i)$ contains $L(a, R_i)$ and, therefore, contains $\mathrm{Ess}_i(F; L(a, R_i))$. Strong monotonicity of F means that an alternative a "survives" not only when it rises, but also when it is dropped "non-essentially". For example, if an alternative x does not belong to the image of F, and the correspondence F is strongly monotone, then $F(R_N)$ depends only on a restriction of R_N to the set $A - \{x\}$.

Both to acquire some familiarity with the strong monotonicity concept and sense better its difference with the monotonicity concept, we consider two examples.

(2.3.6) Example. Three participants and three alternatives: a, x and y. The correspondence F has the form: $F(R_N) = \{a\}$ if a gathers a number of points ≥ 8 points (using the Borda rule), and $F(R_N) = \emptyset$ otherwise. The correspondence F is monotone, albeit not strongly monotone. To see that, we consider two preference profiles

$$R_N = \begin{array}{|c|c|c|} x & a & a \\ a & * & * \\ y & * & * \\ \hline 1 & 2 & 3 \end{array}\,, \quad R_N = \begin{array}{|c|c|c|} x & a & a \\ y & * & * \\ a & * & * \\ \hline 1 & 2 & 3 \end{array}\,.$$

Then $F(R_N) = \{a\}, F(R'_N) = \emptyset$, although the restrictions of R_N and R'_N to the set $\mathrm{Im}F = \{a\}$ coincide.

(2.3.7) Example. The correspondence $U(a)$ considered above, is strongly monotone. This can be seen by considering the above formula for Ess and by applying the following lemma.

(2.3.8) Lemma. *Let a correspondence F be monotone. Suppose that for any participant i and any subset $X \subset A$ the set $\mathrm{Ess}_i(F; X)$ is either empty or equal to X. Then the correspondence F is strongly monotone.*

Indeed, let $a \in F(R_N)$. Then $a \in \mathrm{Ess}_i(L(a, R_i))$, hence $\mathrm{Ess}_i(L(a, R_i))$ is non-empty and, by assumption, equal to $L(a, R_i)$. Suppose now that R'_N is a preference profile satisfying $L(a, R'_i) \supset \mathrm{Ess}_i(F; L(a, R_i))$ for any participant i, then $L(a, R'_i) \supset L(a, R_i)$, for any i, and therefore (from the monotonicity of F) $a \in F(R'_N)$. ■

We give an ultimate general fact about strong monotonicity in the following proposition (compare with Lemma (1.3.5)):

(2.3.9) Proposition. *If F and G are strongly monotone SCCs, then $F \cup G$ is also a strongly monotone SCC.*

Proof. We state the following equality (whose proof is straightforward):

$$\mathrm{Ess}_i(F \cup G, X) = \mathrm{Ess}_i(F; X) \cup \mathrm{Ess}_i(G; X).$$

Now, let $a \in (F \cup G)(R_N)$, and R'_N be another preference profile such that $L(a, R'_i) \supset \mathrm{Ess}_i(F \cup G, L(a, R_i))$ for any $i \in N$. Then

$$L(a, R'_i) \supset \mathrm{Ess}_i(F, L(a, R_i)) \quad \text{and} \quad L(a, R'_i) \supset \mathrm{Ess}_i(G, L(a, R_i))$$

for any i. Since $a \in (F \cup G)(R_N) = F(R_N) \cup G(R_N)$, we can assume, for example, that $a \in G(R_N)$. Now from strong monotonicity of G, $a \in G(R'_N)$ and therefore $a \in (F \cup G)(R'_N)$. ■

(2.3.10) Corollary. *For any SCC F, there exist a (unique) maximal strongly monotone sub-correspondence $F^{sm} \subset F$.*

Indeed, F^{sm} is the union of all strongly monotone sub-correspondences of F. ■

The following theorem explains our interest in the notion of strong monotonicity.

(2.3.11) Theorem. *For any mechanism π, the equilibrium outcome correspondence F_π is strongly monotone.*

Proof. Let $a \in F_\pi(R_N)$. That is $a = \pi(s_N^*)$, where s_N^* is a Nash equilibrium for the game $G(\pi, R_N)$. We assert that for any participant i and each of his strategies $s_i \in S_i$, the following inclusion is satisfied:

$$\pi(s_i, s_{N-i}^*) \in \text{Ess}_i(F_\pi; L(a, R_i)).$$

Remark that once the assertion holds, then it clearly follows that the bundle s_N^* is an equilibrium for any preference profile R'_N such that $L(a, R'_i) \supset \text{Ess}_i(F; L(a, R_i))$, for all $i \in N$.

Suppose the reverse holds, that is:

$$\pi(s_i, s_{N-i}^*) \notin \text{Ess}_i(F_\pi; L(a, R_i))$$

for some $i \in N$ and $s_i \in S_i$. In other words, the element $x = \pi(s_i, s_{N-i}^*)$ is non-essential for i in the set $X := L(a, R_i)$. Form the following preference profile,

$$Q_N = \left|\begin{array}{c|c} \overline{X} & x \cdots x \\ x & * \\ X - \{x\} & * \\ \hline i & N-i \end{array}\right| .$$

The set $L(x, Q_i)$ is equal to X, and x is non-essential in X. Therefore $x \notin F_\pi(Q_N)$, which means that the outcome $x = \pi(s_i, s_{N-i}^*)$ is not an equilibrium for the game $G(\pi, Q_N)$. No participant $j \neq i$ will have any interest in changing x, since it is his best alternative. Hence participant i can improve on the outcome x (as x is not an equilibrium). In other words, there exists a strategy $s'_i \in S_i$ such that $\pi(s'_i, s_{N-i}^*)) \in \overline{X}$. But this contradicts the fact that s_N^* is an equilibrium for the game $G(\pi, R_N)$, since any element from X is prefered to $a = \pi(s_N^*)$ with respect to R_i. ■

In particular, we obtain that the correspondence F from Example (2.3.6) can be implemented by no mechanism, albeit it is monotone. Should the reader be unsatisfied with the possible empty-valuedness of F, he might substitute F by $F \cup \{x\}$. This new correspondence is also a monotone SCC and is implemented by no consistent mechanism.

Convinced as we are of the importance of strong monotonicity, we establish one further useful property of strongly monotone correspondences.

(2.3.12) Proposition. *Let F be a strongly monotone correspondence, let i be a participant and $X \subset A$. The following assertions are equivalent:*

a) *i blocks the set X through F;*

b) $\text{Ess}_i(F; X) = \emptyset$;

c) *$F(R_N) \subset \overline{X}$ for any preference profile R_N , provided R_i satisfies ($\overline{X} \succ X$).*

Proof. We show the implication a) $\Rightarrow$ b). Let $\{i\} B_F X$, that is there exists a preference R'_i such that $F(R'_i, *) \subset \overline{X}$. Suppose now that the set $\text{Ess}_i(F; X)$ is non-empty, and $a = \max R'_i | \text{Ess}_i(F; X)$. Then,

$$\mathrm{Ess}_i(F; X) \subset L(a, R'_i).$$

Since the element a is essential in X, then $a \in F(R_N)$, for some preference profile R_N with $L(a, R_i) \subset X$. Using monotonicity of F and propping a up, if needed, one can assume that $L(a, R_i) = X$. Now, $a \in F(R_N)$ and $L(a, R'_i) \supset \mathrm{Ess}_i(F; L(a, R_i))$. By definition of strong monotonicity we can conclude that $a \in F(R'_i, R_{N-i}) \subset \overline{X}$, which contradicts the relation $a \in X$.

The implication b) $\Rightarrow$ c) follows from the definition of non-essential elements. The implication c) $\Rightarrow$ a) is obvious. ■

From this we deduce an important assertion on equilibrium outcome correspondences.

(2.3.13) Corollary. *Let π be a mechanism, $F = F_\pi$ be the equilibrium outcome correspondence of this mechanism, and $X \subset A$. The following assertions are equivalent:*

1) *a participant i β-blocks X;*
2) *a participant i blocks X through F.*

Proof. Clearly enough, if a participant i β-blocks X then, for any preference R_i of the type $(\overline{X} \succ X)$, no equilibrium outcomes can be in X. This proves the implication 1) $\Rightarrow$ 2).

Prove now the implication 2) $\Rightarrow$ 1). Suppose that a participant i blocks X through F_π. According to Proposition (2.3.12), we know that for any preference R_i of the type $(\overline{X} \succ X)$ each equilibrium outcome is in $\overline{X}$. Picking now an arbitrary strategy $s_{N-i} \in \times_{j \neq i} S_j$, we have to show that there exists a strategy $s_i \in S_i$ such that $\pi(s_i, s_{N-i}) \in \overline{X}$. Let $x = \pi(s_i, s_{N-i})$, where s_i is an element of S_i. If $x \notin X$ then everything is all right. If $x \in X$, then take the profile Q_N shown in the proof of Theorem (3.11). Given these preferences, no participant $j \neq i$ has any interest in changing x, since it is his best alternative. Moreover $x \in X$ and therefore it is not an equilibrium. Hence the participant i can improve upon this outcome, in the sense that he can figure out a strategy s'_i such that $\pi(s'_i, s_{N-i}) \in \overline{X}$. ■

2.4 Nash-implementable Correspondences

The main result of the preceding Section asserts that the equilibrium outcome correspondence associated with any mechanism is strongly monotone. Are there some other properties? The answer is “no”; at least when there are more than two participants. We show in this Section that in the latter case any strongly monotone correspondence can be implemented by some mechanism. The case of two participants is slightly different and thus will be considered in Section 2.5.

(2.4.1) The Maskin Mechanism. Let F be an SCC. We construct here the mechanism $\pi = \pi_F$ which implements F, in the case of strong monotonicity. The idea of this construction was originally put forward by E. Maskin, although some details are new and prompted by the strong monotonicity property.

Thus, let $F : \mathbf{L}^N \Longrightarrow A$ be an SCC, and $A^* = \cup_{R_N} F(R_N)$ be the image of F. A participant's message is a pair (R_N, x), for which $R_N \in \mathbf{L}^N$ and $x \in F(R_N)$. In other words, the set S_i of participant i's messages is equal to the graph of the correspondence $F \subset \mathbf{L}^N \times A$. In a more figurative way, each participant both tries to guess the preferences of the whole group and offers some admissible outcome.

Suppose that participants send the messages $s_i = (R^i_N, x^i)$. We now explain how an outcome is obtained. One needs here to distinguish three cases. In the first case, the *coordinated* case, all participants send the same message (R_N, x). In this case, the outcome is x.

The second case is *almost coordinated.* Here all participants except one (who is determined uniquely and called the *dissident*) send the same message (R_N, x); the dissident sends another message (R'_N, x'). The outcome is equal to x', if x' is contained in the set $\text{Ess}_i(F; L(x, R_i))$, and is equal to x otherwise. R_i is the preference of the dissident from the point of view of the others, $N - \{i\}$. In words, each participant has a right to disagree if the others both misrepresent his preferences too strongly and try to impose on him a "bad" alternative. Note that Maskin proposes use of the set $L(x, R_i)$ in place of the set $\text{Ess}_i(F; L(x, R_i))$; we refine slightly the construction here.

Finally, in all other *uncoordinated* cases, outcomes are determined using the roulette mechanism with values in A^* (see the notion of a composite mechanism and the roulette introduced in (1.5.5) and (1.5.10) of Chapter 1). With this the description of Maskin's mechanism $\pi = \pi_F$ is completed.

Let us study the equilibrium outcome correspondence F_π implemented by the Maskin mechanism $\pi = \pi_F$. Remark that $F \subset F_\pi$ always holds. In other words, any alternative $a \in F(R_N)$ is attained as an equilibrium outcome for the game $G(\pi, R_N)$. For this all participants must send the same message $s^*_i = (R_N, a)$. Then $\pi(s^*_N) = a$. Can somebody improve upon this outcome? No. In fact, every participant i could become a dissident by deviating from s^*_i and enforcing as a new outcome any element of $\text{Ess}_i(F; L(a, R_i))$. However, this set being contained in $L(a, R_i)$, i does not improve upon a, and thus a is an equilibrium outcome.

The reverse inclusion $F_\pi \subset F$ is true only for strongly monotone correspondences F (compare this assertion with Theorem (2.3.11)).

(2.4.2) Theorem. *Let $|N| \geq 3$. If a correspondence F is strongly monotone, then it is implemented by the Maskin mechanism π_F.*

Proof. Without loss of generality one can assume that $A^* = A$. One needs to prove that if s^*_N is an equilibrium for the game $G(\pi, R_N)$, then $a = \pi(s^*_N)$ is contained in $F(R_N)$. We consider three cases.

The first case: the situation s_N^* is coordinated, all participants send the same message (R_N^*, x). By definition of π_F, $a = x$ and a is contained in $F(R_N^*)$. Each participant i can enforce as a new outcome any element of the set $\mathrm{Ess}_i(L(a, R_i^*))$. However (since we are at an equilibrium), he does not do it, which means that it is not profitable for him, that is $\mathrm{Ess}_i(L(a, R_i^*)) \subset L(a, R_i)$. Now from strong monotonicity of F we conclude that $a \in F(R_N)$.

The second case: the situation s_N^* is almost coordinated, and call the dissident i. Suppose the messages of all participants except i are (R_N^*, x). If $a = x$, then we are back to the previous argument. Dissident i could enforce any outcome out of $\mathrm{Ess}_i(L(a, R_i^*))$, but he does not. Consequently, $\mathrm{Ess}_i(L(a, R_i^*)) \subset L(a, R_i)$. Any other participant j could as well decide to break this 'almost coordination' and crown himself as the "king" through the roulette mechanism; consequently, $a = \max R_j$ (i.e. the best-for-j alternative) for any $j \neq i$. As above, $a \in F(R_N)$.

Consider now the subcase, for which $a \neq x$ and the outcome a is determined by the dissident i. As explained above, $a = \max R_j$ for $j \neq i$. The dissident could enforce any element from $\mathrm{Ess}_i(X)$ as an outcome, where $X = L(a, R_i^*)$. Since he chose a, then $a = \max R_i | \mathrm{Ess}_i(X)$. By definition of an essential element, this means that $a \in F(Q_N)$ for some preference profile Q_N, with $L(a, Q_i) \subset X$. By strong monotonicity of F, one can assume that $L(a, Q_i) = \mathrm{Ess}_i(X)$. But now, $Q_N \preceq_a R_N$. Indeed, if $j \neq i$ then $L(a, R_j) = A$; if $j = i$ then $L(a, R_i) \supset \mathrm{Ess}_i(X) = L(a, Q_i)$, a being the best element of $\mathrm{Ess}_i(X)$ relatively to R_i. By monotonicity of F, we conclude that $a \in F(R_N)$.

The third case: the situation s_N^* is uncoordinated. Here any participant i could enforce as a new outcome any element of A^*. Therefore $a = \max R_i$ for any $i \in N$, and $a \in F(R_N)$ both by monotonicity of F and because $A^* = A$. ∎

As a consequence (and in the case of three or more participants), there exists a (unique) maximal implementable sub-correspondence of F, for any SCC F. Of course, this is F^{sm} (the maximal strongly monotone subcorrespondence of F, see the preceding Section).

Thus, to establish the implementability of a correspondence F it suffices to prove its strong monotonicity. We give now some important classes of strongly monotone correspondences.

(2.4.3) Proposition. *If a correspondence F is monotone and neutral, then it is strongly monotone* (and, consequently, it is implemented by the Maskin mechanism for $|N| \geq 3$).

Indeed, by the neutrality of F, the set $\mathrm{Ess}_i(X)$ is either empty or equal to X. The sequel follows from lemma (3.8). ∎

Correspondences satisfying the *no veto power property* provide yet another class of strongly monotone SCCs. We say that a participant i is *weak* if he is not able to block (using correspondence F) any alternative. It is evident now that $\mathrm{Ess}_i(X) = X$ for any $X \subset A$, and again, by Lemma of Section 3,

F is strongly monotone if both F is monotone and all participants are weak. This proves the following

(2.4.4) Proposition. *Suppose that $|N| \geq 3$, F is a monotone correspondence, and all participants are weak. Then the Maskin mechanism π_F implements F.* ■

This assertion was proved by E. Maskin.

Minimal monotone correspondences (see Appendix 1.A3) form the last important class of strongly monotone correspondences.

(2.4.5) Proposition. *If F is a minimal monotone SCC, then F is strongly monotone* (and, consequently, is implemented by the Maskin mechanism for $|N| \geq 3$).

Proof. Let $a \in F(R_N)$, $X_i = L(a, R_i)$ and $E_i = \text{Ess}_i(F; X_i)$. One needs to prove that if Q_N is another preference profile and $L(a, Q_i) \supset E_i$ for any $i \in N$, then $a \in F(Q_i)$. By the minimality criterion (Chapter 1, Appendix 1.A3), there exists a preference profile R'_N such that $F(R'_N) = \{a\}$ and $X'_i = L(a, R'_i) \subset X_i$. By monotonicity of the operator Ess, $\text{Ess}_i(X') \subset E$. Substituing R_N by R'_N , we can assume that $F(R_N) = \{a\}$.

Now we prop up all the non-essential elements of each participant above a, while keeping invariant the arrangement of the remaining alternatives. Formally, we consider preferences $P_i = (R_i|(A - X_i), R_i(X_i - E_i), R_i|E_i)$. If $x \in X_i - E_i$, that is if x is non-essential in X_i, then $x \notin F(P_N)$ because $L(x, P_i) \subset X_i$. Therefore the set $F(P_N)$ does not intersect the set $\cup_i(X_i - E_i)$. Moreover the remaining elements can only but fall, when we move from R_N to P_N. Therefore $F(P_N) \subset F(R_N) = \{a\}$. Since F is a nonempty-valued SCC, then $F(P_N) = \{a\}$. It suffices to note that $P_N \preceq_a Q_N$ and to use the monotonicity of F. ■

(2.4.6) Corollary. *Let F be a monotone nonempty-valued SCC, and $|N| \geq 3$. Then there exists a consistent mechanism π such that $F_\pi \subset F$.*

It suffices to take a minimal monotone sub-correspondence of F and to apply the Maskin mechanism. ■

(2.4.7) Using the results obtained above, we exhibit several concrete implementable correspondences; assuming everywhere that $|N| \geq 3$.

I. Let $F(R_N) = A$ for any preference profile R_N. This SCC is monotone and neutral, hence, Proposition (2.4.3) is applied.

II. The Paretian SCC Par is implementable for the same reason.

III. The union of a few dictatorial SCFs is monotone and neutral, hence, is implementable. However, we have already seen (Section 2.3) that the union of strongly monotone SCCs is also strongly monotone.

IV. The correspondence $U(a)$ from Example (1.3.3).

V. For the same reasons, the Maskin correspondence $M(a)$ is implementable. One can also remark that this correspondence is minimal monotone.

VI. As will be shown in Chapter 4, any core correspondence $C(B, \cdot)$ is strongly monotone and, hence, implementable.

(2.4.8) Let us make two closing remarks on the case of three or more participants. They aim at clarifying possible misinterpretations of Theorem (2.4.2). First we would like to come back to the idea that the Maskin procedure might enable the explicit finding of many consistent mechanisms. It is not so. In fact, the procedure rests crucially on nonempty-valued strongly monotone SCCs, and we hardly know how to construct monotone SCCs.

The second remark concerns the plausibility of the equilibria which were used in proving the consistency of a mechanism. These equilibria exist, but it is quite difficult to understand how participants might built these equilibria through their individual actions. For this, we "just" required that they both correctly *guess* each other's preferences and thereafter *reach an agreement* about an outcome. If the participants could communicate and talk, then they could, perhaps, have a guess at others' preferences. But then they would most probably behave cooperatively, which in turn would imply that we use another solution concept in place of Nash equilibria. Indeed if participants are able to spontaneously reach an agreement, then why would they need a mechanism at all? Without agreement they would most probably find themselves in a uncoordinated state, which is most certainly not an equilibrium. The situation is worse than that, participants are never able to exit uncoordinated states through sequences of best replies. In short, despite the optimistic formulation of Theorem (2.4.2), the Maskin mechanism is non-consistent in practice. We do not know whether "better" mechanisms exist nor do we know what they might implement.

2.5 Implementation: the Case of Two Participants

(2.5.1) The two participants' case occupies a special place in the theory of Nash equilibrium. Indeed the restriction to two participants is crucial for the theorem of Gurvich and we see further that it is specific in other instances. For example, the Maskin mechanism is not defined when $n = 2$ for an odd reason; in effect, in this case it is not obvious which one of the two participants is the dissident. However, the specificity of the two participants' case takes its origins in a more essential cause: in this case, strong monotonicity does not warrant implementation. For example, as we shall soon see, the trivial SCC $F \equiv A$ is not implementable, although it is obviously strongly monotone. Therefore we need to reexamine the necessary conditions of implementability and strengthen them.

In the case of two participants, the inclusion $F \subset IC(B_F)$ (which is true for any number of participants, see Proposition (2.2.2) and Corollary (2.3.13)) begins to play an important role. Additionally in this case the "blocking" B_F is *sub-additive: if participant 1 does not block a set* X_1*, and participant 2 does not block a set* X_2*, then* $X_1 \cap X_2 \neq \emptyset$. The proof follows from the above-mentioned Corollary: if participant 1 does not block X_1, and F is implemented by a mechanism π, then participant 2 forces X_1 through the mechanism π. Similarly, participant 1 forces X_2, hence X_1 intersects X_2.

The non-implementability of the SCC $F \equiv A$ follows from here, because the blocking B_F is not sub-additive. Indeed, here a participant can only block the empty subset.

Actually, implementable SCCs possess a finer property than just sub-additivity of B_F. To formulate it, we need the following notion (compare with the notion of essential elements).

(2.5.2) Definition. Consider the two subsets $X_1, X_2 \subset A$. An alternative a from $X_1 \cap X_2$ is *bi-essential,* if there exists a preference profile $R_N = (R_1, R_2)$ such that $a \in F(R_N)$ and $L(a, R_i) \subset X_i$ for $i = 1, 2$. The set of bi-essential elements is denoted by $\mathrm{Bess}(F; X_1, X_2)$.

(2.5.3) Definition. An SCC F has the *MR-property* (in honor of Moore and Repullo (1990)) if the set $\mathrm{Bess}(F; X_1, X_2)$ is non-empty for any X_1, X_2 such that $i\overline{B_F}X_i$, for $i = 1, 2$.

The MR-property clearly strengthens the sub-additivity property, since by definition $\mathrm{Bess}(F; X_1, X_2) \subset X_1 \cap X_2$.

(2.5.4) Proposition. *Let π be a mechanism with two participants. Then the equilibrium outcome correspondence F_π has the MR-property.*

Proof. Let the sets X_1,X_2 be given, and suppose that no participant i blocks X_i through F. According to Corollary (2.3.13), a participant i does not β-block X_i. This means that participant 2 has a strategy s_2^* such that $\pi(\cdot, s_2^*) \in X_1$. Similarly, participant 1 has a strategy s_1^* such that $\pi(s_1^*, \cdot) \in X_2$. Therefore, the alternative $a = \pi(s_1^*, s_2^*)$ belongs to $X_1 \cap X_2$. Let now R_N be such a preference profile that $L(a, R_i) = X_i$ for $i = 1, 2$. It is obvious that (s_1^*, s_2^*) is an equilibrium for the game $G(\pi, R_N)$, thus $a \in \mathrm{Bess}(F; X_1, X_2)$. ■

(2.5.5) To sum-up: an implementable correspondence F satisfies the following three properties:

1. F is strongly monotone;
2. $F \subset IC(B_F)$;
3. F has the MR-property.

We assert that properties 1-3 are not only necessary, but also sufficient for Nash-implementability. For this purpose, we construct the following implementation mechanism $\mu = \mu_F$. The basic strategies of this mechanism consist of some subsets of A. More exactly, a participant i calls any subset $X_{-i} \subset A$, which is not blocked (through B_F) by the other participant $-i$. Once participants selected respectively X_{-1} and X_{-2}, an outcome is formed using the roulette mechanism with values in $\mathrm{Bess}(F; X_{-2}, X_{-1})$. The latter set is non-empty if F has the MR-property.

(2.5.6) Proposition. *Let F be a correspondence having the MR-property, and $\mu = \mu_F$ be the corresponding mechanism. Then:*

a) if $F \subset IC(B_F)$ then $F \subset F_\mu$;

b) if F is strongly monotone, then $F \supset F_\mu$.

In particular, if F has the properties 1-3, then $F = F_\mu$, that is F is implemented by the mechanism μ.

Proof. a) Suppose that $a \in F(R_N)$. Then, according to the inclusion $F \subset IC(B_F)$, the alternative a is individually rational, hence participant i does not block the set $L(a, R_i)$. Therefore, participant i can send the "coarse" message $X_{-i} = L(a, R_{-i})$. It is obvious that a belongs to $\mathrm{Bess}(F; X_{-2}, X_{-1})$ and that it is the best element of this set for both participants. Therefore, the dictator whom ever he is, chooses a. We assert that this is a Nash equilibrium. Indeed the outcome remains in $L(a, R_1)$ whatever the message of participant 1, assuming participant 2's message $X_{-2}{=}L(a, R_{-2}) = L(a, R_1)$ remains fixed. A similar reasoning applies to any moves of participant 2.

Let us now prove b). Assume the "coarse" messages X^*_{-1},X^*_{-2} yield an equilibrium with outcome a for the game $G(\mu, R_N)$. We show that $a \in F(R_N)$. Suppose that participant 1 both changes his message from X^*_{-1} to $X_{-1} = A$ and puts himself in the king's seat. Then he can enforce any alternative from the set $\mathrm{Bess}(F; X^*_{-2}, A) = \mathrm{Ess}_1(F; X^*_{-2})$ as an outcome. Since we are at an equilibrium then, $\mathrm{Ess}_1(F; X^*_{-2}) \subset L(a, R_1)$. Similarly, $\mathrm{Ess}_2(F; X^*_{-1}) \subset L(a, R_2)$. Since $a \in \mathrm{Bess}(F; X^*_{-2}, X^*_{-1})$, then $a \in F(Q_N)$ for some a preference profile Q_N such that $X^*_{-i} \supset L(a, Q_i)$, $i = 1, 2$. Therefore

$$\mathrm{Ess}_i(F; L(a, Q_i)) \subset \ \mathrm{Ess}_i(F; X^*_{-i}) \subset L(a; R_i) \text{ for } i = 1, 2.$$

From the strong monotonicity of F, we conclude that $a \in F(R_N)$. ■

(2.5.7) Let us give a "typical" example of an implementable correspondence with two participants. Suppose that B is a "blocking", that is a monotone relation between $N = \{1, 2\}$ and 2^A. Let $F = IC(B)$ be the correspondence of individually rational outcomes. We assert that *the correspondence $IC(B)$ is implementable if B is sub-additive.*

According to Proposition (2.5.6), one needs to prove that the correspondence $IC(B)$ satisfies properties 1-3. However we can do it in a simpler way. First, note that $F = IC(B)$ has the MR-property. This results trivially from sub-additivity: if iBX_i for $i = 1, 2$, then X_1 intersects X_2. Let a be an arbitrary element of $X_1 \cap X_2$, and R_N be a preference profile, such that $L(a, R_i) = X_i$. Then $a \in IC(B)$. This proves that $\mathrm{Bess}(IC(B); X_1, X_2) = X_1 \cap X_2$ and

F has the MR-property. Therefore the mechanism $\mu = \mu_F$ is well-defined. We prove directly that $F = F_\mu$. By the same line of reasoning used in proving a) we get the inclusion $F \subset F_\mu$. Further, $F_\mu \subset IC(B^\beta_\mu) \subset IC(B) = F$, since $B \subset B^\beta_\mu$.

Example (2.1.2) shows that an implementable correspondence F can be different from its individually rational core $IC(B_F)$. However, we have the feeling that both these correspondances should not differ too much one from another. For example, if the correspondence $IC(B_F)$ is nonempty-valued, then F will also be nonempty-valued. Indeed, if $IC(B_F)$ is nonempty-valued, then B_F is a fully-fledged blocking, maximal by force of sub-additivity. Therefore, by the Gurvich theorem, the implementing mechanism is consistent. Nevertheless, we do not know how we should formulate precisely the idea of 'proximity' between F_π and $IC(B^\beta_\mu)$.

The mechanism μ seems much more satisfactory (than the Maskin mechanism) in the sense that it happens to be easier for participants to find equilibrium strategies. In effect, even if participants do not know each other's preferences, they can use some "natural" strategies consisting in blocking alternatives ranked low in their preferences. And it is very likely that these actions lead to an equilibrium, which on top might be Pareto optimal. We explain this in the following example.

(2.5.8) Example. There are five alternatives, and each participant can block any two of them. Suppose that preferences of the participants are $R_1 = R_2 = (x \succ y \succ z \succ u \succ v)$. Then $IC(B, R_N) = \{x, y, z\}$, and all three outcomes x, y, z can be equilibria. However, the likelihood of their respective occurences is quite different! The outcome z occurs only in the (quite implausible) case when both participants block their best alternatives x and y. The occurence of y is also rather improbable. On the contrary, it is quite natural to think of x as an equilibrium when the first participant blocks $\{u, v\}$ and the second blocks $\{z, y\}$. Furthermore this example emphasizes the "richness" of the mechanism π in comparison to the implemented SCC F_π. We mean by this that among all Nash equilibria for the game $G(\pi, R_N)$ some might stand out of the crowd, in the sense that they appear more probable in comparison to others. Whereas in $F_\pi(R_N)$ all alternatives appear rather colourless.

2.6 Acceptable Mechanisms

We have seen that Nash equilibria can be non-efficient. All the more interesting are the mechanisms whose equilibrium outcomes are all Pareto optimal.

(2.6.1) Definition. A mechanism π is *acceptable* if it is consistent and $F_\pi \subset Par$.

In other words, we require here that equilibria exist for every preference profile, and that all of them be Pareto optimal.

We have already seen three examples of acceptable mechanisms: the dictatorial mechanism, the kingmaker (see Example (2.1.3)) mechanism and the mechanism presented in Example (2.1.6). Can we find other examples of such mechanisms? What if any, are the obstacles that prevent acceptability? In this section, we provide a set of partial answers to these questions.

Let us begin with the case of two participants. Generally one might expect that acceptable mechanisms with two participants be dictatorial (at least, in the sense of the associated blocking). Indeed, as we have already suggested, the equilibrium outcome correspondence of a mechanism with two participants, is close to the individual core, which contains many non-optimal outcomes if the associated blocking is not dictatorial. As it happens this is more than a mere expectation; it is true.

(2.6.2) Proposition (Hurwicz, Schmeidler). *Any acceptable mechanism with two participants is dictatorial.*

The proof is based on the following general fact:

(2.6.3) Lemma. *Let π be a mechanism. Suppose that there exists an alternative blocked by both a coalition K and its complement $\overline{K}$. Then the mechanism is not acceptable.*

Proof. Suppose a coalition K blocks some alternative a using a strategy $s_K \in S_K$, and suppose the coalition $\overline{K}$ blocks a using a strategy $s_{\overline{K}} \in S_{\overline{K}}$. Then the outcome $x = \pi(s_K, s_{\overline{K}})$ is different from a. Consider now a preference profile R_N, in which all R_i are identical and equal to $(a \succ x \succ ....)$. We assert that $(s_K, s_{\overline{K}})$ is an equilibrium for the game $G(\pi, R_N)$. Indeed, participants of the coalition K can not improve upon the outcome x (that is to move it towards a), because the coalition $\overline{K}$ prevents it using the strategy $s_{\overline{K}}$; similarly for $\overline{K}$. However, the outcome x is not efficient, since $Par(R_N) = \{a\}$. ■

(2.6.4) Corollary (Dutta). *Suppose that a mechanism π is acceptable, and that the associated blocking B_π is maximal. Then the blocking B_π is generated by a simple game* (see (1.5.7)).

Proof. We need to show that if a coalition K blocks some alternative a, then K forces any alternative. So, let K block a. By the previous lemma, its complement $\overline{K}$ does not block a. Therefore from the maximality of B_π, the coalition K blocks the set $A - \{a\}$. In particular, K blocks any alternative of A. Repeating the preceding line of reasoning, we get that K forces any alternative. ■

Proposition (2.6.2) follows now from the maximality of the blocking B_π associated to a consistent mechanism with two participants (Theorem (2.2.3)). ■

(2.6.5) Remark. One ought to evaluate the somewhat "pessimistic" Hurwicz-Schmeidler proposition with the same prudence as all other theorems related to the use of Nash equilibria. Their result is obtained at the expense of somewhat "implausible" equilibria. Consider, for instance, the mechanism described in Example (2.1.8). In this mechanism, participant 1 blocks some set $X \subset A$, while participant 2 (using an order R) picks an element out of $A - X$. The second participant has a dominant strategy which consists in choosing his true preference. And indeed he might use it. In this case, the whole situation is different. If participant 2 uses the strategy $s_2^* = R_2$, then any equilibrium outcome is efficient. Indeed, let $a = \pi(X, R_2)$ be an equilibrium outcome. Participant 1 can force the outcome to be any element x of the set $A - L(a, R_2)$; for this, he needs only to block the set $X - \{x\}$ instead of X. But he does not do it, which means that all these possible alternatives x are worse for him than a, that is $a \in Par(R_1, R_2)$. Therefore, this mechanism is acceptable in practice.

Thus we can find many "practically acceptable" mechanisms, even if formally speaking, except for dictatorial mechanisms, there are no other acceptable mechanisms in the case of two participants.

Let us examine the three-and-more-participants case. In this case, there are many acceptable mechanisms. Take any efficient and strongly monotone SCC, then its Maskin mechanism is acceptable. More specific examples of such SCCs are (see Section 2.4 above): the correspondence Par, the Maskin correspondence $M(a)$, any union of dictatorial SCFs, any core correspondence for a blocking with property B4. However, as we already noted, Maskin mechanisms formally have equilibria, but in all practical cases these equilibria are not attainable.

We now discuss the issue of participants' power within acceptable mechanisms. A participant i is *strong* if he is able to block (in the sense of B_π) at least one alternative; otherwise he is *weak*. Generalizing Proposition (2.6.2), we prove the following assertion:

(2.6.6) Theorem. *Let π be an acceptable mechanism. If some participant is strong then all other participants are weak.*

Proof. Let us suppose conversely that there are two strong participants i and j. Participant i blocks an alternative x, and participant j blocks an alternative y. By lemma (2.6.3), $x \neq y$. Moreover, by the same lemma, participant i β-enforces x, and the participant j β-enforces y. Now if their preferences are such that $\max R_1 = x$ and $\max R_2 = y$, then there are no equilibria. ■

In the preceding examples of acceptable mechanisms, all participants were weak. Therefore we can wonder whether we might find an acceptable mechanism with a strong participant, who is not a dictator. The answer is affirmative, as is shown in the following example.

(2.6.7) Example. There are three (or more) participants. A strategy of a participant consists in choosing a function $u_i : A \longrightarrow \mathbf{R} \cup \{-\infty\}$, where

$u_i(.) \neq -\infty$ for all participants (except participant 1,whose $u_1(.)$ at alternative a is possibly equal to $-\infty$). Further, all participants except participant 1 also call a number $\varepsilon_i \in \{0,1\}$. We define an outcome of this mechanism as follows: the outcome belongs to Argmax(u_1), whenever some of the ε_i is different from 0; the outcome maximizes the function $\delta_a \cdot u_1 + \sum_{i \neq 1} u_i$, where $\delta_a(x) = 0$ at $x \neq a$ and $\delta_a(a) = 1$, whenever all $\varepsilon_i = 0$.

We assert that:

a) the mechanism is consistent. Indeed, let all $\varepsilon_i = 1$, and the function u_1 be a representant of the true preference of participant 1. Then the outcome is maxR_1, and this is an equilibrium. In fact, no participant (excepting participant 1) is able to change the outcome.

b) any equilibrium outcome is Pareto optimal. This is obvious if $\varepsilon_i = 1$ for some i. Suppose now that all ε_i are equal to 0. If a is the only Pareto optimal alternative, then it is an equilibrium outcome since participant 1 β-enforces a. If $Par(R_N) \neq \{a\}$, then the effectiveness of the outcome follows from the fact that any participant (different from 1) can enforce any alternative (different from a) to be an outcome.

c) participant 1 can block the alternative a. To do so, he might select a u_1 such that $u_1(a) = -\infty$. Thus he is strong.

d) participant 1 is not the dictator, since equilibrium outcomes can be different from maxR_1. (However, they can not be worse than a for participant 1.) ■

(2.6.8) In conclusing Chapter 2, we reckon that the Nash equilibrium concept is somewhat unsatisfactory. On the one hand, there are many absurd-looking equilibria; it is unlikely that reasonable people might choose such strategies. Apparently, one needs to use a stronger or more refined equilibrium concept, which decreases the number of equilibria. On the other hand, it remains unclear how participants could attain a "nice" equilibrium (when and if it exists); compare this with the discussion about the Maskin mechanism in (2.4.8). We would expect that mechanisms help to find "nice" equilibria and that they monitor somewhat our way towards "nice" outcomes.

In the next chapter, we consider in more detail the concept of dominant strategies equilibrium, which is more rigid than the Nash concept. This concept is free, in many respect, from previous objections. Some reinforcements of the Nash concept in the sense of coalitional stability will be considered in Chapter 5.

2.A A Simple Mechanism for the Implementation of Walrasian Equilibria

(2.A.1) In Chapter 2, we dealt with a finite set of alternatives A and arbitrary weak orders on it. Here we show how that approach works in an "economic environment", namely in a pure exchange economy. More exactly, we show that Walrasian equilibria are implementable by a simplified version of the Maskin mechanism.

We briefly sketch the pure exchange set-up at stake. The agents of some group N want to trade their initial endowments. Let G be a (finite) set of types of goods. A commodity bundle is an element of the space $V = \mathbf{R}^G$. The initial endowment of agent i is denoted by w_i. An after-trade allocation (x_i) satisfies the balance equality $\sum x_i = \sum w_i$. Any such allocation is termed *feasible*. Thus, the set A of alternatives consists of feasible allocations. Moreover we suppose that each agent only cares about the bundle he gets at the end of the day. Therefore agent i's preferences depend only upon x_i. Furthermore we assume preferences to be strictly monotone ($x > y$ implies that x is strictly preferred to y) weak orders R_i on V.

Suppose now that the commodity prices are given, that is we have a non-negative linear functional p on the commodity space V. A pair (p, x_N) is a *competitive* (or Walrasian) *equilibrium* if the allocation x_N is feasible and if the commodity bundle x_i is the most preferred (in the sense of R_i) bundle in the budget set $B(i,p) = \{x \in V, px_i \leq pw_i\}$ of every agent i. The set of equilibrium allocations x_N is denoted by $W(R_N)$.

(2.A.2) Here we consider a simple mechanism f implementing the Walrasian correspondence W in the case of three or more agents. This mechanism is a version of the Maskin mechanism seen in (2.4.1). Let us recall that in the Maskin mechanism π_F the (basic) strategies have the form (R_N, a) where $a \in F(R_N)$. However in that set-up the sets $L(a, R_i)$ play the crucial part. In our (as in Dutta et al. (1995)) mechanism, prices p will stand for the preference profile R_N whereas budget sets will stand for sets $L(a, R_i)$.

More precisely, agents strategies (or messages) take the form (p, x_N) where p is a non-negative linear functional on V and x_N is a feasible allocation satisfying $px_i = pw_i$ for every agent i. As in the Maskin mechanism, we define the outcome for each of the following three cases.

Case 1: all agents send the same message. The outcome is x_N.

Case 2: all agents except agent i send the same message, (p, x_N), whereas i sends (p', x'_N). The outcome is x'_N, if $px'_i \leq pw_i$; or x_N otherwise.

Case 3. The "roulette" mechanism acts as soon as case 1 or case 2 do not obtain.

In words this means that each agent proposes his pair of prices and allocation. The outcome of the mechanism is the common proposal when all proposals coincide. Now an agent can always deviate and try to obtain another bundle x'_i such that $px'_i = pw_i$, if it turns out that the bundle x_i offered to him is not best in his budget set $B(i,p)$. To do so he will have to send a message (p, x'_N) where $x'_j = x_j - (x'_i - x_i)/(n-1)$ for $j \neq i$.

(2.A.3) Proposition. *If R_N is a profile of strictly monotonic preferences the $W(R_N) = f(NE(f, R_N))$.*

The proof is based on the same arguments which were used in proving Theorem (2.4.2). If (p, x_N) is a Walrasian equilibrium, then it can be chosen as a strategy in our mechanism. If all agents send the message (p, x_N), we then have a Nash equilibrium. This proves the inclusion $\subset$. Conversely, suppose we have an equilibrium strategy profile. It is easy to understand that it should be a profile of coinciding messages (p, x_N). Indeed, if not some agent could activate the roulette mechanism and improve his outcome (we use here the strict monotonicity of preferences assumption). We now only need to check that (p, x_N) is a Walrasian equilibrium. This implies that we show that the bundle x_i is among the most preferred bundles in the budget set $B(i,p)$, for every agent i.

Suppose, conversely, that agent i prefers bundle x'_i to x_i and that x'_i is in his budget set $B(i,p)$. By monotonicity, we can consider that $px'_i = pw_i$. Therefore this agent can (by deviating from the agreed upon message as described above) get hold of this better bundle x'_i. But this contradicts our assumption that $((p, x_N)_{i \in N})$ is a Nash equilibrium. ■

The reader will note that we do not worry about the existence of equilibria issue (be it Nash or Walrasian), nor do we worry about the domain of definition of agents' preferences, assuming that they are defined on the whole commodity space V. Our main goal was to design a simple mechanism (the strategies are pairs of "prices-allocations") whose outcomes be precisely the feasible allocations.

Bibliographic Comments

The concept of Nash equilibria appeared in Nash (1951). Luce and Raiffa (1957), Moulin (1981b), Myerson (1991) and others discuss this conceptextensivly. The reader may consult van Damme (1987) and Myerson (1991) for presentations of finer equilibrium concepts (refinements).

Gurvich (1975) launched the investigation of Nash-consistent mechanisms, (he called them solvable positional forms). He proved the important Theorem (2.2.3); the proof given here appears to be simpler than the original one.

The issue of implementation was initiated by Maskin (1977). He noticed the close relation between Nash implementability and monotonicity and designed the implementation mechanism from Section 2.4. The concept of strong monotonicity and its relation to Nash implementation are elaborated by Danilov (1992). Another solution (applicable not only to the universal environment) of the Nash-implementation problem was presented in Moore and Repullo (1990); it influenced our exposition of the two participants' case of Section 5.

There are many results about non-cooperative implementation, related to other solution concepts (refinements): subgame perfect equilibrium (Moore and Repullo (1988), Abreu and Sen (1990)), backward induction (Dutta and Sen (1990)), dominance solvability (Moulin (1979, 1983)). Roughly speaking, these results claim that any SCC can be implemented. See also a survey by Moulin (1982). Implementation via strong (or coalitional) equilibria will be considered in Chapter 5.

The optimality of Nash equilibria issue has interested investigators for a long time. On the one hand, there is no reason to expect that selfish behavior of agents might always lead to the social optimum. On the other hand, Walrasian equilibrium (which is essentially a Nash equilibrium) is efficient. The concept of acceptable mechanism was introduced as a possible answer to the Gibbard-Satterthwaite theorem. (Peleg (1978) provided another interesting answer, see Chapter 5). Hurwicz and Schmeidler (1978) showed that in the case of two participants, any acceptable mechanism is dictatorial (see also Dutta (1984)). When there are three or more participants, there exist rather many acceptable mechanisms. Two interesting "acceptable (with respect to considering an economic environment)" mechanisms are presented in Walker (1981) and Hurwicz (1979) (see also Kim (1993) and Tian (1993)); we discuss them in Appendix 3.A2 of Chapter 3.

3. Strategy-proof Mechanisms

In this chapter we examine strategy-proof mechanisms, i.e., mechanisms that endow every agent with the best (called dominant) strategy for each permissible preference profile. Based on the revelation principle, we construct, for every strategy-proof mechanism, an equivalent direct non-manipulable mechanism. The key characteristic of such a mechanism is the agent's effective region in the set of outcomes. From this point of view, we study the structure of non-manipulable mechanisms in both the universal and the single-peaked environments (Sections 3.1 and 3.2). The convex structure of the outcome set yields an affine environment and allows us to mix strategy-proof mechanisms. In Section 3.3 we conjecture that any non-manipulable mechanism (within an affine environment) is a probability mixture of duplet and unilateral non-manipulable mechanisms. In the following two sections, we study the properties of Groves mechanisms in transferable environments, in particular, the issue of efficiency. We present some efficiency evaluations and efficiency criteria for Groves mechanisms.

We conclude the chapter with two Appendices. Appendix 3.A.1 contains comparative efficiency data for both Groves and Clark mechanisms, depending on the number of agents. Appendix 3.A.2 elaborates upon the Walker and Hurwicz concept of acceptable mechanisms in a transferable environment.

3.1 Dominant Strategies. The Revelation Principle

(3.1.1) One of the main drawbacks of Nash equilibrium is the difficulty for participants to find an equilibrium, even if they know that it exists. For example, in discussing the Maskin mechanism, we saw that finding an equilibrium essentially implied agents' knowledge of each others' preferences. Mechanisms theory addresses cases in which agents know for sure their own preference, while possibly having a vague idea about others' preferences. This is quite a different case from that in which the preferences of all agents are known.

Thus we would not expect that agents acting independently would be able to attain a Nash equilibrium. More realistically we expect that some preliminary negotiations will take place. The participants send trial messages, possibly changing them if they find better ones and so on. In other words, an iterative procedure precedes the attaining an equilibrium. In fact, this iterative procedure amounts to transiting from an initial mechanism $\pi : \prod_i S_i \to A$ to some "informational" extension $\pi^* : \prod_i S_i^* \to A$, of the mechanism π whose strategies are the strategies of negotiation and exchange of information between agents. But we have to deal with the same issues about equilibrium and etc. for the mechanism π^*.

There is another approach where an agent, not knowing the behavior of any of his partners, follows his subjective beliefs about other agents. This unties the system and allows us to model the behavior of every agent individually based on his preferences and beliefs. In this approach, difficulties arise with the forming of beliefs corresponding to the real behavior of other agents.

However, a large part of these difficulties disappears when an agent happens to have a dominant strategy, i.e., a strategy which is optimal for any strategies of the other agents. Due to this property, dominant strategies are sometimes called absolutely optimal. The concept of dominant strategy is not very interesting for game theory. An agent endowed with such a strategy is in fact not a player because he only needs to stick to his dominant strategy. (Of course, he might decide not to use it, but then he would be pursuing some other kind of goal, differing from utility maximization. There are other reasons which can explain deviating from dominant strategies, such as coalitional effects or threats. However, if agents do not communicate with each other, then their using dominant strategies seems quite plausible.) On the other hand, the notion of dominant strategy is of great interest for the theory of social choice mechanisms. Indeed, one of its main tasks consists in constructing mechanisms in which agents have dominant strategies in order to forecast both the behaviour of participants and the resulting outcome.

We now give a few formal definitions. Let $\pi : S_N \to A$ be a social choice mechanism.

(3.1.2) Definition. A strategy s_i^* of agent i, endowed with a preference R_i, is *dominant* if

$$\pi(s_i^*, s_{N\setminus i}) R_i \pi(s_i, s_{N\setminus i})$$

for any $s_i \in S_i$ and $s_{N\setminus i} \in S_{N\setminus i}$.

It is important here that the optimal strategy s_i^* be independent of $s_{N\setminus i}$, although it may depend on R_i. Usually (and this is almost a rule) there are either no dominant strategies or several. Denote the set of dominant strategies $Dom_i(\pi, R_i)$ or simply $Dom_i(R_i)$.

(3.1.3) Definition. Let $R_N = (R_i)$ be a preference profile. A strategy profile $s_N^* = (s_i^*)$ is a *dominant strategy equilibrium* or (*DS-equilibrium*) if $s_i^* \in Dom_i(\pi, R_i)$ for every agent $i \in N$.

Note that a DS-equilibrium is a Nash equilibrium, but the reverse is not true.

(3.1.4) Definition. A mechanism π is *dominant strategy mechanism* (or DS-mechanism, or strategy-proof mechanism) in the environment $Prof \subset \mathbf{L}^N$ if for any feasible profile of preferences $R_N \in Prof$, the game $G(\pi, R_N)$ has at least one DS-equilibrium.

Clearly the DS-property for mechanism survives any restriction of the environment. On the contrary, the wider the environment the lesser DS-mechanisms, and it is all the more easy to describe them all. The number of DS-mechanisms is smallest in the universal environment $\mathbf{L}^N$.

(3.1.5) The Revelation Principle. This principle enables us to construct from any general form DS-mechanism an "equivalent" direct DS-mechanism. In what follows, we shall assume an environment of the form $\mathcal{P}_N = \mathcal{P}_1 \times ... \times \mathcal{P}_n$, where $\mathcal{P}_i \subset \mathbf{L}$ is a domain of "feasible" preferences of agent i. Assume that a mechanism $\pi : \prod_i S_i \to A$ is strategy-proof in the environment $\mathcal{P}_N = \prod_i \mathcal{P}_i$. This means that the set $Dom_i(R_i)$ is non-empty, for any $R_i \in \mathcal{P}_i$. And we choose an element $d_i(R_i)$ in $Dom_i(R_i)$. We construct in this way a series of mappings $d_i : \mathcal{P}_i \to S_i$, $i \in N$. Let us now construct the new mechanism

$$\pi^* = \pi \circ d_N : \prod_i \mathcal{P}_i \overset{\prod d_i}{\to} \prod_i S_i \overset{\pi}{\to} A.$$

The feasible preferences become strategies for the mechanism π^*, so this is a direct mechanism. It is easy to understand that π^* is also a DS-mechanism. More exactly, agent i (endowed with the preference R_i)'s dominant strategy is R_i in the mechanism π^*. Thereby the mechanism π^* induces participants to reveal their true preferences, i.e., it is *non-manipulable*. In some sense, the mechanisms π and π^* are "equivalent" (we shall not define precisely this concept here). The *revelation principle* is the process which leads from a DS-mechanism π to the direct non-manipulable associated mechanism π^*. From now on, when we speak of DS-mechanisms we shall mean simply non-manipulable mechanisms.

Remark. When we defined the mechanism π^*, we arbitrarily chose a selector d_i in Dom_i. When preferences are linear orders, the resulting mapping π^* does not depend on this arbitrary choice of a selector. In the more general case of weak orders, the mapping π^* may depend on the choice of a selector. Nevertheless, we shall consider only direct mechanisms.

(3.1.6) The Single Agent Case. Let us examine for a while the construction of non-manipulable mechanisms in the single agent case. We might wonder why we should need any mechanisms in this specific case. The issue is the following. Let $\pi : \prod_i S_i \to A$ be a direct DS-mechanism with several participants. Supposing the strategies of all agents (different from agent i) are fixed, we end up with a DS-mechanism $\pi(\cdot, R_{N \setminus i}) : \mathcal{P}_i \to A$ depending only on agent i. And it is useful to understand its construction.

Let $\mathcal{P} \subset \mathbf{L}$ be a domain of preferences and

$$\pi : \mathcal{P} \to A$$

be a single agent non-manipulable mechanism, set within the environment $\mathcal{P}$. Its most important characteristic is the set $Z = \pi(\mathcal{P})$. We will call the set Z agent i's *effective region* (or option set, as S.Barbera calls it). The set Z plays a crucial role in the recovering (uniquely) of mechanism π; indeed for any $R \in \mathcal{P}$

$$\pi(R) = \max R \mid Z.$$

This is simply a rephrasing of the non-manipulability condition: for any $R, R' \in \mathcal{P}$ the relation $\pi(R)R\pi(R')$ should take place.

This remarkable feature prompts a method for constructing any arbitrary non-manipulable mechanism. Take a set $Z \subset A$ and define the mechanism $\pi_Z : \mathcal{P} \to A$ by the formula

$$\pi_Z(R) = \max R \mid Z$$

for $R \in \mathcal{P}$. Obviously the mechanism π_Z is non-manipulable.

Note, incidentally, that the set $\pi_Z(\mathcal{P})$ may turn out to be smaller than Z. In short, different sets $Z \subset A$ may yield the same mechanisms. However this does not matter too much; one can choose, for example, the smallest or the largest among these sets Z. Thus the single agent case is clear-cut.

Dealing with weak orders, there is a caveat: $\pi(R)$ need only be one of the maximal points of the relation $R \mid Z$, i.e. $\pi(R) \in \max R \mid Z$. And therefore, in this case, the set Z does not allow a recovering of the mechanism π in a unique fashion, although it remains its most important characteristic.

Having examined the single agent case, we return to the general case. Let $\pi : \prod_{i\in N} \mathcal{P}_i \to A$ be a non-manipulable mechanism. We fix an agent i and a preference profile $R_{N\setminus i} \in \mathcal{P}_{N\setminus i}$. As was explained above, any non-manipulable single agent mechanism $\pi(\cdot, R_{N\setminus i})$ determines and is determined by the set $Z = Z_i(R_{N\setminus i})$, which in turn depends on the environment $R_{N\setminus i}$. In particular, the mechanism π is completely given by the family of sets $(Z_i(R_{N\setminus i}))$ where $R_{N\setminus i}$ runs the set $\mathcal{P}_{N\setminus i}$. We could almost think that we have our answer since we are able to describe any DS-mechanism! Alas, this is all too hasty a conclusion! Indeed, up to now we used solely the fact that agent i has a dominant strategy. But we need to consider the remaining agents. And in fact the requirement of non-manipulability for other agents is expressed through some strong interdependence of sets $Z_i(\cdot)$ for the environment. The whole matter and the whole mystery, the key to understanding structure of DS-mechanisms, lies within this dependence. The following sections will clarify these matters. Meanwhile, we illustrate this method in the universal environment setting.

(3.1.7) Universal Environment. Let every $\mathcal{P}_i = \mathbf{L}$. Then, a non-manipulable direct mechanism is simply a non-manipulable SCF, with which we became familiar in Chapter 1. We already know the answer: this non-manipulable mechanism is either a duple or a unilateral SCF (the Gibbard-Satterthwaite theorem (1.A2.7)). However, for the time being, for the sake of simplicity, we will assume that there are only two agents (in any case, this does not change essence of the matter). Here the environment consists only in giving a preference R_2 for agent 2. This yields a family $(Z(R)), R \in \mathbf{L}$, of non-empty subsets of A; the mechanism π acts by the formula

$$\pi(R_1, R_2) = \max R_1 \mid Z(R_2).$$

If all sets $Z(R)$ are singletons, then agent 1 does not at all influence the outcome and the mechanism π is unilateral with a "dictator", namely agent 2. Assume now that for some R the set $Z(R)$ contains more than one element. If $Z(R)$ does not change with R, then the mechanism is unilateral and agent 1 is the dictator. We consider in more detail the dependency of $Z(R)$ on R. The non-manipulability condition implies the following properties for this dependency: suppose that we move from the linear order R to R', then

a) if an element disappears from $Z(R)$, then all elements ranking higher with respect to the order R also disappear with it;

b) new elements appearing in $Z(R')$ with respect to those in $Z(R)$, are ranked below any element of $Z(R)$.

For example, if $R = \{x \succ y \succ z \succ a \succ b \succ c \succ d \succ e)$ and $Z(R) = \{y, a, b\}$ then $Z(R')$ could look like $\{a, b, d, e\}$, but not like $\{x, a, d\}, \{y, b, c\}$ or $\{z, a, b\}$.

Indeed, suppose that contrary to assertion a) an alternative x disappeared from $Z(R)$, but that an alternative y situated higher than x remained. Let $R_1 = (x \succ y \succ ...)$. Then $\pi(R_1, R) = x,\ \pi(R_1, R') = y$ and yRx. This implies that agent 2 is able to exert some manipulation: the untruthful preference R' is clearly more advantageous (for agent 2). Assertion b) is proved analogously.

Now we turn to proving the fact that $Z(R)$ is independent of R.

It is convenient to proceed to the comparison of the sets $Z(R)$ and $Z(R')$, by moving from order R to R' step by step, namely starting with R and switching only two elements at every step. Assume the alternative x precedes the alternative y, in the ordering R :

$$R = (... \succ x \succ y \succ ...).$$

The transition from the order R to an order $R' = (... \succ y \succ x \succ ...)$ where the other alternatives stay at their places is called a *switching*. There is now a standard result from permutation group theory, which states that one can proceed from any one order to another by a finite number of switchings. With this in mind, our task is slightly simpler: we need only compare the sets $Z(R')$and $Z(R)$ when R' is obtained from R after switching two given elements. If $Z(R') \neq Z(R)$ then (w.l.o.g.) there is $z \in Z(R) \backslash Z(\acute{R}')$. According to a) $Z(R')$ lies below z with respect to R, and according to b) $Z(R')$ is above with respect to R'. This is possible (at $|Z(R)| > 1$) only if $z = x$,

$Z(R) = \{x, y\}$, $Z(R') = \{y\}$. Thereafter we conclude that if $|Z(R)| \geq 3$ the set $Z(R)$ remains unchanged and agent 1 is the "dictator"; if $|Z(R)| = 2$ the mechanism is a duple (whose efficiency region can sometimes be reduced to one element).

3.2 Single-Peaked Environment

(3.2.1) We now undertake a more detailed study of DS-mechanisms for a few specific environment classes. As we have already mentioned, we ordinarily obtain a narrower environment, by imposing some additional structure (with which the agents' preferences are to some extent compatible) on the set of alternatives. Then the appearance of new DS-mechanisms (other than the dictator or the duple mechanisms) follows from new opportunities for compromising, for softening out extreme viewpoints. The single-peaked environment is the most significative example of such cases.

We have a single-peaked environment when the set of alternatives is somewhat similar to a "line", or possesses a "linear structure". If we name (conventionally) one end of this "line" the left and the other the right, this amounts to defining a linear order R^0 on A, which expresses the idea of moving from the left to the right (however the choice of left and right gives us two possible ways of doing it). We choose a linear structure on A. A preference $R \in \mathbf{L}(A)$ is *single-peaked* or *unimodal*, if moving from alternative $\max R$ either to the left or to the right, the utility function is decreasing. In terms of the utility function, the following picture can be drawn (Fig.2).

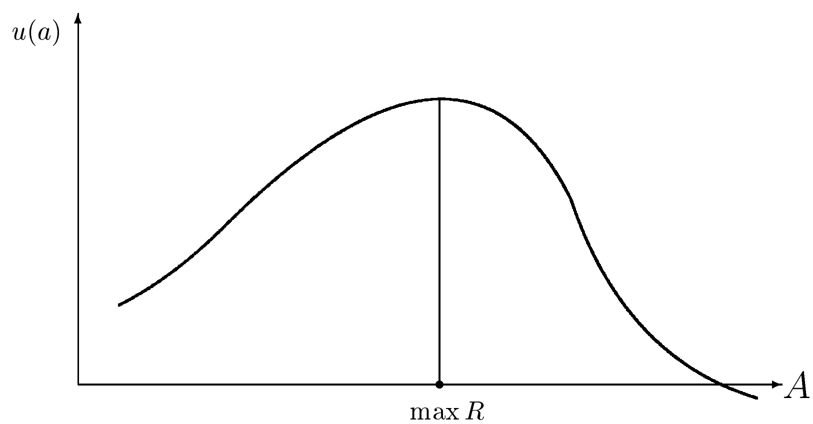

Fig. 2. A unimodal utility function.

The alternative $\max R$ is the happiness or the *bliss point* of R. In many interesting cases, preferences are indeed single-peaked. For example, we can always imagine representing political parties on a "left-right" scale. Each voter has a most acceptable political program; moves to the left or to the right (of this most acceptable program) correspond to decreasingly attractive programs. One can devise preference upon indices (such as defense expenditure indices and so on) in the same fashion.

We denote by $\mathcal{U}$ the set of all single-peaked preferences. The single-peaked environment is defined as $\mathcal{U}^N \subset \mathbf{L}^N$. Single-peaked preferences are not determined uniquely by the giving of a happiness point, but it is a crucial item of information. Henceforth we consider mechanisms for which the agents' messages consist in giving their happiness points, not their whole preference, and non-manipulability will be understood in relation to these happiness points. Thus a mechanism, in this section, will be a mapping

$$\pi : A^N \to A.$$

The non-manipulability property is expressed as follows: for every agent $i \in N$ and for every single-peaked preference $R_i \in \mathcal{U}$, $\max R_i$ belongs to $Dom_i(\pi, R_i)$. In other words, for every agent, reporting the true happiness point is a dominant strategy. Can we find non-manipulable mechanisms, other than dictatorial mechanisms (a mechanism being dictatorial if $\pi(x_N) = x_i$ for a fixed "dictator" i) ? It turns out that we can find some (and sufficiently many of them). We now provide a few examples.

(3.2.2) Example. The *classical median.* Let there be an odd number of agents and let them name alternatives (points) $x_i \in A, i \in N$. Among these points, there exists one point x^* such that both to its left and to its right there are $< n/2$ points x_i. This point x^* is unique and is called *the median* of the family of points (x_i). We construct the classical median mechanism by associating to each bundle (x_i) its median x^*.

We affirm that this mechanism is non-manipulable. Pick an arbitrary participant i and assume that his bliss point x_i is located to the left of x^*. He is not able (through any strategy) to pull the point x^* to the left, therefore closer to his bliss point. He is able to move the median outcome x^* to the right, just by setting his "bliss-point signal" x_i' to the right of x^*, but it is not interesting a strategy for him.

(3.2.3) Example. *Left (or right)-dictator mechanism.* The point $\min R^0 \mid \{x_1, ..., x_n\}$, located at the extreme left of all points x_i, is chosen to be the outcome of a "left-dictator" mechanism $\pi(x_1, ..., x_n)$. The non-manipulability of this mechanism in a single-peaked environment is proved as above in Example (3.2.2).

We define the "right-dictator" mechanism similarly by:

$$\pi(x_N) = \max R^0 \mid \{x_1, ..., x_n\}.$$

These three mechanisms are anonymous. How do we construct anonymous non-manipulable mechanisms in general? The following example (Moulin (1980), see also Moulin (1985), p.23) gives us the answer.

(3.2.4) Example. Fix $n+1$ alternatives $a_0, \ldots, a_n$, in the set A. We define

$$\pi(x_1, \ldots, x_n) = \text{median of } (x_1, \ldots, x_n,\ a_0, \ldots, a_n).$$

Again it is not difficult to verify that this mechanism is non-manipulable.

There are many non-anonymous and non-manipulable mechanisms. Moreover, we surely would hope to have a more complete picture of the set of all such mechanisms. As it appears we can, we now show how to construct any arbitrary non-manipulable mechanism.

(3.2.5) The Median Operation. We start with a general remark. The setting of an additional structure on the set A often permits us to construct new mechanisms from known ones. We explain how through an example, set in the single-peaked environment. The remarkable feature of a linear structure on a set A, is that picking any three alternatives x, y and z, one of them will be located *between* the two others. This "middle" alternative is called the *median* of the triple (x, y, z) and is denoted by $\mu(x, y, z)$. The inf and sup operations (with respect to the order R^0) can be rephrased using μ as follows:

$$\inf(x, y) = \mu(\Lambda, x, y), \quad \sup(x, y) = \mu(x, y, \Pi),$$

where $\Lambda = \min R^0$ is the left end of A,and $\Pi = \max R^0$ is the right end. The median, by the way, can also be expressed through inf and sup:

$$\mu(x, y, z) = \sup(\inf(x, y), \inf(y, z), \inf(x, z)).$$

The nice thing about using the median formulation is that it does not require that we first determine a (out of the two possible directions) direction on A, i.e. we need not define which end will be called the left and which the right.

If now we take three mechanisms $\pi_j : A^N \to A, j = 1, 2, 3$, then we can conceive a new mechanism $\pi = \mu(\pi_1, \pi_2, \pi_3)$ as follows,

$$\pi(x_N) = \mu(\pi_1(x_N), \pi_2(x_N), \pi_3(x_N)).$$

(3.2.6) Lemma. *If π_1, π_2, π_3 are non-manipulable, then $\pi = \mu(\pi_1, \pi_2, \pi_3)$ is non-manipulated.*

The proof boils down to the following straightforward assertion. Let $a, y_1, y_2, y_3, z_1, z_2, z_3$ be seven points from A. Assume $y_i \in [a, z_i]$ for $i = 1, 2, 3$, where $[x, y]$ denotes a segment in A with ends x and y. Then $\mu(y_1, y_2, y_3) \in [a, \mu(z_1, z_2, z_3)]$.■

So given any three DS-mechanisms, we construct a new DS-mechanism using the median operation. We now show how this composition principle helps understanding the construction of any DS-mechanisms within a single-peaked environment.

(3.2.7) Structure of Non-Manipulable Mechanisms. We begin with the single agent case. In Section 1 we argued that a non-manipulable mechanism π is determined by its effective region $Z = \pi(A)$. This is true whatever the environment. In this case, the peculiarity is that the set Z is a segment with respect to the linear structure of A.

Indeed, assume conversely that there exists a point $x \notin Z$ such that there are points from Z located to its right and to its left. Assume the point $\pi(x)$ lies at the right of x, and that there is another point $z = \pi(x'), x' \in A$, lying at the left of x. Now let R be a single-peaked preference with a bliss point x, such that $zR\pi(x)$ (see Fig.3).

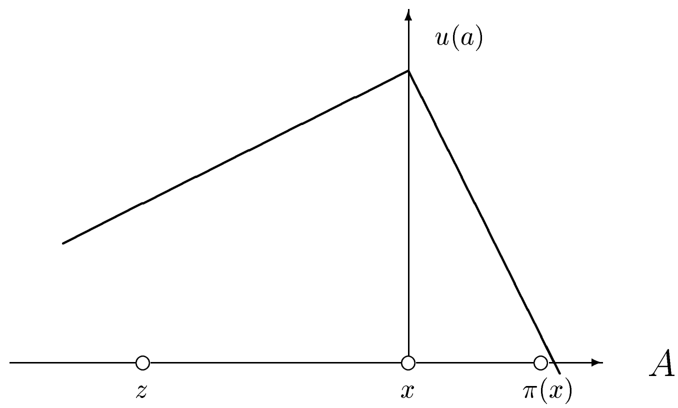

Fig. 3

In this case, the agent is better of not naming his true bliss point x, but the point x' instead. Contradiction.

Thus we proved that Z is a linear segment. Its ends are obviously $\pi(\Lambda)$ and $\pi(\Pi)$. More generally, $\pi(x)$ is the closest point to x in Z. This assertion can be rephrased more compact as

$$\pi(x) = \mu(\pi(\Lambda), x, \pi(\Pi)).$$

Thus in the single agent case, any non-manipulable mechanism π is the median of the three (non-manipulable) mechanisms: two of them are constant with outcomes $\pi(\Lambda)$ and $\pi(\Pi)$ and one is dictatorial (or identical $\pi(x) \equiv x$).

Now we examine the n agents' case. Let $\pi : A^N \to A$ be a non-manipulable mechanism. Single out agent 1. From the single agent case, for any bundle $x_{N\setminus 1} \in A^{N\setminus 1}$ we have

$$\pi(x, x_{N\setminus 1}) = \mu(\pi(\Lambda, x_{N\setminus 1}), x_1, \pi(\Pi, x_{N\setminus 1})).$$

This means that our mechanism π is the median of three mechanisms: a dictatorial mechanism for agent 1 and two non-manipulable mechanisms

$\pi(\Lambda, \cdot)$ and $\pi(\Pi, \cdot)$ depending on the strategies of $n-1$ agents. The previous considerations hold for both these mechanisms and so on. At the end of all iterations, the mechanism π will be represented as a repeated application of the median operation to a few dictatorial and constant mechanisms. For example, in a case of two agents,

$$\pi(x_1, x_2) = \mu(\mu(\pi(\Lambda, \Lambda), x_2, \pi(\Lambda, \Pi)), x_1, \mu(\pi(\Pi, \Lambda), x_2, \pi(\Pi, \Pi))).$$

The constant mechanisms that appear in this decomposition are given by the values of the mechanism π for "extremal" profiles, when a section of participants (a coalition $K \subset N$) opts for the right end Π and the remainder (the coalition $\overline{K}$) opts for the left end Λ. Denote such a profile by $(\Pi_K, \Lambda_{\overline{K}})$ and set

$$\alpha(K) = \alpha_\pi(K) = \pi(\Pi_K, \Lambda_{\overline{K}}).$$

We see that a non-manipulable mechanism is defined uniquely by its family of "constants" $(\alpha(K)), K \subset N$. This family is not arbitrary; clearly, the more "at right" the profile the more "at right" the outcome. In other words, if $K \subset K'$, then $\alpha(K) \leq \alpha(K')$, where $\leq$ represents the order from the left to the right in A. Understandably, this is the unique condition on $(\alpha(K))$.

We sum up this result in the following theorem.

(3.2.8) Theorem. *Any non-manipulable (for the single-peaked environment) mechanism $\pi : A^N \to A$ can be obtained by applying (iteratively) the median operation to dictatorial and constant mechanisms.*

The mechanism π is then uniquely determined by its values at extremal profiles. So there exists a bijective correspondence between the set of non-manipulable mechanisms $\pi : A^N \to A$ and that of monotone mappings $\alpha : 2^N \to A$.

Moulin (1980) gives the following more explicit expression of a mechanism π_α in terms of the family of constants $(\alpha(K))$,

$$\pi_\alpha(x_1, ..., x_n) = \sup_{K \subset N} [\inf_{i \in K} \{x_i, \alpha(K)\}].$$

For example, taking two agents only,

$$\pi(x_1, x_2) = \sup\{\alpha(\emptyset), \inf(x_1, \alpha(1)), \inf(x_2, \alpha(2)), \inf(x_1, x_2, \alpha(1, 2))\}.$$

Recently, Berga (1998) generalized Moulin's maxmin formula to the case of single-plateaued environment.

We make here a few remarks.

First, note that when the set A consists of two alternatives, then any order on A is single-peaked, and the difference between universal and single-peaked environments disappears. In particular, this theorem yields the previously mentioned (1.3.10) result by Monjardet (1978) about representation of majorities systems through dictatorial families using the median operation.

The second remark deals with finding out the conditions for which a non-manipulable mechanism π is efficient. It is clear that the following criterion is true: a mechanism π is efficient if and only if $\pi(x, ..., x) = x$ for any $x \in A$, or when $\pi(\Lambda, ..., \Lambda) = \Lambda$, and $\pi(\Pi, ..., \Pi) = \Pi$.

Thirdly, in the single-peaked setting, a non-manipulable mechanism is also coalitionally non-manipulable. In other words, no coalition can improve upon the outcome for all its members. This is a specific property of single-peaked environments. Note that the Oil Producing Cartel (OPEC) uses mechanisms of this type for decision-making when a production level has to be chosen (Border, Jordan (1983)).

The fourth remark concerns the relation between anonymity and coalition size. The associated constants $\alpha(K)$ of an anonymous mechanism π depend on $| K |$ only. Thus it suffices to fix $n+1$ constants $a_0, a_1, ..., a_n$ (cp. Example (2.4)) instead of the 2^n required constants $\alpha(K)$.

Lastly, note that we nowhere made use of the set A's finiteness. In fact, the crucial point is that A possess a linear structure. Thus the above-mentioned results remain valid when A is the real line $\mathbf{R}$. It is slightly inconvenient that for the real line it is difficult to talk about a left and a right end, but this is easily rectified by adding formally $-\infty$ and $+\infty$ to $\mathbf{R}$.

(3.2.9) Generalization on Trees. The replacement of a linear structure on A by a more general "tree" structure yields yet more interesting generalizations. Recall that a *tree* is a connected (non-oriented) graph without cycles. Given two nodes x and y of the tree, we will denote by $[x, y]$ the minimal connected subgraph in A containing x and y. Intuitively, this is the shortest path between x and y.

The notion of single-peakedness is readily transferable to trees. A linear order R is *single-peaked* if it is single-peaked on any "segment" $[x, y] \subset A$. Else, the further we move from the bliss point $a = \max R$ the lower the agent's utility; more formally: if $x \in [a, y]$ then xRy.

(3.2.10) Example. Suppose that w is a probability measure on the set of participants N. For any strategy profile $x_N = (x_i)$, we form a function $p : A \longrightarrow \mathbf{R}$, $p(a) = \sum_i \rho(a, x_i)w(i)$, where $\rho(a, x_i)$ is the distance (on the tree) between the node a and the node x_i. Suppose, for simplicity, that the weight $w(K)$ of any coalition K is not equal to $1/2$. Then the function p has a unique minimum; the minimum point is denoted by $\mu(x_N)$. This yields the mechanism $\mu : A^N \longrightarrow A$. One can easily prove that the mechanism μ is non-manipulable for the single-peaked environment on the tree A.

The crucial (in our view) property of a tree is that the notion of the median of three points is well-defined.

(3.2.11) Lemma-definition. *For any three nodes x, y, z of a tree, the intersection $[x, y] \cap [y, z] \cap [z, x]$ consists exactly of one node. This node is called the median and is denoted by $\mu(x, y, z)$.*

It suffices to look at Fig. 4 to prove this lemma. The unique difference with respect to the linear case, considered above, is that the median does not have to coincide with any one of the nodes x, y or z.

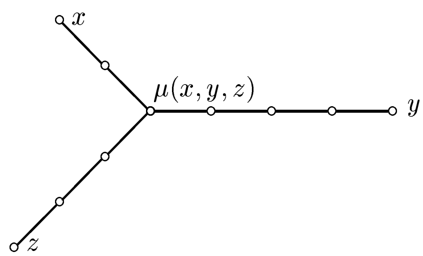

Fig. 4. A tree and the median of the nodes x, y, z.

As in the linear case, median mechanisms can always be built from any three mechanisms: $\pi = \mu(\pi_1, \pi_2, \pi_3)$. The mechanism π is non-manipulable as soon as π_1, π_2 and π_3 are non-manipulable. Here again, any non-manipulable rule can be obtained from dictatorial and constant rules using median operations.

The case of a single agent again plays a pivotal role. Let $\pi : A \to A$ be a non-manipulable rule and $Z = \pi(A)$. As in the linear case, it is easy to establish that Z is a convex subset of A, i.e. $[x, y] \subset Z$ for any $x, y \in Z$. Besides, it is obvious that $\pi(x)$ is the closest point to x in Z.

Further we can express this last property through medians, thereby generalizing the formula $\pi(x) = \mu(\pi(\Lambda), x, \pi(\Pi))$ for the linear case. To do so we need a generalization of the notions of left and right ends, used previously. A node a of a tree A, is thus called *extremal* if $A \backslash \{a\}$ is still a tree. We denote by ∂A the set of extremal points of the tree A. $Z = \pi(A)$ is obviously the convex hull of $\pi(\partial A)$.

Now we explain how to express $\pi(x)$ in terms of x and of the nodes $\pi(a)$, $a \in \partial A$. To do so, we start by enumerating arbitrarily all extremal points $\partial A = \{a_0, ..., a_m\}$ of A and determine by induction

$\pi_1(x) = \mu(x, \pi(a_0), \pi(a_1))$,

$\pi_2(x) = \mu(x, \pi_1(x), \pi(a_2))$,

....

$\pi_m(x) = \mu(x, \pi_{m-1}(x), \pi(a_m))$.

We assert that $\pi(x) = \pi_m(x)$. This can be seen considering the two following facts. First, the point $\pi_m(x)$ belongs to any segment $[x, \pi(a_j)], j = 0, ..., m$. This assertion is obtained inductively: for any j the point $\pi_j(x)$ belongs to any segment $[x, \pi(a_0)], ..., [x, \pi(a_j)]$. Second, the point $\pi_m(x)$ belongs to Z.

Thus $\pi(x)$ can be expressed by a somewhat tedious formulation using the median operation through $\pi(a_j)$ and x. By induction and as in the linear case, we can figure out a generalization for any number of agents.

(3.2.12) Theorem. *Any non-manipulable rule $\pi : A_N \to A$ on the tree A can be obtained using the median of some dictatorial and constant rules; it is uniquely determined by its values for extremal profiles $(\partial A)^N$.*

(3.2.13) We do not give here an explicit formula *a la* Moulin. But we provide (without proof) a necessary and sufficient condition on the mapping $\pi : (\partial A)^N \longrightarrow A$ which can be extended to suit a non-manipulable mapping $A_N \to A$. The giving of an extremal profile $N \to \partial A$ amounts to allocating (somehow) agents to extremal points. Imagine now that an agent i moves from an extremal point a to another extremal point a'; how does $\pi(\cdot)$ change? There are two possible cases. In the first, the outcome $\pi(a)$ does not belong to the segment $[a, a']$; in this case the outcome does not change, $\pi(a) = \pi(a')$. In the second, $\pi(a) \in [a, a']$; in this case the outcome $\pi(a')$ also belongs to $[a, a']$ and can move to a'. This condition is a generalization of the monotonicity requirement of $\alpha(K)$ in Theorem (3.2.8).

A last remark. A tree can be viewed as realizing the idea of a compromise. Extremal alternatives (from ∂A) can be considered as basic or initial non-structured alternatives. Social choice on this set is impossible. However embedding the set ∂A into a tree A by adding to ∂A a few compromise alternatives makes the solving of a social choice problem more satisfactory. Moreover in some sense we think that we can not do better.

3.3 Linear Environment

(3.3.1) Lotteries. One other way to express the idea of compromise is to take mixtures of alternatives (in a probabilistic sense) or lotteries. Von Neumann was the first to propose working not only with "pure" alternatives, but also with their probability mixtures, launching a tradition which is now firmly anchored in the practice of mathematical economics (especially in game theory). Let A be a finite set of initial "pure" alternatives. A *lottery* on A is a probability measure on A, or in other words, a formal combination $\mu = \Sigma_{a \in A} \mu(a) a$, where all $\mu(a) \geq 0$ and $\Sigma_{a \in A} \mu(a) = 1$. Of course, one can devise probability measures on more general or complex sets (such as, for instance, infinite sets), but the technical details involved with such constructions would lead us too far away from our purpose here.

A mechanism, whose outcomes are lotteries on A, can be described as follows. An alternative from A is determined, through the rolling of a dice (or for the matter, any more modern probabilities generator), after the collecting of all agents' messages s_i and the forming of a lottery $\pi(s_N) = \Sigma_a \mu(a) a$. Here $\mu(a)$ is the probability of occurrence of alternative a.

To evaluate the outcomes of this kind of mechanism, the utilities of agents should be defined on lotteries. Denote by $\Delta(A)$, the set of all lotteries on A; geometrically, this set can be represented by a $(|A| - 1)$-dimensional simplex whose vertices correspond (one to one) to the elements of A. It is usual to assume that preferences on lotteries satisfy the von Neumann and Morgenstern axioms and are represented by affine functions $u : \Delta(A) \to \mathbf{R}_+$ Such a function is defined by its values for vertices of the simplex $\Delta(A)$, i.e. for elements of A. Given the numbers $u(a), a \in A$, the "utility" of a lottery $\mu = \Sigma\mu(a)a$ is given by the formula

$$u(\mu) = \Sigma_{a \in A}\mu(a)u(a).$$

Note that these preferences on $\Delta(A)$ represented by affine functions are not linear orders, but weak orders. Practically we now have to deal with indifference classes, and this brings about quite a few mathematical inconveniencies (however nothing that can't be dealt with).

(3.3.2) For most of the material we present in this section, it is inessential that $\Delta(A)$ be a simplex, but it is important that it be a convex set. Therefore in this section we shall work with the following set-up. The space of alternatives is an infinite convex set V (this is why we do not call it A, to draw the reader's attention to the fact that it is infinite). A preference on V is called *affine,* if it is represented by an affine utility function on V. The latter means that $u(\alpha x + (1-\alpha)y) = \alpha u(x) + (1-\alpha)u(y)$ for any $x, y \in V$ and $0 \leq \alpha \leq 1$. The set of all affine preferences is denoted $Aff(V)$; we call *affine environment* the environment $Aff(V)^N$.

The definition of strategy–proof mechanisms in an affine environment is straightforward. As we noted in Section 3.1, "equivalent revealing" direct mechanisms are not uniquely determined. Nevertheless, in what follows we shall deal mainly with mechanisms of the following form:

$$\pi : Aff(V)^N \to V.$$

Such mechanisms are called non-manipulable in a affine environment, if the following inequality is fulfilled,

$$u_i(\pi(u_N)) \geq u_i(\pi(u_i', u_{N\setminus i}))$$

for any profile of preferences $u_N = (u_1, ..., u_n) \in Aff(V)^N$, and any agent $i \in N$ and any $u_i' \in Aff(V)$.We do not distinguish here affine preferences and their affine representations.

(3.3.3) The affine environment (as the universal environment) possesses two "trivial" classes of non-manipulable mechanisms. The first one consists of dictatorial, or, more accurately, *unilateral* mechanisms. Fix an agent i, fix some subset $Z \subset V$ (an efficiency region) and define a mechanism π as follows,

$$\pi(u_N) \in \max(u_i \mid Z).$$

This mechanism depends only on the messages of agent i and indeed, fully favors him in Z. Obviously, this is a non-manipulable mechanism. Of course, the previous formula makes sense if the maximum is attainable. To ensure this, we assume everywhere that the convex set V has finite dimension, and the subset Z is compact.

We can already at this point sense an inconvenience due to the presence of indifference sets. The mechanism π is not defined uniquely by the set Z, which, by the way, we can take as being convex. In fact, the maximum of π on Z might not be unique. However, multi-valuedness of the maximum is not generic; in fact, the outcome $\pi(u_N)$ is determined uniquely in most cases (i.e., for "general" profile $u_N \in Aff(V)^N$). It seems that an adequate formalization is required in order to better understand the whole issue. One should most probably gather in a single equivalence class all the mechanisms which differ only with respect to very "specific" profiles. These subtleties are not essential to the matter under investigation, so we shall not dwell further upon them.

A second large class of non-manipulable mechanisms in affine environments consists in affine *duple* mechanisms. One dimensional convex sets, i.e., segments, are the analogue of two element sets in the affine environment. There are only two affine preferences on the line corresponding to the setting of two possible directions on this line (if we neglect the "total indifference" preference). We investigate these mechanisms more in detail.

(3.3.4) One-dimensional Mechanisms. Assume first that $V = \mathbf{R}$ (the real line). The three affine preferences on the real line are represented by the functions $u(x) = x, u(x) \equiv 0, u(x) = -x$. We denote them conventionally by the symbols $+, 0, -$. A direct mechanism thus takes the form:

$$\pi : \{-, 0, +\}^N \to \mathbf{R}.$$

This mechanism is non-manipulable if and only if the mapping π is monotone (where $- < 0 < +$ and the order on $\mathbf{R}$ is the natural order). We give a more concrete example.

(3.3.5) Example. Let there be two agents, $\pi(+,+) = 1$ and $\pi(\cdot,\cdot) = 0$ for all other profiles. This mechanism is non-manipulable and non-unilateral.

Generally, a mechanism $\pi : Aff(V)^N \to V$ is called *one-dimensional* or *linear* if its image $\pi(Aff(V)^N)$ lies into some line $L \subset V$. Non-manipulability of a one-dimensional mechanism π implies the following property: suppose that $(R_1, ..., R_n)$ is a profile of (affine) preferences, suppose that R_i' is another (affine) preference, such that the restrictions of R_i' and R_i to L coincide and are not such that all elements of L are equivalent, then $\pi(R_1, ..., R_i, ..., R_n) = \pi(R_1, ..., R_i', ..., R_n)$.

Thus the construction of general one-dimensional mechanism resembles that of the mechanism constructed above for $V = \mathbf{R}$. Namely, given a line $L \subset V$, the set of affine preferences can be divided into three classes. Preferences belong to either class, when their restriction to L yields either one of the directions (chosen on the line L) or indifference. We denote, as above, these classes conventionally by $+, -$, and *class* 0. A one-dimensional non-manipulable mechanism takes the following form:

$$Aff(V)^N \to \{-, \text{class } 0, +\}^N \xrightarrow{\sigma} L \subset V,$$

where σ is a monotone (in a natural sense) mapping.

As above, the main part of the mechanism is given by the mapping $\{-,+\}^N \to L$, although the latter does not account for the outcomes of the mechanism for those rare profiles, for which the preferences of some agent might fall into the class 0.

(3.3.6) Mixing Mechanisms. The above two non-manipulable mechanisms' classes do not exhaust all non-manipulable mechanisms, for a reason similar to that cited in (3.2.5). In fact, the convexity of the set V enables us to construct new mechanisms by forming convex combinations (or mixtures) of both alternatives and mechanisms. More exactly, let

$$\pi_t : Aff(V)^N \to V, \ t \in T,$$

be a finite family of mechanisms and $\lambda_t \geq 0$ be real numbers, such that $\Sigma_t \lambda_t = 1$. Then one can form a mechanism $\pi = \Sigma_t \lambda_t \pi_t$ by formula

$$\pi(u_N) = \Sigma_t \lambda_t \pi_t(u_t),$$

which is non-manipulable if all mechanisms π_t are so. Here are two examples.

(3.3.7) Example. *Random dictator.* Let $V = \Delta(A)$, i.e. we are in the lottery setting. We form a rule:

$$\pi(u_1, ..., u_n) = \Sigma_{i \in N} \frac{1}{n} \max(u_i \mid A).$$

This amounts to taking the arithmetic mean of dictator rules $\pi_i(u_N) = \max u_i \mid A$ (we do not pay attention to possible non-uniqueness of max). Of course, one could decide to give different weights to participants.

(3.3.8) Example. Analogously, one can take a mixture of duple rules. For simplicity, we consider the case of one agent and $A = \{x, y, z\}$. The rule π is given by the formula:

$$\pi(u) = \frac{1}{3} \max(u \mid \{x,y\}) + \frac{1}{3} \max(u \mid \{y,z\}) + \frac{1}{3} \max(u \mid \{x,z\}).$$

This is a non-manipulable *unilateral* rule whose efficiency region is depicted in Fig.5.

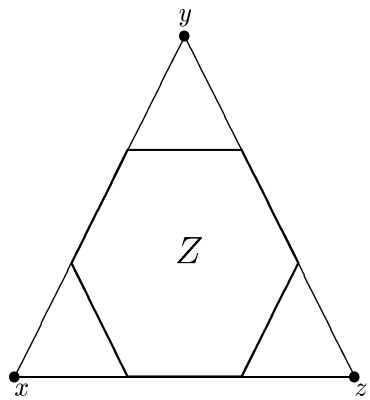

Fig. 5. The hexagon Z is the effectivity region.

Of course, when mixing we need not to take only finite convex combinations; we could use, in fact, any arbitrary probability measures on the space of all non-manipulable mechanisms. We shall not give precise definitions and only present a simple example.

(3.3.9) Example. Again, there is one agent, $V = \mathbf{R}^2$ is the euclidean plane. Given an angle φ, $0 \leq \varphi \leq 2\pi$, we denote by π_φ a one-dimensional mechanism, whose efficiency region is the segment I_φ, whose ends are fixed at the origin of coordinates 0 and at the point $a_\varphi = (\cos\varphi, \sin\varphi)$ (see Fig.6).

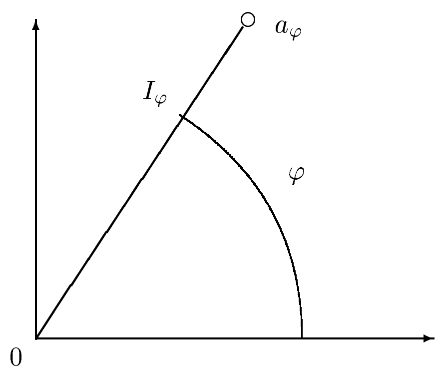

Fig. 6. One-dimensional mechanism with effectivity region I_φ

The performing of this mechanism is clear enough. Assume, for example, that the utility u is given by the second coordinate on $\mathbf{R}^2$. Then

$$\pi_\varphi(u) = \begin{cases} a_\varphi, & \text{if} \quad 0 < \varphi < \pi \\ 0, & \text{if} \quad \pi < \varphi < 2\pi. \end{cases}$$

Let us construct the mechanism θ by integrating the mechanisms θ_φ:

$$\theta = \frac{1}{2\pi}\int_0^{2\pi} \theta_\varphi d\varphi.$$

The value of the mechanism θ for the same utility u is

$$\theta_\varphi(u) = \frac{1}{2\pi}\int_0^{2\pi} \theta_\varphi(u)d\varphi = (\frac{1}{2\pi}\int_0^{\pi} \cos\varphi\, d\varphi, \frac{1}{2\pi}\int_0^{\pi} \sin\varphi\, d\varphi) = (0, 1/\pi).$$

The efficiency region of the mechanism θ is the disk Z with radius $1/\pi$, this is a consequence of the obvious isotropy property of this mechanism (Fig.7). The value $\theta(0)$ is not determined, though $\theta(0) \in Z$.

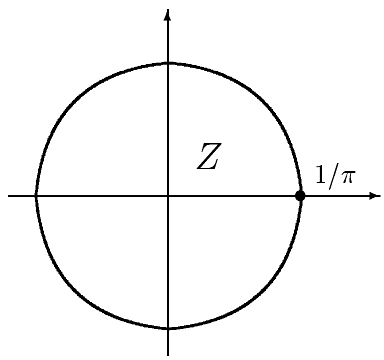

Fig. 7. Unilateral mechanism with effectivity region - the circle Z.

One can construct, in an analogous fashion, a number of other non-manipulable mechanisms, taking different convex combinations (probability mixtures) of unilateral and/or one-dimensional non-manipulable mechanisms.

In this Section, we make the central conjecture that all non-manipulable mechanisms can be obtained in this manner.

(3.3.10) Conjecture. *Any non-manipulable mechanism (in an affine environment) is a probability mixture of unilateral and one-dimensional non-manipulable mechanisms.*

This conjecture can be understood as follows: the set M of all non-manipulable mechanisms is a convex subset in the space F of all mappings $Aff(V)^N \to V$. The convexity of M is clear from what was explained above, we argued that a mixture of non-manipulable rules yields a non-manipulable rule. The convexity of M can also be seen from the following representation of M as an intersection of half-spaces . We fix $n+1$ elements $u_j, j \neq i, u_i, u'_i$ from $Aff(V)$ and determine a "half-space"

$$\Pi(u_N, u_i') = \{\pi : Aff(V)^N \to V,\ u_i(\pi(u_i, u_{N\setminus i})) \geq u_i(\pi(u_i', u_{N\setminus i}))\}.$$

This is a closed set in the topology of pointwise convergence in the space F. The set M is obviously the intersection of all the "half-spaces" $\Pi(u_N, u_i')$. Consider the following two subsets M_u and M_o of M. M_u is the subset of unilateral non-manipulable mechanisms, and M_o is the subset of one-dimensional non-manipulable mechanisms. Now, our conjecture can be rephrased as follows

$$M = \overline{conv}(M_u \cup M_o),$$

where $\overline{conv}$ designates the closed convex hull in a topological convex space F. Proving this conjecture would certainly help in providing a more precise formulation, however we were not successful in proving it. Nevertheless we should like to discuss some insights here.

A major difficulty in proving this conjecture is related to issues about integrals of one-dimensional mechanisms. A one-dimensional mechanism is an analogue of a linear segment (see Examples (3.3.9) and (3.3.5)). Thus it is interesting to examine the issue of integrals of linear segments. These integrals are convex bodies called *zonoids* (see Bolker (1969)). Unfortunately the criterion qualifying a zonoid is rather cumbersome to use. There is however one important case in which things become simpler: this is when we deal with convex polytopes. Namely, a convex polytope is a zonotope if and only if all its facets are centrally symmetric. With this we might hope that the case of mechanisms with finite number of strategies might be easier to study. And, in fact, this is so. Specifically, Gibbard (1978) showed that if $\pi : \prod_{i \in N} S_i \to \Delta(A)$ is a SP-mechanism with a finite set of strategies S_i ,then π is a convex combination of finite set of unilateral and one-dimensional SP-mechanisms.

From this result, Gibbard derives quite an interesting corollary. He shows that if a non-manipulable mechanism $\pi : \prod_{i \in N} S_i \to \Delta(A)$ is efficient, then it is a probability mixture of dictatorial rules, thus has the following form

$$\pi(u_1, ..., u_n) = \Sigma_i \lambda_i \cdot \max(u_i \mid A),$$

where $\lambda_i \geq 0$ and $\Sigma_i \lambda_i = 1$. Hylland (1980) generalizes this to the case of an arbitrary strategies set. This assertion about efficiency would easily follow from the conjecture given above.

The above conjecture was partly confirmed by Barbera et al. (1998). They proved that any twice continuously differentiable non-manipulable mechanism is a convex combination of unilateral non-manipulable mechanisms. Note also that Barbera (1977, 1978) describes all anonymous and neutral SP-mechanisms in the lottery environment.

3.4 The Transferable Environment. Groves Mechanisms

(3.4.1) Another interesting possibility to attain compromise consists in making compensatory transfers, labeled in units of a desirable good (which for brevity, we call *money*). Assume that apart from the "basic" alternatives forming the set A (set within which the choice is made), there is a divisible and desirable good - money. Then, one might think of some compensation schemes (by judicious reallocations of money), if ever a switching from one alternative to another should take place.

The presence of money slightly changes the issue of choice. First instead of the initial set of alternatives A, we deal with the infinite set $A \times \mathbf{R}^N$. Elements of this set are bundles of the type $(a, t_1, ..., t_n)$, where $a \in A$ is the social outcome and t_i is the monetary tranfer received (or given away) by agent $i \in N$. One might consider only those transfer bundles $(t_1, ..., t_n)$ for which $\Sigma t_i = 0$ (or at least $\Sigma t_i \leq 0$). This is indeed reasonable a requirement. However for the time being we shall not take this constraint into account, coming back to it in the next Section. Therefore, one can imagine that any deficit in money (should it occur) is covered at the expense of the mechanism designer.

Second, we assume that the utility of any bundle $(a, t_1, ..., t_n)$ for agent i depends only on a and t_i. In other words, an agent only cares about the social outcome and the monetary amount he gets in the deal. Many situations do not fit into this framework, but this does not make this set-up less interesting.

Lastly, we assume that the preferences of agents, on the set $A \times \mathbf{R}$ are given by utility functions of the form:

$$u(a, t) = v(a) + t,$$

where $v(a)$ represents the money valuation of the utility of alternative a. It might seem that we step into the realms of interpersonal utility comparisons here, but that isn't the case. It means only that individual indifference curves on the space $A \times \mathbf{R}$ result one from another, by a parallel shift along the axis $\mathbf{R}$. Essentially, the last assumption means that the individuals' valuations of alternatives are independent of individual wealth.

(3.4.2) This environment is called *transferable* or more precisely *quasi-linear*. Preferences of agents are given here by valuation functions $v : A \to \mathbf{R}$ and run the set $V = \mathbf{R}^A$. However, sometimes, we shall consider more restricted environments of the form $V_1 \times ... \times V_n$, where the domains V_i are in V. We now give a few examples of such domains:

a) A is a topological space; the valuations v are continuous functions satisfying a condition of upperlevel compactness: $v^{-1}([\alpha, \infty))$ is a compact in A for any $\alpha \in \mathbf{R}$. The last condition is satisfied automatically when A is compact;

b) A is a convex and compact set, functions v are concave;

c) $A = \mathbf{R}^m$, functions v are quadratic concave.

(3.4.3) Typical Issues. Consider the following simple typical choice issue in a transferable environment: take an object (a picture, a vase or anything similar) and a group of participants, each of them ready to acquire it. How should one organize the allocation procedure (or auction)? We start with a simple instance, in which we assume that the reservation price (in dollars, for example) v of the object is known to everybody. Each agent i evaluates the object, and this valuation u_i is known to him only.

Let us begin with a "naive" auction. Participants send their bids (reflecting their true valuation or not) to a collection center. Those agents whose bids are such that $u_i < v$, exit the auction. The object is given to some participant i whose bid is such that $u_i \geq v$. This participant pays v to the seller. Now the question is which participant wins the object? It seems that one should give it to the agent, whose bid u_i is highest, otherwise the auction's outcome is not Pareto optimal. But if it were so, then it would pay each agent to raise his bid, increasing with respect to others' bids. But in fact this would bring us nowhere. All agents will indulge in lying, and such a game does not have any Nash equilibrium.

Now let us examine another organisational principle. Assume the agents' bids are ranked in decreasing order: $u_1 \geq u_2 \geq ... \geq v$ and assume that again agent 1, whose bid u_1 is highest, wins the object. But now assume that he pays $u_1$to the seller. This seems to be more reasonable a principle. However, if agent 1 knows the bid u_2 of the (next) agent 2 and if $u_1 > u_2$, then he would gain by not bidding u_1, but $u_2 + \varepsilon$,where ε is small, and positive. Here as well, agent 1 has an incentive to lie.

It turns out that one can organize auctions such that all agents find it profitable to bid " their true valuation u_i" . It suffices to attribute the object to the agent whose bid is highest (denoted agent 1), but have him pay u_2 (assuming that $u_2 \geq v$) or v (if $u_2 < v$). In this case, the payment of agent 1 does not depend on the valuation u_1 and it is never profitable for him to lie. The same argument holds for the other agents. The payment u_2 can be viewed as a loss that agent 1 bears upon other agents. More exactly, agent 2 bears the totality of the loss, for he would win the object if u_1 had been $< u_2$.

The reader will note here that this auction is only individually non-manipulable (and differs in this from the mechanisms studied in Section 2). In fact, agents 1 and 2 (acting together) are able to push the payment as low as u_3 dividing the profit $u_2 - u_3$.

Next, we complicate the matter slightly. Assume now that the value of the object (or the seller's valuation for the object) v is known only to the seller. For simplicity, we assume that there is only one buyer with valuation u. How can one organize a non-manipulable auction? In this set-up, an exchange occurs only if $u \geq v$ and the agreed-upon price $p \in [v, u]$ (for the seller has to get no less than v and the buyer should not pay more than u). Should the object be sold at the price $p = v$, then we understand that the seller might be tempted, in certain cases and namely when $u > v$, to pretend that his valuation v is higher than it is. Should the object be sold at the price $p = u$, then the buyer might be interested in pretending that his valuation u is lower than it actually is. Now if price is equal to $(u + v)/2$, both will indulge in lying. In general it is not possible to conceive a non-manipulable exchange mechanism unless we are ready to accept that the monetary transfer be unbalanced (see Theorem (3.5.11)). What if we accept unbalanced transfers? Then we can think of the following mechanism. Assume the seller asks v, the buyer bids u. When $u \leq v$, nothing happens. When $u > v$ the buyer gets the object and pays v, while the seller receives u. Of course, this trade is possible only if we bring some additional amount $u - v$ of money in the deal.(This is called financial or budget unbalancedness.) Where does this money come from? We do not answer to this yet, but we present later some (albeit not fully satisfying) explanations.

(3.4.4) We now proceed to a more detailed study of social choice mechanisms in quasi-linear environments. Recall that the utilities of agents are generated by both valuations functions $v_i : A \to \mathbf{R}$ and monetary transfers. Thus the choice set is $A \times \mathbf{R}^N$, the set of bundles of the form $(a; t_1, ..., t_n)$. According to the revelation principle, we can restrict ourselves to the studying of the mechanisms of the following form:

$$\pi : \prod_{i \in N} V_i \to A \times \mathbf{R}^N,$$

$$\pi(v_1, ..., v_n) = (a(v_N); t_1(v_N), ..., t_n(v_N)),$$

where V_i are domains in the space $V = \mathbf{R}^A = \{v : A \to \mathbf{R}\}$ of all valuations. Indeed, for simplicity, we assume both that the set A is finite and that every $V_i = V$. A mechanism is non-manipulable if for any $v_N \in \prod_{i \in N} V_i$, for any agent $i \in N$, and for any valuation $v_i' \in V_i$, the following inequality is satisfied:

$$v_i(a(v_N)) + t_i(v_N) \geq v_i(a(v_i', v_{N \setminus i})) + t_i(v_i', v_{N \setminus i}).$$

Once this definition is posed, we have to answer a crucial question: can we find non-manipulable mechanisms of this kind? If we can, then how many of them are there? Surprisingly, we can answer both these questions and even more we can fully characterize all non-manipulable mechanisms.

More precisely, we shall consider only those mechanisms, which satisfy an additional weak efficiency condition. A mechanism $\pi : V_N \to A \times \mathbf{R}^N$ is *weakly efficient* if the chosen main alternative $a(v_N)$ maximizes the sum of utilities $v_1 + ... + v_n$, i.e. if for any $v_1, ..., v_n$ we have

$$a(v_N) \in \operatorname{Arg\,max}(\sum_{i \in N} v_i).$$

When the outcome is not weakly efficient, the agents could chose another alternative and compensate each other in such a way that everybody would be better off. Note that weak efficiency implies at the same time anonymity and neutrality with respect to the choice of an alternatives from A. One could consider more generally the maximization of a function $v_0 + \sum_{i \in N} v_i$, where v_0 is a fixed function ("utility of a phantom").

A mechanism π is *efficient* if it is weakly efficient and balanced in the following sense: for any v_N,

$$\sum_{i \in N} t_i(v_N) = 0 \text{ (or shorter } : \Sigma t_i \equiv 0).$$

Efficiency is a very desirable property for a mechanism and we discuss it further in the next Section.

(3.4.5) Groves Mechanisms. We now construct an example of a weakly efficient non-manipulable mechanism. For that, we fix a selector $a^* : V_N \to A$ of a correspondence $v_N \mapsto \text{Argmax}(\sum_{i \in N} v_i)$. Then we form a mechanism $\pi^* : V_N \to A \times \mathbf{R}^N$ using the following formula

$$\pi^*(v_N) = (a^*(v_N); t_1^*(v_N), ..., t_n^*(v_N)),$$

where

$$t_i^*(v_N) = \sum_{i \neq j} v_j(a^*(v_N)).$$

In this mechanism, the monetary payoff of agent i is equal to what all other agents "gain" from accepting the project a^*. It is almost obvious from here, that this mechanism π^* is non-manipulable. Indeed, agent i derives utility from the alternative $a^*(v_N)$ and from money, $t_i^*(v_N)$, as follows,

$$u_i(v_N) = v_i(a^*(v_N)) + t_i^*(v_N) = (\sum_{i \in N} v_i)(a^*(v_N)).$$

By definition of a^*, the sum of valuations $v = \sum_{i \in N} v_i$ is maximal at the point $a^* = a^*(v_N)$. So for any other signal v_i', his utility is

$$u_i(v_i', v_{N \setminus i}) = v(a^*(v_i', v_{N \setminus i})) \leq v(a^*(v_N)) = u_i(v_N).$$

This proves the non-manipulability of this mechanism.

One can explain the mechanism π^*'s underlying scheme as follows: the payoffs of agents are chosen in such a way as to equalize gains of all agents. The aims of all participants coincide and the difference between individual and group interests disappears;the group becomes a team in the sense defined by Hurvicz (1972). And in fact, in the process of investigating the issue of stimulation in teams, Groves ended up with precisely such a mechanism.

Let us, however, stress that we have to 'pay a price' in order to get this kind of unanimity and unity within the group. This 'price' is the sum of the carefully designed (by the organizer of the mechanism) set of financial compensations. So, an amount covering these expenses has to be provided externally. Indeed, the total of all monetary compensations of participants is equal to:

$$\sum_{i\in N} t_i^*(v_N) = \sum_{i\in N}(\sum_{i\neq j} v_j(a^*(v_N))) =$$

$$(n-1)(\sum_{i\in N} v_i)(a^*(v_N)) = (n-1)\max_{a\in A} \nu(a),$$

(where again $v = \sum_{i\in N} v_i$). Thus, in order to compensate every agent with the transfer $\max_A v$ one should pump into the system, an amount of money equal to $(n-1)\max_A v$. The fact that a mechanism depends on such a condition is not satisfactory.

(3.4.6) Note, the functions v_i are determined *up to a constant.* Adding a constant to the function $v_j, j \neq i$, changes the value of transfer $t_i^*(v_N)$ independently of agent i's messages. More generally, take n functions $h_i(v_{N\setminus i}), i \in N$, (dependent of v_j $(j \neq i)$ but *independent of* v_i), and then, form a new mechanism π as follows

$$\pi(v_N) = (a^*(v_N); t_1(v_N), ..., t_n(v_N)),$$

where

$$t_i(v_N) = t_i^*(v_N) + h_i(v_{N\setminus i}) = \sum_{j\neq i} v_j(a^*(v_N)) + h_i(v_{N\setminus i}),$$

This new mechanism differs from the former (π^*) only through the modifications in monetary transfers brought by the functions h_i. This kind of mechanism is called a *Groves mechanism.* Since h_i do not depend on v_i (i.e. are constant for agent i), then using the argument above, one easily proves the following important theorem.

(3.4.7) Theorem. *Groves mechanisms are non-manipulable.* ■

We prove the converse assertion in the following section. Non-manipulable mechanisms with different additional properties are constructed by selecting sets of appropriate functions h_i.

(3.4.8) Clark mechanism. To conclude, we discuss an interesting monetary transfer correction procedure (Clark (1971)). Assume agent i is indifferent to all alternatives from A, so that his valuation function v_i is equal to 0 (or a constant). Then it seems natural that he should not be levied with any monetary transfer. One can then propose to normalize functions t_i as follows: $t_i(0, v_{N\setminus i}) \equiv 0$. Therefore functions $h_i(v_{N\setminus i})$ will be given by the formula

$$h_i(v_{N\setminus i}) = -\sum_{i \neq j} v_j(a^*(0, v_{N\setminus i})).$$

Note that Clark mechanism remains unchanged when adding constants to the valuations v_i.

Let us further develop Clark's argument. Suppose agent i is indifferent to all alternatives from A, then the desired outcome will be $a_i^* = a^*(0, v_{N\setminus i})$. If he sends the message w_i, then the outcome is a point $a^* = a^*(w_i, v_{N\setminus i})$. At this point, the other agents $j \neq i$ experience a loss equal to $v_j(a^*) - v_j(a_i^*)$. The "loss-causing" agent i is imposed a penalty t_i equal to $\sum_{i\neq j}(v_j(a^*) - v_j(a_i^*))$, where the penalty is non-positive, as a_i^* maximizes $\sum_{i\neq j} v_j$. Nevertheless, it is easy to check that this monetary penalty is inferior to what he gains from going $a_i^* \to a^*$; this non-monetary gain is equal to $v_i(a^*) - v_i(a_i^*)$, where v_i is his true valuation. Therefore agent i has no incentive to distort his valuation function v_i; the function of a "message" for him consists in moving the social outcome from the point a_i^* to the point a^*, which gives maximum to his full utility equal to $v_i(x) + \sum_{i\neq j}(v_j(x) - v_j(a_i^*)) = v(x)+$ constant.

It is worthwhile, once again, to note that the penalty, imposed on agent i for the loss caused to other agents, does not serve to compensate (even partially) the loss of other agents. Indeed the "wronged" agents may, in fact, really never receive this money. The penalty's role here is to force agent i to think twice by weighing his gains against his money losses.

We have already shown that in a Clark mechanism all $t_i \leq 0$. Thus, and this is a peculiarity of this mechanism, it does not require that the organizer subsidize the mechanism. When agent i does not influence the social outcome by sending his message v_i (i.e. $a^* = a_i^*$), then his transfer is equal to 0. In this set-up, only those agents who actively influence the selection of the outcome will be taxed. These agents are called *pivotal,* or leading. This is why, in the literature, the Clark mechanism is often called the *pivotal* mechanism.

3.5 Further Properties of Groves Mechanisms

(3.5.1) Characterization. In the previous Section, we discovered a whole class of non-manipulable mechanisms set-up in a quasi-linear environment. Do there exist other non-manipulable mechanisms with nice properties? We show that, under sufficiently weak assumptions on the environment, all non-manipulable mechanisms are Groves mechanisms.

Let us start by assuming that the set of projects A is finite and $V_i = V = \mathbf{R}^A$, that is any valuations are feasible. Let $\pi : V^N \to A \times \mathbf{R}^N$ be a non-manipulable weakly efficient mechanism such that $a^*(v_N) \in$Argmax$(\sum_{i\in N} v_i)$ for any profile v_N.

(3.5.2) Theorem. *Under these assumptions, π is a Groves mechanism.*

Proof. We have to show that for any agent i the function

$$t_i(v_i, v_{N\setminus i}) - \sum_{i\neq j} v_j(a^*(v_i, v_{N\setminus i}))$$

is independent of v_i; then we denote it $h_i(v_{N\setminus i})$. Since in the sequel the valuations v_j, $j \neq i$, remain fixed, for brevity we shall not write further $v_{N\setminus i}$ and we shall denote the function $\sum_{i\neq j} v_j$ by w. So we should check now that the function

$$h(v_i) = t_i(v_i) - w(a^*(v_i))$$

is constant. First, note that non-manipulability implies that the monetary transfer t_i remains unchanged if the selected project a^* does not change.

Assume now that $h(v_i) = h(v_i^{'}) + \varepsilon$, where $\varepsilon > 0$. We form an auxiliary valuation $\widetilde{v_i}$ setting

$$\widetilde{v_i}\ (x) = \begin{cases} -w(x) + \varepsilon/2, & \text{if} \quad x = a^*(v_i^{'}), \\ -w(x), & \text{if} \quad x \neq a^*(v_i^{'}). \end{cases}$$

Since the function $\widetilde{v_i}\ + w$ attains its maximum at the point $a^*(v_i^{'})$, then $a^*(\widetilde{v_i}) = a^*(v_i^{'})$. Thus we have $t_i(\widetilde{v_i}) = t_i(v_i^{'})$ due to the preceding remark. Hence,

$$\widetilde{v_i}\ (a^*(v_i) + t_i(v_i) = -w(a^*(v_i)) + (w(a^*(v_i)) + h(v_i)) =$$

$$h(v_i^{'}) + \varepsilon = t_i(v_i^{'}) - w(a^*(v_i^{'})) + \varepsilon = t_i(\widetilde{v_i}) + \widetilde{v_i}\ (a^*(\widetilde{v_i})) + \varepsilon/2.$$

Thus when the true valuation is $\widetilde{v_i}$, agent i gains by sending the message v_i, which contradicts non-manipulability of π. ∎

When A is an arbitrary compact and V_i consists of all continuous functions on A (as well as for many other cases), a slight modification of this argument does the job. Holmström (1979) obtains a more general result:

(3.5.3) Theorem. *If all domains V_i are convex (as subsets in the space of functions on A), then every weakly efficient non-manipulable mechanism is a Groves mechanism.*

Note that the requirement of convexity of V_i is essential, it can not be replaced by a requirement of connectedness. We can sense it in a one agent example with $A = [0, 1]$. Assume the domain of valuations consists of functions of the form $v_\alpha(x) = -\mid x - \alpha \mid, \alpha \in [0, 1]$. Clearly, the mechanism where $t \equiv 0$ is non-manipulable. However, the mechanism, whose associated transfer is $t(\alpha) = \alpha/2$ (where in place of v_α the agent calls simply α), is non-manipulable as well, but is not a Groves mechanism.

These results provide an additional reason to consider only Groves mechanisms.

(3.5.4) Let us make one further comment. In the previous theorems we assumed mechanisms to be weakly efficient. Replacing this assumption with the weaker attainability condition (any alternative from A is attained for an appropriate profile v_N), then the following result by Roberts (1979) is true: there exist a function $v_0 : A \to \mathbf{R}$ and a non-zero non-negative weight vector $k = (k_1, ..., k_n)$ such that:

a) $a^*(v_N) \in \text{Argmax}(v_0 + \sum_{i \in N} k_i v_i)$,

b) $k_i v_i(a^*(v_N) + t_i(v_N)) = (v_0 + \sum_{i \in N} k_i v_i)(a^*(v_N)) + h_i(v_{N \setminus i})$.

Roughly speaking, the case is almost similar to that in Groves with two slight differences: agents are assigned weights k_i and a valuation of a "phantom" v_0 is added. Roberts notes that Gibbard's theorem (A2.7 of Chapter 1) can be derived from his result, provided transfers be forbidden, i.e. $t_i \equiv 0$.

(3.5.5) Efficiency. Non-manipulability is not an end per se. However it is important since it simplifies the strategic behavior of agents and therefore enables to predict outcomes as well as to evaluate their efficiency. If we know the outcome then we can appreciate how well a given mechanism works and whether it outperforms any other mechanisms. Within the large list of desirable properties for mechanisms, we discuss here the sole efficiency, or financial balancedness, property. (Green and Laffont (1979) discuss some other properties.) Remember (3.4.4) that financial balancedness means

$$\sum_{i \in N} t_i(v_N) = 0$$

for any $v_N \in V_N$. If $\sum_{i \in N} t_i(v_N) > 0$, then the group must be subsidized by this amount; the problem is that subsidies are not always available. If $\sum_{i \in N} t_i(v_N) < 0$, then the group ends up throwing out money, which obviously doesn't please any of its members.

The issue now is: do there exist Groves mechanisms,

$$\pi : V_1 \times ... \times V_n \to A \times \mathbf{R}^N,$$

such that $\sum_{i \in N} t_i(v_N) \equiv 0$? Let us say at once that balanced Groves mechanisms exist only for very special environments. Assume we do not succeed in nullifying the function $\sum_{i \in N} t_i(v_N)$, then we could think of minimizing its deviation from zero. We shall take deviation from zero here in the "uniform" sense. We call *inefficiency measure* of a mechanism π the following number,

$$\varepsilon(\pi) = \sup_{v_N \in V_N} \left| \sum_{i \in N} t_i(v_N) \right| .$$

If $\varepsilon(\pi) = 0$, then the mechanism π is efficient; the smaller $\varepsilon(\pi)$ the more efficient π. We introduce the following notation:

$$W(v_N) = (n-1) \max_{a \in A} \sum_{i \in N} v_i(a),$$

where $v_N = (v_1, ..., v_n)$ is a valuation profile. The function $W : V^N \to \mathbf{R}$ depends only on the environment $(A; V_1, ..., V_n)$. The following relation emphasizes its role:

$$\sum_{i \in N} t_i(v_N) = W(v_N) + \sum_{i \in N} h_i(v_{N \setminus i}),$$

or in short, $\sum_{i \in N} t_i = W + \sum_{i \in N} h_i$. From here we derive an

(3.5.6) Efficiency Criterion. *A Groves mechanism $\pi : V_N \to A \times \mathbf{R}^N$ with correction functions $h_i(v_{N \setminus i})$ is efficient if and only if*

$$W(v_N) + \sum_{i \in N} h_i(v_{N \setminus i}) \equiv 0.$$

(3.5.7) Thus one sees readily that the issue of finding a mechanism with a small $\varepsilon(\pi)$ is closely related to the issue of approximating the function W of n variables $v_1, ..., v_n$ by a sum of functions h_i, each of which depending on $(n-1)$ variables. The latter problem being connected with a finite differences issue. We introduce the following notions.

Take a function $f : V_N \to \mathbf{R}$ and two points $v_N, v_N^{'} \in V$. We pose

$$D(f; v_N, v_N^{'}) = \sum_{K \subset N} (-1)^{|K|} f(v_K, v_{\bar{K}}^{'})$$

to be the $n-$th mixed difference. For $n = 1$, this is equal to $f(v^{'}) - f(v)$; for $n = 2$, this is

$$f(v_1^{'}, v_2^{'}) - f(v_1^{'}, v_2) - f(v_1, v_2^{'}) + f(v_1, v_2).$$

Now, on top, we pose

$$\delta = \sup_{v_N, v_N' \in V_N} \mid D(W; v_N, v_N') \mid .$$

The number δ depends only on the function W, i.e., on the environment, and can be understood as a measure of the complexity of a given environment $(A; V_1, ..., V_n)$.

It is easy to note that for any function $h_i(v_{N \setminus i})$ which does not depend on the variable v_i, the mixed difference $D(h_i; \cdot, \cdot)$ is identically equal to zero. Thus the term Σh_i does not influence the mixed difference and we get the equality:

$$D(\sum_{i \in N} t_i; \cdot, \cdot) = D(W, \cdot, \cdot),$$

and prove the following result

(3.5.8) Theorem. *The inequality $\varepsilon(\pi) \geq \delta/2^n$ holds for every Groves mechanism π.*

Indeed,

$$\left|D(W;v_N,v_N^{'})\right| \leq \left|D(\sum t_i;v_N,v_N^{'})\right| \leq$$
$$\leq \Sigma_{K\subset N}\left|\sum(-1)^{|K|}t_i(v_K,v_{\overline{K}}^{'})\right| \leq 2^n\varepsilon(\pi).$$

Therefore $\delta \leq 2^n\varepsilon(\pi)$. ■

(3.5.9) Corollary. *If there exists an efficient Groves mechanism then* $\delta = 0$.

Thus, to warrant existence of at least one efficient Groves mechanism, the environment $(A;V_1,...,V_n)$ has to satisfy quite a stringent requirement. Namely, for any $v_N,v_N^{'} \in V_N$ the following equality (*Walker condition*) has to be fulfilled:

$$D(W;v_N,v_N^{'}) = 0.$$

(Below, we shall show that Walker condition suffices to warrant the existence of an efficient Groves mechanism). We should note that Walker's condition obtains very rarely.

(3.5.10) Example. Consider a simple situation, where A consists only in two alternatives: a "status-quo" **0** and a proposed project **1**. We normalize valuations v as follows $v(\mathbf{0})=0$, and identify v with the real number $v(\mathbf{1})$. The space of all valuations is the real line $\mathbf{R}$. The function $W:\mathbf{R}^N \to \mathbf{R}$ takes the following explicit form:

$$W(x_1,...,x_n) = (n-1)\max\{0,x_1+\ldots+x_n\}.$$

We now calculate the mixed difference $D(W)$ at points $x_N = (-n,...,-n)$ and $x'_N = (1,...,1)$. $W(x_K,x'_{\bar{K}}) = 0$ for all "intermediate" points $(x_K,x'_{\bar{K}})$ with $K \neq \emptyset$. And $W(x'_N) = n(n-1)$ is true only at one point, $(x_\emptyset,x'_N) = x'_N$. Finally

$$D(W;x_N,x_N^{'}) = n(n-1) \neq 0.$$

Thus even in this simple environment, Groves mechanisms are not efficient. In Appendix 3.A1, we come back to evaluating $\varepsilon(\pi)$ in this simple environment.

The same arguments prove the following general result.

(3.5.11) Theorem. *Let A be compact, let V_i consist of all continuous functions on A. Then no Groves mechanism can be efficient.*

An analogous result was obtained by Walker (1980). More precisely, Walker showed that if A is an open convex set and V_i consists of all strictly concave functions on A with compact upperlevels, then for almost every profile $v_N \in V_N$, the sum $\sum_{i\in N} t_i(v_N)$ is not equal to zero.

Since efficient mechanisms are extremely rare, we are forced back to the issue of finding Groves mechanisms with small $\varepsilon(\pi)$. We now mention a general, but a rather weak proposition.

(3.5.12) Theorem. *There exists a Groves mechanism* $\widetilde{\pi}$ *with* $\varepsilon(\widetilde{\pi}) \leq \delta$.

Proof. Fix a profile $v_N^0 \in V_N$. Then for any $v_N \in V_N$, the relation

$$W(v_N) + \sum_{K \neq \emptyset} (-1)^{|K|} W(v_K^0, v_{\overline{K}}) = D(W; v_N^0, v_N),$$

obtains. When $K \neq \emptyset$, the functions $W(v_K^0, v_{\overline{K}})$ are independent of the variables v_i, $i \in K$. Thus one can think of grouping them to form the functions $h_i(v_{N \setminus i})$. If we now take a Groves mechanism $\widetilde{\pi}$ with those correction functions h_i, we get

$$\sum_{i \in N} \widetilde{t}_i\ (v_N) = D(W; v_N^0, v_N),$$

whence $\varepsilon(\widetilde{\pi}) \leq \delta$. ■

The proof of Theorem (3.5.12) is constructive in the sense that it yields a mechanism whose inefficiency is "small" enough. However the valuation given by this theorem is rather coarse (see Appendix 3.A.1). Theorems (3.5.8) and (3.5.12) provide a useful and simple existence criterion for efficient Groves mechanisms:

(3.5.13) Criterion. *Given the environment* $(A; V_1, ..., V_n)$, *there exists an efficient Groves mechanism if and only if* $\delta = 0$.

We now give two examples in which we apply this criterion.

(3.5.14) Example. Assume that the preferences of one among the agents (say, the first) are fixed and known, such that V_1 consists of a unique element v_1. Then, the function $W(v_N)$ is tautologically constant on v_1 and $D(W; \cdot, \cdot) \equiv 0$, so that $\delta = 0$ and hence it is possible to have an efficient Groves mechanism.

This example is not very interesting, for any disbalance is written off at the expense of the 1st agent. Agent 1 has no discretionary power, therefore one can impose on him any kind of tax without influencing his incentive to reveal his true preferences. Of course this follows from our focusing on non-manipulability and our disregarding important issues such as individual rationality, fairness, and so on.

(3.5.15) Example. This example is more interesting. Here $A = \mathbf{R}$ is the real line, the space V_i consists of "quadratic" valuations functions of the form $\theta x - x^2/2$, where $\theta \in \mathbf{R}$ is a "parameter" (see Fig.8). It is easy to understand what θ represents: the function $\theta x - x^2/2$ attains its maximum precisely at $x = \theta$. Therefore we will assume that agents simply call their value θ_i.

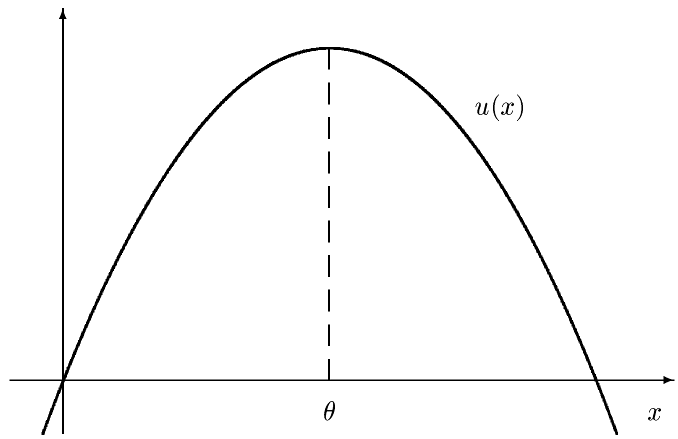

Fig. 8. Quadratic function with parameter θ.

Let us now express the function $W(\theta_1, ..., \theta_n)$. It is easy to see that the following function

$$\sum_{i\in N} v_i = \sum_{i\in N}(\theta_i x - x^2/2)(\sum_{i\in N}\theta_i)x - nx^2/2$$

is maximal at the point $(\sum_i \theta_i)/n$, and is equal then to $(\Sigma\theta_i)^2/2n$. Thus $W(\theta_1, ..., \theta_n) = (n-1)(\Sigma\theta_i)^2/2n$.

Let us show now that $\delta = 0$ for $n \geq 3$. Let us recall that if $W(v_N)$ is an arbitrary differentiable function, then

$$D(W; v_N, v'_N) = \int \frac{\partial^n W}{\partial v_1 \cdots \partial v_n} dv_1 \cdots dv_n$$

where integration is taken on a parallelepiped with opposite vertices at the points v_N and v'_N. Thus the condition $\delta = 0$ is equivalent to

$$\frac{\partial^n W}{\partial v_1 \ldots \partial v_n} \equiv 0.$$

(Generally δ has an upper bound:

$$\delta \leq \max_{v_N \in V_N} \left| \frac{\partial^n W}{\partial v_1 \ldots \partial v_n} \right| \prod_{i \in N} (\text{“diameter” of } V_i).$$

Here the first factor expresses the “local complexity of the environment”, while the second relates to the “general width of the environment”.)

Let us come back to the quadratic case. Assume that $n \geq 3$; then

$$\frac{\partial^n}{\partial\theta_1 \dots \partial\theta_n}(\theta_1 + \cdots + \theta_n)^2 = 0.$$

Hence $\delta = 0$, and by the former criterion, there exist efficient Groves mechanisms. Moreover one can select one that is anonymous, for which all functions h_i are identical. To do so one needs only to decompose explicitly $(\theta_1 + \cdots + \theta_n)^2(n-1)/n$ into a sum of n "identical" terms, each depending on θ_{N-i}. Finally, skipping a lot of cumbersome computations, the monetary transfers take the form:

$$t_i(\theta_N) = -\frac{n-1}{2n^2}(\theta_i - \mu_{-i})^2 + \frac{1}{2n(n-2)}\Sigma_{j\neq i}(\theta_j - \mu_{-i})^2,$$

where $\mu_{-i} = \frac{1}{(n-1)}\Sigma_{j\neq i}\theta_j$. In words, agent i:

a) is penalized proportionally to the square of deviation between θ_i and the average value of θ_j (of the remaining agents);

b) is rewarded proportionally to the mean square deviation the remaining agents from μ_{-i}.

Groves and Ledyard (1977) use the quadratic environment to construct a "nice-looking" tax mechanism in an economy with public goods.

(3.5.16) In Example (3.5.15), we assumed that $n \geq 3$. When $n = 2$, no efficient mechanism exists, for $\delta \neq 0$ (note that the case $n = 2$ was also peculiar in Section 2.6). This not by chance, for when $n = 2$ efficient mechanisms are very rare, which does not mean that there are no such mechanisms at all. Assume that $A = A_1 \times A_2$ and assume that the valuation of agent i depends only on the projecting of alternatives on A_i. Then it is easy to construct an efficient Groves mechanism, moreover a mechanism for which $t_i \equiv 0$. However, the construction of such an efficient mechanism rests on the fact that the environment is degenerate. More exactly, assume that the set of valuations V_i $(i = 1, 2)$ includes indifference $(v_i \equiv 0)$. Then, if $\delta = 0$, $v_1 \in V_1$ and $v_2 \in V_2$ have the same maximum. Indeed, by Theorem (3.5.2), we have $D(W; (0,0), (v_1, v_2)) = 0$, or $\max(v_1 + v_2) = \max v_1 + \max v_2$.

Thus when $n = 2$, efficiency arises because of coinciding agents' aims, when there is no conflict of interests.

(3.5.17) We conclude with two brief remarks. First, there exist non-manipulable mechanisms and such mechanisms are Groves mechanisms. Second, they all are coalitionally manipulable (in-efficiency just means that the whole coalition is able to manipulate; as evidenced in the case $n = 2$ studied above, any group of size 2 is also able to manipulate). Therefore when discussing coalitional aspects of mechanisms, we should select a weaker equilibrium concept than that of dominant strategies, and choose, for example, the strong equilibrium concept, as in Chapter 5.

3.A1 The Simple Transferable Environment Case

We come back to the simple transferable environment case from Example (3.5.10), in which $A = \{\mathbf{0,1}\}$. We shall assume that V_i is the segment $[-1, 1]$, i.e. that $0 \leq v_i(\mathbf{1}) \leq 1$. We showed that

$$W_n(v_N) = (n-1)\max\{0, v_1 + \cdots + v_n\}.$$

Denote the sum $v_1 + \cdots + v_n$ by v, so

$$W_n(v_N) = (n-1)\max\{0, v\}, v \in [-n, n].$$

Our goal is to estimate the efficiency of Groves mechanisms in such an environment, as a function of the number of participants n. One can show that $\delta = \delta_n$ depends on n as showed in Table 1.

Table 1.

n	2	3	4	5	6	...	n
δ_n	2	3	12	20	60	...	$\approx 2^n\sqrt{n/2\pi}$

Due to Theorem (3.5.12), we know that there exists a mechanism π_n whose $\varepsilon(\pi) \approx 2^n\sqrt{n/2\pi}$. We show that this estimation is rather coarse and that one can refine it improving on the efficiency of the mechanism. To this end, we construct two arrays of more efficient mechanisms .

The construction principle of the first array of mechanisms consists in approximating the function $W_n(v)$ on the segment $[-n, n]$ by a degree $n-1$ polynomial $P_{n-1}(v)$. This polynomial is a sum of $n-1$ variables polynomials, in which only $n-1$ variables appear (thus the monomial $v_1 \cdot ... \cdot v_n$ is never encountered).

We shall approximate not W_n but more convenient function $\frac{n-1}{2}|v|$. It differs from W_n by a linear function. The approximation of the function $| v |$ by polynomials is a classical problem (see, for example, Dzyadyk (1977)) which we shall not elaborate. In brief, the function W_n can be approximated by a polynomial of power $n-1$ up to the efficiency factor $0,15n$. Thus there exists a Groves mechanism $\tilde{\pi}_n$ with $\varepsilon(\tilde{\pi}_n) \approx 0.15n$.

The second array of mechanisms is the now familiar Clark mechanism, see Section 3.4, which we denote by π_n^C. Recall that only those "pivotal" agents, who influence the social outcome through their messages, are imposed a transfer. One checks that the number of "pivotal" agents inferior to $n/2$ and that each pays a money amount ≤ 1. Thus in the case of the Clark mechanism, the following inequalities $| \Sigma t_i(v_N) | \leq n/2$ and $\varepsilon(\pi_N^C) \leq n/2$ obtain. The Table 2 summarizes these data.

Table 2.

n	$\delta_n/2^n$, i.e. lower bound for $\varepsilon(\pi_n)$	$\varepsilon(\tilde{\pi}_n)$	$\varepsilon(\pi_n^C)$	the bound from Theorem 5.12
2	.59	.64	1.00	2
3	.38	.96	1.50	3
4	.75	1.27	2.00	12
5	.62	1.59	2.50	20
6	.94	1.91	3.00	60
$n \gg 1$	$\sqrt{n/2\pi}$	$.15n$	$n/2$	$2^n\sqrt{n/2\pi}$

Remarking that in a Clark mechanism $t_i \leq 0$,for each i, one can reduce the efficiency factor to $n/4$. This speaks of the effectiveness of Clark mechanisms. This feature, added to simplicity, speaks in favor of Clark mechanisms. However we do not know whether the above mechanism is the "absolute best"

3.A2 Acceptable Mechanisms in Transferable Environment

Non-manipulable mechanisms can be designed in a quasi-linear environment, but they rarely select efficient outcomes. We can find efficient mechanisms, however, when and if we accept weakening the equilibrium concept, going from dominant strategies to, for example, Nash equilibrium concept as in Chapter 2. More precisely, we shall focus here on mechanisms, whose equilibrium outcome is Pareto optimal for any quasi-linear preference profile. We disregard here the existence of equilibrium issue, therefore taking a different approach from that of Chapter 2. We provide now two acceptable mechanisms. In what follows, A is a finite set (or a compact) and the valuations functions are continuous.

(3.A2.1) The Walker Mechanism. Take $n \geq 3$ agents, arranged around a circle, so that $N = \mathbf{Z}/n\mathbf{Z}$. The strategy spaces S_i are the same as in the direct mechanism case, that is consist of valuations of the form $v : A \to \mathbf{R}$. We obtain an outcome $\pi^*(v_N) = (a^*(v_N); t_1(v_N), ..., t_n(v_N))$ as follows:

$$a^*(v_N) \in \text{Argmax}(v_1 + \cdots + v_n),$$
$$t_i(v_N) = v_{i+1}(a^*(v_N)) - v_{i-1}(a^*(v_N)).$$

Note that $\sum_i t_i(v_N) = 0$ for any v_N, implying that this mechanism is tautologically financially balanced.

We check now that every Nash equilibrium outcome is efficient. Let $v_N^* = (v_i^*)_{i \in N}$ be a Nash equilibrium for the true valuations v_i and assume that the outcome, associated to v_N^*, is not efficient, that is that there exists an alternative $a \in A$ and payoffs $t_1, ...t_n$ such that $\sum_i t_i = 0$ and for any participant $i \in N$

$$(a, t_i) \succ_i (a^*(v_N^*), t_i(v_N^*)).$$

Every agent i can, by choosing an appropriate valuation v_i', force the outcome to be a, where $a = a^*(v_i', v_{N\setminus i}^*))$. Since v_N^* is the array of equilibrium strategies, we have

$$(a^*(v_N^*), t_i(v_N^*) \succeq_i (a, t_i(v_i', v_{N\setminus i}^*)).$$

By transitivity, $(a, t_i) \succ_i (a, t_i(v_i', v_{N\setminus i}^*))$, i.e.

$$t_i > t_i(v_i', v_{N\setminus i}^*) = v_{i+1}(a) - v_{i-1}(a).$$

Summing up all these inequalities yields a contradiction: $0 = \sum_i t_i > 0$.

(3.A2.2) The Hurwicz Mechanism. Here $n \geq 2$. The agents' messages consist in pairs $s = (v, \alpha)$, where $v : A \to \mathbf{R}$ figures the valuation and α is a real number. The mechanism is structured as follows. The choice $a^*(s_N) \in A$ is carried out on the base of the maximization of $v_1 + \cdots + v_N$, so that

$$a^*(s_N) \in \text{Argmax}(\sum_{i \in N} v_i).$$

The monetary transfers $t_i(s_N)$ take the form

$$t_i(s_N) = C(a^*(s_N)) \cdot (\sum_{i \neq j} a_j) - (\Sigma_{j \in N} \alpha_j)^2,$$

where $C : A \to \mathbf{R}$ is an auxiliary function, denoting the cost of a project. (This function is more important for the existence of equilibrium than for optimality issues).

We check that a Nash equilibrium outcome is optimal. Let $s_N^* = ((v_i^*, \alpha_i^*))$ is a Nash equilibrium. Then

I. $\sum_{i \in N} \alpha_i^* = 0$ because the choice of α_i influences only the monetary transfers t_i through the term $(\sum_{i \in N} \alpha_j)^2$, which can always be made equal to zero by any agent.

II. From here, we get that $t_i(s_N^*) = -\alpha_i^* C(\alpha^*(s_N^*))$ and $\sum_i t_i(s_N^*) = 0$.

III. Assume now that the outcome at s_N^* is not optimal, i.e. there is a better outcome $(a; t_1, ..., t_n)$ for which $\sum_{i \in N} t_i = 0$ and for any $i \in N$

$$(a, t_i) \succ_i (a^*(s_N^*), t_i(s_N^*)).$$

Again having chosen an appropriate signal $s_i' = (v_i', \alpha_i^*)$ agent i can force outcome at $(s_i', s_{N\setminus i}^*)$ to be equal to a. Since s_N^* is an array of equilibrium messages we have,

$$(a^*(s_N^*), t_i(s_N^*)) \succeq_i (a, -\alpha_i^* C(a)).$$

So $(a, t_i) \succ_i (a, -\alpha_i^* C(a))$, i.e. $t_i > -\alpha_i^* C(a)$. Summing up these inequalities, we get a contradiction:

$$0 = \sum_i t_i > -C(a) \cdot \sum_i \alpha_i^* = 0.$$

Bibliographic Comments

Dominant strategy mechanisms have been the central focus of the literature on social choice mechanisms for a number of reasons. First, the mere possibility to yield truthful revelations of participants' preferences was felt to be a very attractive feature of mechanisms in general. Investigators were therefore curious to know at what price this could be fulfilled. Second, the focus on dominant strategy mechanisms was connected somehow with an anterior literature centered around Arrow's result. Finally, dominant strategy mechanisms were interesting because one was often able to describe them fully.

The appeal of DS-mechanisms also stems from practical purposes (see Green and Laffont (1979)). Take for example incentive theory, which aims at smoothing differences of interests between a planning (or managing) center and its subordinated agents. Another practical case in which DS-mechanisms were felt to be important was that of planned economies. In these economies, in order to manage ressources efficiently and to establish plans, truthful information had to reach the center. The issue was then how to design incentives such as to reconcile interests of all agents as well as to elicit truthful behaviour from them.

Hurwicz (1972) was the first to study the incentive compatibility problem. He showed that it is possible to influence the equilibrium outcome by either manipulating initial endowments or distorting preferences in a pure exchange setting. Farquharson (1969) investigating voting procedures showed that, to his knowledge, all known voting procedures were manipulable. Gibbard (1973) and Satterthwaite (1975) proved this in a general setting (see Appendix 1.A2). These results were as negative as those of Arrow. By then, however, Clark (1971) and Groves (1973) had uncovered the existence of non-manipulable mechanisms in a restricted (quasi-linear) environment (see Sections 3.4 and 3.5), launching the study of DS-mechanisms.

Gibbard (1973) introduces the revelation principle. The case of the single agent case was considered by Barbera (1983), Burkov and Enaleev (1985).

As we mentioned above, narrowing the environment brings new and interesting DS-mechanisms. And a classic example is the single-peaked environment (Section 3.2). Black in 1948 had already noted that in single-peaked environments, nontrivial rules of aggregating preferences obtain. It so happens that these rules can be transformed in non-manipulable social choice rules. Moulin (1980) provides an exhaustive description of such mechanisms (for linear structures on A) and see also Border and Jordan (1983). Demange (1982) generalized these results to trees. Generalization to a multi-dimensional case is rather difficult (however some results were obtained in the case of euclidean environment, see Danilov (1988)).

Gibbard (1977) launched the investigation of DS-mechanisms in a lottery environment (see Section 3). Gibbard (1978) further explored the issue. See also Barbera (1977, 1978), and Freixas (1984).

Transferable (or quasi-linear) environments elicited great interest as germane to public goods economies and auction settings. Green and Laffont (1979) and Groves (1979) offer surveys on this issue. Moulin (1982) has another very interesting survey departing from both DS-mechanisms and quasi-linear environment set-ups. See also a survey by Afanasiev and Lezina (1982) and an article by Makowski and Ostroy (1992).

Two interesting topics remain on which we shall not elaborate. The first concerns dynamic mechanisms. The procedures exposed in Chapters 4 and 5 will give the reader a flavor of the latter mechanisms. Green and Laffont (1979) develop the issue more thoroughly. The second topic concerns matching and marriage problems (Roth and Sotomayor (1990), Sönmez (1996)) and related to it, the house market issue (Shapley and Scarf (1971)). Some general results concerning strategy-proof mechanisms in such models were obtained by Sönmez (1999).

4. Cores and Stable Blockings

This chapter is devoted to the issue of stable outcomes, that is those outcomes which are rejected by no coalition of agents. The existence of such outcomes depends only on the stability of the blocking generated by a given mechanism. We investigate here stable blocking relations. We begin with a few examples and give some useful instruments (Section 4.1). In Sections 4.2-4.4 we discuss three classes of stable blockings: additive blockings, almost additive blockings, and convex blockings. The main finding is that for almost additive blockings a family of coalitions which reject alternatives out of the core, can be equipped with a laminar structure (Theorem (4.4.7)). Section 4.5 reviews a series of necessary conditions to warrant the stability of a given blocking. In particular, convexity and almost-additivity turn out to be necessary for the stability of maximal blockings. In Section 4.6, we develop a veto-procedure in order to find elements in the core. The procedure yields single-element outcomes for any maximal convex blocking.

We conclude with three Appendices. Appendix 4.A1 introduces one more class of stable blockings namely balanced blockings which bear some resemblance to Scarf's balanced cooperative games. Appendix 4.A2 briefly tackles the issue of blockings with infinite number of alternatives. Appendix 4.A3 provides a proof of the lemma on harems.

4.1 Stable Outcomes

(4.1.1) The two last chapters of this book are devoted to the investigation of solution concepts which enable agents to unite into coalitions and coordinate their actions within a coalition. We assume that a necessary condition for coalition formation is the *strict* improving of the outcome for *all* members of the coalition. A coalition K will reject a given state a only when the result of joint actions of its members happens to be strictly better than a for all $i \in K$. Of course, this result will depend not only on the actions of a "rebelling" coalition K, but also on the reactions of the complementary coalition $\overline{K}$. We will consider in detail only two extreme cases of reactions. One extreme case is where members of the complementary coalition keep their previous strategies. We delay investigation of this case until the next chapter. In the other case, any reaction of the complementary coalition (even if it damages the coalition $\overline{K}$) is admissible. However, the difference between these cases is not as great as it may seem, and thus chapters 4 and 5 are closely related.

We now give more precise definitions. Let $\pi : S_N \to A$ be a mechanism, and R_N be a preference profile. An *equilibrium with threats* in the game $G(\pi, R_N)$ is a strategy profile $s_N^* = (s_i^*)_{i \in N}$ such that, for any coalition K and a coalitional strategy $s_K \in S_K$, the complementary coalition has a strategy (namely, a threat) $s_{\overline{K}} \in S_{\overline{K}}$ such that $\pi(s_N^*) R_i \pi(s_K, s_{\overline{K}})$ for some agent $i \in K$.

Denoting by $a = \pi(s_N^*)$, the preceding relation can be rephrased as

$$\pi(s_K, s_{\overline{K}}) \in L(a, R_K).$$

Or alternatively, recalling the definition of a blocking (1.5.2): any coalition$\overline{K}$ β-forces the set $L(a, R_K)$. This is equivalent to saying that no coalition K can block the set $L(a, R_K)$. One can see immediatly that this definition of an equilibrium outcome with threats depends less on mechanism π than on the blocking B_π generated by it. Hence, in this chapter we will deal with blockings rather than with mechanisms. Let us reformulate in this case the definition of an equilibrium (or stable) outcome. Although, it is more convenient to state the conditions for which an outcome is not stable, that is, when there exists a coalition K which can definitly improve it.

(4.1.2) Definitions. A coalition K *rejects* an alternative (given a blocking B and a preference profile R_N) if K blocks the set $L(a, R_K)$. An alternative which is rejected by no coalition is *stable* (for the game without side payments $G(B, R_N)$). The set of all stable outcomes is called *the core* and is denoted by $C(B, R_N)$.

The set of stable outcomes may be empty. Let us consider the following two examples.

(4.1.3) Example. Take three agents and three alternatives, and assume that the simple majority rule prevails (see Example (2.1.5)). Let R_N be the following cyclic profile,

x	y	z
y	z	x
z	x	y
1	2	3

There is no stable outcome in this case and the core is empty. For example, outcome x is not stable since agents 2 and 3 by concluding an agreement and calling z can force the outcome to be z, whatever the actions of agent 1. Moreover $z \succ x$ for 2 and 3. Thus x is an unstable outcome, which is rejected by coalition $\{2, 3\}$. By the same kind of argument, y is rejected by coalition $\{1, 3\}$ and z - by coalition $\{1, 2\}$.

Identical considerations obtain for the case of the kingmaker mechanism (see Example (2.1.3)) which generates the same blocking.

(4.1.4) Example. There are three agents and four alternatives. Consider the following blocking: any coalition of k agents is able to block at most k alternatives. We consider the profile R_N :

$$R_N = \begin{array}{|c|c|c|} y & a & x \\ a & z & a \\ z & y & z \\ x & x & y \\ \hline 1 & 2 & 3 \end{array}.$$

Alternative x is rejected by agent 1 or agent 2. Alternative y is rejected by agent 3. Alternative z is rejected by the coalition {1,2,3}. No coalition rejects alternative a. So $C(B, R_N) = \{a\}$.

In the case of the profile R'_N below,

$$R'_N = \begin{array}{|c|c|c|} x & y & z \\ y & z & x \\ z & x & y \\ a & a & a \\ \hline 1 & 2 & 3 \end{array},$$

the core consists of x, y and z.

We now introduce a useful concept that sometimes helps check the stability of some alternative a.

(4.1.5) Definition. Assume an alternative a and a profile R_N. *A supporting scheme* for this alternative is a family of coalitions $K(x), x \in A \backslash \{a\}$, such that

a) aR_ix for any $i \in K(x)$, $x \in A\backslash\{a\}$, i.e. the coalition $K(x)$ consists of "opponents" of x;

b) $K(X) = \cup_{x\in X} K(x)$ blocks X for any subset $X \subset A\backslash\{a\}$.

(4.1.6) Lemma. *If there exists a supporting scheme for alternative a given the profile R_N, then $a \in C(B, R_N)$.*

Proof. Take an arbitrary coalition K and pose $L = L(a, R_K)$. If $x \notin L$ then $x \succ a$ for all members of K; thus $K(x)$ and K do not intersect. Therefore the coalition $K(\overline{L})$ does not intersect K, and according to b) it blocks $\overline{L}$. Hence K does not block L and a is stable. ■

The complete relation between supporting schemes and blockings is exposed below.

(4.1.7) Lemma. *Let B be a blocking. The following two assertions are equivalent:*

1) *for every preference profile there exists a supporting scheme for some alternative;*

2) *the blocking B is maximal and admits stable outcomes for every preference profile.*

Proof. 1) $\Longrightarrow$ 2). By virtue of lemma (4.1.6), we need to check the maximality of the blocking. Consider the following profile

$$\begin{array}{|c|c|} \overline{X} & X \\ X & \overline{X} \\ \hline K & \overline{K} \end{array}.$$

According to 1) there exists a supporting scheme K for an alternative (denoted by a). And assume that $a \in \overline{X}$. Then according to Definition (4.1.5a) $K(x) \subset K$ for any $x \in X$ and according to b) $K(X)BX$. It follows from this and axiom B1 that KBX.

2) $\Longrightarrow$ 1). Let B be maximal and $a \in C(B, R_N)$. Consider the family $(K(x),\ x \in A \backslash \{a\})$ with

$$K(x) = \{i \in N, aR_ix\}.$$

We check now that this family satisfies the condition b) of Definition (4.1.5). Let X be a subset of $A\backslash\{a\}$; and take the coalition $K = \overline{K(x)}$. $X \succ a$ for all members of K, therefore $L(a, R_K) \subset \overline{X}$. If $K(X)$ does not block X, then, by maximality, K blocks $\overline{X}$ and moreover $L(a, R_K)$. This contradicts the stability of a.■

(4.1.8) Stable Blockings. A blocking B is *stable* if $C(B, R_N) \neq \emptyset$ for any preference profile $R_N \in \mathbf{L}^N$.

Stability can be understood as consistency (see (1.4.3)) with respect to the concept of equilibrium with threats. The natural following questions arise: do stable blockings exist, are there many of them, how can stability be checked, how do we construct stable blockings? Let us start by examining a few blockings, whose stability can be established in an elementary way.

(4.1.9) Example. *Oligarchy.* Let $O \subset N$ be a non-empty coalition; we can always associate to O the following "oligarchic" blocking B: a coalition K blocks a set $X \neq A$ if and only if $K \supset O$; otherwise it blocks nothing (except $\emptyset$, see Example (1.5.7)). All in all this means that unanimity should prevail within the oligarchy O. (Note two particular cases of oligarchy: dictatorship ($O = \{i\}$) and unanimity ($O = N$).) It is easy to understand that the blocking is stable. In effect the core here is the set of Pareto optimal outcomes for the oligarchy O :

$$C(B, R_N) = Par(R_O).$$

This set is always non-empty.

(4.1.10) Example. Any blocking with two agents is stable. In (1.5.8) we introduced the notion of individual core to be the set $IC(B, R_N)$ of those alternatives a, which are rejected by no agent. Clearly this set is always non-empty. To compute the core we need however to consider one last coalition, $N = \{1, 2\}$. Therefore,

$$C(B, R_N) = Par(R_N \mid IC(B, R_N))$$

which is also non-empty.

Identically, if A consist of two alternatives, then any blocking is stable.

(4.1.11) Example. *Maskin blocking.* There is a blocking with many agents where the core is given by the formula from Example (4.1.10), i.e. where the only genuine coalitions are formed by either singletons or the whole group N. Fix an alternative a. Now form the blocking B for which every agent blocks any subset $X \subset A$, which does not contain a, and for which the only coalition which is able to block a is the whole group. In this case, the individual core $IC(B, R_N)$ coincides with the set $U(a, R_N)$ from Example (1.3.3), and the core

$$C(B, R_N) = Par(R_N \mid IC(B, R_N))$$

coincides with $M(R_N)$ from Example (1.3.4).

(4.1.12) The Core Correspondence. The *core correspondence* of the blocking B,

$$C(B, \cdot) : \mathbf{L}^N \Longrightarrow A,$$

associates to every profile R_N the core $C(B, R_N)$.

The blocking B is stable when the $C(B, \cdot)$ is non-empty valued. We now discuss a few useful properties of core correspondences.

1. *The core correspondence is monotone* (see (1.3.2)). We check that if $a \in C(B, R_N)$ and $R'_N \succeq_a R_N$ then $a \in C(B, R'_N)$. Indeed, for any coalition K we have $L(a, R'_K) \supset L(a, R_K)$. Since K does not block $L(a, R_K)$ then neither does it block $L(a, R'_K)$.

2. *The core correspondence is strictly monotone* ((2.3.5)). Due to the monotonicity of $C(B, .)$ we need only check that for any $X \subset A$ and $i \in N$ the set $\mathrm{Ess}_i(X)$ is empty or equal to X. In this case, we can even formulate explicitely this set

$$\mathrm{Ess}_i(X) = \begin{cases} \emptyset & \text{if } \{i\}BX \\ X & \text{otherwise} \end{cases} .$$

Indeed if agent i blocks X then all elements X are i-non-essential and $\mathrm{Ess}_i(X) = \emptyset$. From the other side, let $x \in X$ be an $i-$non-essential alternative (for $C(B, .)$). This means that for the following profile R_N

$\overline{X}$	x
x	$*$
$X\backslash\{x\}$	$*$
i	$N\backslash\{i\}$

x does not belong to $C(B, R_N)$, that is x is rejected by some coalition. The only coalition that can block X in this case is that formed by agent $\{i\}$.

3. *If* $B \subset B'$ *then* $C(B, \cdot) \supset C(B', \cdot)$. This is obvious.

4. Let B be a stable blocking. Then *the blocking generated by the correspondence* $C(B, \cdot)$ (see (1.5.6)) *contains* B, $B \subset B_{C(B,\cdot)}$.

Indeed, let KBX. If the coalition K uses a profile $R_K = (\overline{X} \succ X)$ then given any profile of the complementary coalition, the coalition K rejects all alternatives from X. Hence $C(B, (R_K, *)) \subset \overline{X}$. This means that $KB_{C(B,\cdot)}X$.

Core correspondences are interesting tools for the construction of (direct) mechanisms. Let a blocking B be stable and $f : \mathbf{L}^N \to A$ be a selector from $C(B, \cdot)$, i.e.

$$f(R_N) \in C(B, R_N)$$

for every $R_N \in \mathbf{L}^N$. Consider f as a direct mechanism. Property 4 implies that the mechanism f is readily compatible with B in the sense that $B \subset B_f$. If additionally, the blocking B is maximal, then $B = B_f$ and stable (so that f is a core mechanism in the sense (5.4.2)). However when B is non-maximal, the blocking B_f can be non-stable.

(4.1.13) Example. We consider again a case with three agents and three alternatives; the blocking B is generated by an oligarchy $O = \{2, 3\}$, as in Example (4.1.9). Assume the selector f takes the following form,

$$f(R_N) = \max(R_1 \mid C(B, R_N)).$$

Then B_f coincides with the simple majority rule, which on top is non-stable (see Example (4.1.3)).

In Theorem (4.5.11) we give a few useful characterizations of stable and maximal blockings. They are based on various necessary and sufficient stability conditions, which we shall elaborate in Sections 4.3 and 4.4. But beforehand, we wish to familiarize the reader with a few classes of blockings which will help him through the coming sections.

4.2 Additive Blockings

(4.2.1) We discuss here some possible ways of representating blockings. We have already seen two of them. The first (and universal) representation is based on mechanisms (see (1.5.2)). It is not the most convenient though, because blockings are simpler and coarser objects than mechanisms. And it seems slightly paradoxical to represent simple things with complex objects.

The second representation is based on simple games (Example (1.5.7)). Note that simple games yield very specific power allocations for coalitions: a coalition will be either all powerful or powerless. Such a strong dichotomy in the degree of power of a coalition often leads to the non-stability of such blockings (see Proposition (4.5.2)). We discuss now an alternative representation which allows for smoother variations in the coalition's power.

Let $\mu : N \to \mathbf{Z}_+$ and $\beta : A \to \mathbf{Z}_+$ be two integral functions. In other words, we associate a non-negative integer $\beta(x)$ (the weight of alternative x) to every alternative $x \in A$ and associate a non-negative integer $\mu(i)$ (the weight of agent i) to every agent $i \in N$. We suppose in the sequel that

$$\beta(A) > \mu(N).$$

(From now on, denote by $\mu(K) = \sum_{i \in K} \mu(i)$ for a coalition $K \subset N$, and analogously $\beta(X)$ for $X \subset A$.)

Define now the blocking $B = B_{\mu,\beta}$

$$KBX \Leftrightarrow \mu(K) \geq \beta(X).$$

In other words, a coalition K blocks a set X as soon as the weight of K is superior or equal to the weight of X. The reader will be easily convinced that this is really a blocking, that is that axioms B1-B3 are fulfilled. Such blockings are called *additive blockings.* Note that neither the weight μ nor β are defined uniquely for a given blocking B, but this is not essential.

The sovereignty property B4 is fulfilled if

$$\beta(A) \leq \mu(N) + \min_{x \in A}\{\beta(x)\}.$$

An additive blocking is maximal if

$$\beta(A) = \mu(N) + 1.$$

Indeed if K does not block X, then $\mu(K) < \beta(X)$. And since both numbers are integers then $\mu(K) \leq \beta(X) - 1$. But then

$$\mu(\overline{K}) = \mu(N) - \mu(K) \geq \beta(A) - 1 - \beta(X) + 1 = \beta(\overline{X})$$

and so $\overline{K}$ blocks $\overline{X}$.

Additive blockings, besides the simplicity of their representation, are interesting because they happen to be stable.

(4.2.2) Theorem (Moulin, Peleg). *Additive blockings are stable.*

Proof. We provide an explicit procedure which yields a stable alternative for any profile (a more general case is discussed in Section 4.6). Imagine that every agent i is given $\mu(i)$ tokens, which he can pile up on alternatives. If alternative x has received a number of tokens $\geq \beta(x)$, it is "propped down" and eliminated. One can think of many different rounds by which agents might distribute their tokens among alternatives. Here we assume that the first agent has to lay out all his tokens, before the second can do so, and so on. Of course to do so we have to start by enumerating our agents, $N = \{1, 2, ..., n\}$.

Agents, in laying out their tokens should be guided by a natural and simple rule: to lay their next token on their worst alternative among those that have not yet been eliminated. An alternative x is eliminated as soon as $\beta(x)$ tokens are piled up on it. As the last agent lays his tokens, there will remain some non-eliminated alternatives (since $\beta(A) > \mu(N)$). We claim that anyone of these remaining alternatives lies in the core.

Let us prove it. Let a be a non-eliminated alternative. Given our definition of a round, every agent i piles his tokens on alternatives ranking worse than a or on a, i.e. on alternatives from $L(a, R_i)$. Indeed he might pile his tokens on alternatives ranking higher than a only after eliminating a. Suppose now that some coalition K blocks the set $L(a, R_K)$ so that $\mu(K) \geq \beta(L(a, R_K))$. We have just seen that the coalition K piled its $\mu(K)$ tokens on alternatives from $L(a, R_K)$. Since by our assumption $\mu(K) \geq \beta(L(a, R_K))$, then there should be no less than $\beta(a)$ tokens piled on alternative a, which means a should be eliminated as well. This contradiction proves the assertion and the theorem.■

(4.2.3) Remark. This procedure underlies the construction of strongly consistent mechanisms ((5.2.1)), which provide one alternative proof to the Moulin-Peleg theorem. There are two other proofs of this theorem. The first one follows from a more general result about the stability of convex blockings (see (4.4.4) and (4.3.4)). The second one is slightly more intricate and we only sketch it here. The point is that additive blockings are balanced in the sense of H. Scarf. And then, by the Scarf theorem, balanced games without side payments $G(B, R_N)$ have non-empty cores, see Appendix 4.A1.

Many stable blockings encountered above were additive, but not all of them. Consider two examples.

(4.2.4) Example. Take the Maskin blocking (from Example (4.1.11)) with a fixed alternative a. Let us show that it is additive, thus stable.

We set the weight of all alternatives $x \neq a$ to be equal to 1. All agents are given the same weights, which is equal to $m - 1$, where $m = |A|$. The weight of a is determined from the equation

$$\beta(A) = \mu(N) + 1,$$

$\beta(a) = (n-1)(m-1)+1$. The reader will easily check that this set of weights yields precisely the Maskin blocking.

(4.2.5) Example. Assume that $N = \{1, 2, 3, 4, 5, 6\}$,and that A consists of two alternatives. The blocking B is given by the following simple game W.

A coalition K is a wining coalition if $\mid K \mid \geq 4$ or $\mid K \mid = 3$ and $K = \{i, j, k\}$ where the sum of numbers $i + j + k$ is even. Since $1 + 2 + ... + 6 = 21$ is odd the simple game W and the corresponding blocking B are maximal. B is stable for there are only two alternatives. However one can check that the blocking is not additive.

We should distinguish two subclasses of additive blockings.

(4.2.6) Peleg blockings. These blockings have the following systems of weights μ and β : where μ is identically equal to 1 ($\mu(i) = 1$ for any $i \in N$) and

$$\beta(A) = \mu(N) + 1 = \mid N \mid +1.$$

Peleg blockings are "natural" for $\mid A \mid \ll \mid N \mid$. Their core correspondence possesses an additional specific property.

(4.2.7) Proposition. *If B is a Peleg blocking then its core correspondence $C(B, \cdot)$ is a minimal monotone SCC.*

Proof. One should check minimality of $C(B, .)$. To do so, we use Moulin's criterion (1.A3.2). We show that for any $a \in C(B, R_N)$ there is a profile $R'_N \approx_a R_N$ such that $C(B, R'_N) = \{a\}$.

We will do it as follows. For $x \neq a$, pose

$$K(x) = \{i \in N, aR_i x\}$$

(i.e. a coalition of "opponents to" x). As it was shown in lemma (4.1.7), for any set $X \subset A \backslash \{a\}$, the following inequality obtains

$$\mid K(X) \mid \geq \beta(X).$$

These inequalities represent a necessary and sufficient condition in the "Harems lemma" (see Appendix 4.A3). According to this lemma, we exhibit subcoalitions $\Pi(x) \subset K(x), x \neq a$, such that

a) the coalitions $\Pi(x)$ do not intersect each other;

b) $\mid \Pi(x) \mid = \beta(x)$ for any $x \neq a$.

Now we change the preferences of every agent i, belonging to $\Pi(x)$, from R_i to R'_i having propped x down,

$$R'_i = (R_i \mid A \backslash \{x\}, x).$$

Since $aR_i x$ then $R'_i \approx_a R_i$. Therefore the new profile $R'_N = (R'_i)$ is a-equivalent to R_N. Finally the coalition $\Pi(x)$ rejects x at the profile R'_N hence $C(B, R'_N) = \{a\}$.■

The following example shows that the Peleg blocking assumption is essential to the result.

(4.2.8) Example. Assume that $N = \{1, 2, 3, 4\}$, and take three alternatives. We set the weight of every alternative to be equal to 2; $\mu(1) = 2$,

$\mu(2) = \mu(3) = \mu(4) = 1$. This is a maximal additive blocking. However the core correspondence $C(B, .)$ is not a minimal monotone SCC. In fact, consider the profile R_N :

x	y	z	z
y	x	x	y
z	z	y	x
1	2	3	4

The reader will easily be convinced that $C(B, R_N) = \{x, y\}$. However, there exists no profile $R'_N \preceq_x R_N$ such that $C(B, R'_N) = \{x\}$. One can check it straightforwardly.

The Peleg blocking belongs to a wider class of *anonymous* blockings, where all agents have equal blocking power. Anonymous (not necessarily additive) blockings are given by a function $b : 2^A \to \mathbf{Z}_+$, where $b(x)$ is the minimal size of the coalition, which is able to block the set X.

(4.2.9) Moulin Blockings. They are dual to Peleg blockings, in the following sense: alternatives and agents change places. More exactly Moulin blockings are blockings $B_{\mu,\beta}$ for which the function β is identically equal to 1 (i.e. $\beta(x)f = 1$ for any alternative $x \in A$) and

$$| A |= \beta(A) = \mu(N) + 1.$$

Moulin blockings are "natural" for $| A |\gg| N |$. They also have a specific property.

(4.2.10) Proposition. *If B is a Moulin blocking, then its core correspondence $C(B, .)$ is a minimal monotone SCC.*

The proof ressembles that of Proposition (4.2.7), where one should replace "horizontal" blocks $(KB\{x\})$ by "vertical" ones $(\{i\}BX)$. We again use Moulin's criterion. Let $a \in C(B, R_N)$; for agent i, pose

$$X(i) = L(a, R_i)\backslash\{a\} = \{x \in A, a \succ_i x\}.$$

Then, for any coalition $K \subset N$, we have $\mu(K) \leq| X(K) |$. Indeed, otherwise

$$\mu(K) \geq| X(K) | +1 =| X(K) \cup \{a\} |$$

and K blocks $L(a, R_K) = \{a\} \cup X(K)$ that contradicts the stability assumption about a.

Again by the " Harems lemma", to every agent i we associate a set $Z(i) \subset X(i)$ such that:

a) the sets $Z(i)$ are disjoint for different i;

b) $| Z(i) |= \mu(i)$.

We modify again R_i, having propped $Z(i)$ down:

$$R'_i = (R_i|A\backslash Z(i), R_i|Z(i)).$$

It is clear that $R'_N \approx_a R_N$. We affirm that $C(B, R'_N) = \{a\}$. Every alternative from $Z(i)$ is rejected by agent i at the profile R'_N by virtue of b). Moreover by a)

$$| Z(N) | = \sum_i | Z(i) | = \sum_i \mu(i) = \mu(N) = | A | -1,$$

hence $Z(N) = A \backslash \{a\}$.

The sole non-rejected alternative is a.■

Moulin blockings belong to the class of neutral blockings. Now we consider a class of blockings (not necessarily additive) which are anonymous and neutral.

(4.2.11) Anonymous Neutral Blockings. The fulfilling of the relation KBX depends only on $| K |$ and $| X |$ for anonymous neutral blockings. Therefore, they can be characterized by its veto-function (see (1.5.8))

$$v : \{0, 1, ..., n\} \rightarrow \mathbf{Z}_+,$$

where $v(k)$ means that a $k-$ member coalition blocks any set of alternatives of size $\leq v(k)$. Of course, the function v defines a blocking if the following two conditions are fulfilled:

1) $v(n) < | A |$;

2) if $k_1 + k_2 \leq n$ then $v(k_1 + k_2) \geq v(k_1) + v(k_2)$.

It turns out that we can give a very simple stability criterion for these blockings using a *proportional veto-function* v^*. This function is defined as follows:

$$v^*(k) = [k \cdot |A| / n - 1] = [k \cdot |A| / n] - 1,$$

where $[t]$ is the smallest integer superior or equal to t. Or in other words, $v^*(k)$ is the greatest integer inferior to $k \cdot |A| / n$.

Roughly speaking the blocking force of a coalition is proportional to its size, $v^*(k) \sim k \cdot |A| / n$. We affirm that the blocking associated to v^* is additive. Indeed, let d be the greatest common divisor of n and $m = | A |$. The equation

$$m\beta = n\mu + d$$

has a solution with arbitrarily large natural numbers μ and β. Assume that (μ, β) is such a solution. Fix agents' weights to be equal to μ and alternative weights to be equal to β, then the corresponding blocking coincides with that generated by v^*. This with Theorem 1 yields the stability of v^*. Of course, functions $v \leq v^*$ are also stable. The converse also holds.

(4.2.12) Theorem (Moulin). *An anonymous and neutral blocking with veto-function v is stable if and only if $v \leq v^*$.*

We do not prove this result (the reader should refer to Moulin (1988)), but we explain briefly in next Section why it is true.

(4.2.13) This theorem is important one. The proportional veto principle is based on fairness considerations, it expresses "minority rights". Coalitions' power is proportional to their size, whether big or small. However Theorem (4.2.12) comes as an unexpected support to this principle from the viewpoint of stability. If we want to allocate power among coalitions (in an anonymous and neutral way) in order to warrant the existence of stable outcomes, we end up with the proportional veto principle.

Moulin (1983) opposes the "minority" principle to the "majority" principle at the core of the Condorcet rule. We illustrate this in two examples.

First example: assume society is divided into two groups K and $\overline{K}$ of approximately equal size, where the first is slightly bigger than the second. Let the preferences inside groups be identical, but assume that they are opposite in different groups, for example

$a \succ b \succ c \succ d \succ e$ for agents from K, and

$e \succ d \succ c \succ b \succ a$ for agents from $N \backslash K$.

According to the majority principle, the outcome will be a. If on the contrary, we use the minority principle, then the outcome will be the middle alternative c. Note also that the alternative c is stabler with respect to small changes in the respective cardinalities of K and $\overline{K}$.

Second example: let there be three agents and four alternatives. Take the following preference profile

a	a	d
b	d	b
c	c	c
d	b	a
1	2	3

.

In this case, we have a Condorcet winner: alternative a. Nevertheless the unique stable outcome, for proportional veto, is alternative c, though it loses for pairwise comparisons with any other alternative.

Note also that proportional veto is maximal if and only if the numbers m and n are coprime.

4.3 Convex Blockings

In this section we consider another stability providing property of blockings.

(4.3.1) Definition. A blocking B is *convex* if the following property holds: $K_1 B X_1, K_2 B X_2$ implies $(K_1 \cup K_2) B (X_1 \cup X_2)$ or $(K_1 \cap K_2) B (X_1 \cap X_2)$.

We give two example of convex blockings.

(4.3.2) Example. Additive blockings are convex. This can be seen from the equalities

$$\mu(K_1) + \mu(K_2) = \mu(K_1 \cup K_2) + \mu(K_1 \cap K_2)$$

and

$$\beta(X_1) + \beta(X_2) = \beta(X_1 \cup X_2) + \beta(X_1 \cap X_2).$$

(4.3.3) Example. Let the blocking B be generated by an oligarchy as in Example (4.1.9). Then obviously B is convex. Conversely, if a convex blocking is generated by a simple game and $\mid A \mid \geq 3$, then the simple game is an oligarchy. We leave it to the reader to check this simple fact.

Notice that both additive blockings and oligarchic blockings are stable. This suggests that convex blockings are also stable.

(4.3.4) Theorem (Peleg). *Convex blockings are stable.*

We mention here two possible proofs of this crucial result. The first was proposed by Peleg (1984), and consisted in using the cooperative game without side payments $G(B, R_N)$. This game is convex (for a definition of convex NTU-games, see Vilkov (1977) or Greenberg (1985)); by Vilkov's theorem its core is not empty.

The second is more straightforward and operates with alternatives and blockings. The proof is carried out by induction on the number of agents. Given a profile R_N we construct a new blocking B' for the group $N' = N \backslash \{i\}$ by eliminating the agent i.

(4.3.5) Elimination of an Agent. Assume that this agent is agent 1 without loss of generality. In order to define this new blocking B', we only need to know the preferences R_1 of agent 1. Let L be the maximal lower contour of R_1, which agent 1 can block. That is $\{1\}BL$, and if $a = \min(R_1 \mid \overline{L})$ agent 1 does not block the set $L(a, R_1) = \{a\} \cup L = L^+$. Let $N' = N \backslash \{1\}$, $A' = A \backslash L$ and define the relation B' between $2^{N'}$ and $2^{A'}$, for $K' \subset N'$ and $X' \subset A'$, as follows

$$K'B'X' \Leftrightarrow \text{either } K'BX' \text{ or } (K' \cup \{1\})B(X' \cup L^+).$$

In other words, the coalition K' can use the help of agent 1 if it, in its turn, helps him to block $L(a, R_1)$.

(4.3.6) Lemma. *Let B be a convex blocking. Then*
a) *B' is also a blocking;*
b) *the blocking B' is convex*;
c) $C(B', R_{N'} \mid A') \subset C(B, R_N)$.

It is easy to see that Theorem (4.3.4) follows from this lemma, since by inductive assumption the core $C(B', R_{N'} \mid A')$ is not empty. The initial step of the induction is true, since any blocking with one (or even two) agents is stable. We develop (namely in Section 4.6) this inductive line of reasoning into a procedure aimed at selecting stable elements. Theorem (4.3.4) and Example (4.3.2) yield the Moulin-Peleg theorem (4.2.2).

Proof of the lemma.

a) Axioms B1 and B3 are obvious enough. Thus we merely verify superadditivity. Let coalitions K_1', $K_2' \subset N'$ be disjoint and $K_j' B X_j', j = 1, 2$; one should check that $(K_1' \cup K_2') B' (X_1' \cup X_2')$. Three of the four arising cases are trivial. We only look at the fourth, namely $(K_j' \cup \{1\}) B (X_j' \cup L^+), j = 1, 2$. We need to used the convexity of B here. Since $(K_1' \cup \{1\}) \cap (K_2' \cup \{1\}) = \{1\}$ and $(X_1' \cup L^+) \cap (X_2' \cup L^+) \supset L^+$ and $\{1\} \bar{\bar{B}} L^+$, we conclude that $(K_1' \cup K_2' \cup \{1\}) B (X_1' \cup X_2' \cup L^+)$, i.e. $(K_1' \cup K_2') B' (X_1' \cup X_2')$.

b) We verify convexity in an analogous way. Let $K_j' B' X_j', j = 1, 2$. Again we have four cases to check but we only check one, where $K_1' B X_1'$ and $(K_2' \cup \{1\}) B (X_2' \cup L^+)$. By convexity of B we have $(K_1' \cup K_2' \cup \{1\}) B (X_1' \cup X_2' \cup L^+)$, i.e. $(K_1' \cup K_2') B' (X_1' \cup X_2')$, or $(K_1' \cap (K_2' \cup \{1\})) B (X_1' \cap (X_2' \cup L^+)$. Since $(K_1' \cap (K_2' \cup \{1\}) = K_1' \cap K_2'$ and $X_1' \cap (X_2' \cup L^+) \supset X_1' \cap X_2'$, then $(K_1' \cap K_2') B' (X_1' \cap X_2')$.

c) Let $x \in A'$ and suppose that it is rejected (for the blocking B) by some coalition $K \subset N$, i.e. $KBL(x, R_K)$. One has to show that x is rejected for the blocking B'. We consider two cases separately. In the first one, assume that $1 \notin K$. Then $K \subset N'$ and since $L(x, R_K \mid A') \subset L(x, R_K)$ then $KB'L(x, R_K \mid A')$ and $x \notin C(B', R_{N'} \mid A')$. In the second case, we assume that $1 \in K$. Since $x \in A'$ then $L(a, R_1) \supset L^+$ and so $KB(L(x, R_1) \cup L^+)$. If we set $K' = K - \{1\}$ then $K'B'L(x, R_{K'} \mid A')$.■

(4.3.7) Who Rejects Unstable Outcomes? Convex blockings possess one additional interesting property, which was obtained by Demange (1987).

(4.3.8) Theorem (Demange). *Let B be a convex blocking and R_N be a preference profile. Then for any alternative x outside of $C(B, R_N)$, there exists an alternative $a \in C(B, R_N)$ and a coalition K such that*

a) *$aR_i x$ for any $i \in K$,*

b) *coalition K rejects x.*

Thus every unstable alternative is rejected by a coalition of agents who have a preferred stable alternative.

Proof. Let an alternative x be rejected by a coalition K and K be minimal among those rejecting x. We show how to find an alternative $a \in C(B, R_N)$ satisfying property a). We use here an argument reminiscent of that used in the proof of Theorem (4.3.4).

We know that the coalition K blocks the set $L(a, R_K)$. Choose some agent in K, for instance 1, and let L_1 be the maximal lower contour of R_1 such that K blocks $L = L_1 \cup L(a, R_K)$. Let $N' = N \backslash K$, $A' = A \backslash L$ and define a new blocking B' on the pair (N', A'). To do so, we define $a_i = \min R_i \mid A'$ and $Z(i) = L(a, R_i)$ for agent $i \in K$. Finally, given coalition $S \subset N'$ and a set $X \subset A'$, we define as follows: $SB'X \Leftrightarrow$ there exists a coalition $T \subset K$ such that $(S \cup T)B(X \cup Z(T))$, where as usual $Z(T) = \cup_{i \in T} Z(i)$.

By the same line of arguments as used in the proof of the previous lemma, it is easy to check that B' is a convex blocking. We do not go through the somewhat tedious computations and check only that $N' \ \overline{B}' \ A'$. Conversely, suppose that $N'B'A'$, i.e. that $(N' \cup T)B(A' \cup Z(T))$ for some coalition $T \subset K$. We know that K blocks $L = A \backslash A'$. Thus we conclude both from convexity of B and axiom B3, that the coalition $(N' \cup T) \cap K = T$ blocks the set $(A' \cup Z(T)) \cap (L_1 \cup L(x, R_K))$, which obviously contains $L(a, R_T)$. But this means that $N' \cup K$ blocks $A' \cup Z(K) = A$ which contradicts axiom B3.

We established that B' is a convex blocking. By Theorem (4.3.4), the core $C(B', R_{N'} \mid A')$ is not empty: it contains some alternative a. It is clear that aR_ix for all $i \in K$ and it remains to be checked that $a \in C(B, R_N)$. Suppose the converse, that a is rejected by some coalition $S \subset N$. We assert then that a is rejected (for the blocking B') by a coalition $S' = S \backslash K$. This contradicts the stability of a. In fact, since aR_ia_i for $i \in S \cap K$ then $L(a, R_i) \supset Z(S \cap K)$, i.e. $S'B'L(a, R_{S'} \mid A')$.■

4.4 Almost Additive Blockings

(4.4.1) Let an additive blocking be given by its weights μ and β as in (4.2.1). Then it has the following qualitative (or structural) property. Take two coalitions K_1 and K_2 where K_1 does not block X_1, and K_2 does not block X_2 and assume that $X_1 \cap X_2 = \emptyset$; then $K_1 \cup K_2$ does not block $X_1 \cup X_2$. Indeed, since K_j does not block X_j then $\mu(K_j) < \beta(X_j)$, $j = 1, 2$. Therefore

$$\mu(K_1 \cup K_2) \leq \mu(K_1) + \mu(K_2) < \beta(X_1) + \beta(X_2) = \beta(X_1 \cap X_2).$$

(4.4.2) Definition. A blocking B is *almost-additive* if $K_1 \overline{B} X_1$, $K_2 \overline{B} X_2$ and $X_1 \cap X_2 = \emptyset$ imply $(K_1 \cap K_2)\overline{B}(X_1 \cap X_2)$.

Note that by monotonicity of the blocking, the coalitions K_1 and K_2 in Definition (4.4.2) can be thought of as being disjoint.

The requirement of almost-additivity hinders the sharp growth of blocking power when coalitions join. For example, it is easy to check the following fact. Let $K_1 \cup K_2$ block a set $X \subset A$. Then there exist two subsets $X_1, X_2 \subset X$ such that K_jBX_j for $j = 1, 2$ and X differs from $X_1 \cup X_2$ by not more than one element.

(4.4.3) Lemma. *Any almost-additive blocking is convex.*

Proof. Let K_1BX_1 and K_2BX_2 but $K_1 \cap K_2$ does not block $X_1 \cap X_2$. By almost-additivity of B the coalition $K_1 \backslash (K_1 \cap K_2)$ blocks $X_1 \backslash (X_1 \cap X_2)$. Then by superadditivity (axiom B2) coalition $K_2 \cup (K_1 \backslash K_2) = K_1 \cup K_2$ blocks $X_2 \cup (X_1 \backslash X_2) = X_1 \cup X_2$. ■

Lemma (4.4.3) and Theorem (4.3.4) yield the following corollary.

(4.4.4) Corollary. *Any almost-additive blocking is stable.*

Unlike general stable blockings, almost-additive blockings have a additional feature, namely that blocking coalitions can be arranged in a "nested structure". We now define this idea, but first introduce the useful (especially for Sections 5.4 and 5.5) notion of elimination scheme.

(4.4.5) Definition. Let B be a fixed blocking. An *elimination scheme* for a set $X \subset A$ is a mapping $K : X \longrightarrow 2^N$ which satisfies two conditions:

C1. For any $x, x' \in X$ either $K(x) \cap K(x') = \emptyset$ or $K(x) \subset K(x')$, or $K(x') \subset K(x)$.

C2. Every coalition $K(x)$, $x \in X$, blocks the set $\{y \in X,\ K(y) \subset K(x)\}$.

An elimination scheme can thus be thought of as a way to allocate the tasks of blocking alternatives from X among coalitions. Given an alternative x, the coalition $K(x)$ will be responsible for ensuring the blocking of x. This coalition not only blocks x, but also helps reinforce the blocking power of all coalitions which include it (according to the hierarchy principle enclosed in condition C1).

The following lemma formulates a few useful properties of elimination schemes.

(4.4.6) Lemma. *Let K be an eliminating scheme for a set X, and Y be a subset of X. Then $K \mid Y$ is an eliminating scheme for Y and the coalition $K(Y) = \cup_{y \in Y} K(y)$ blocks Y.*

Proof. The first assertion is obvious. In order to check the second, one can take $Y = X$. Choose among the coalitions $K(x)$, $x \in X$, those that are maximal by inclusion and call them $K_1, ..., K_r$. Obviously $K(X) = K_1 \cup ... \cup K_r$. By C1 the coalitions $K_1, ..., K_r$ are pairwise disjoint. So by B2 and C2, the coalition $K(X)$ blocks $L_1 \cup ... \cup L_r$, where $L_j = \{x \in X,\ K(x) \subset K_j\}$. Since every alternative $x \in X$ lies into one of the L_j, then $K(X)$ blocks X. ■

We return to almost-additive blockings.

(4.4.7) Theorem. *Let B be an almost-additive blocking, $R_N \in \mathbf{L}^N$, and $X \subset A$. Assume that for every $x \in X$ there is a coalition $S(x)$ rejecting x. Then there exists an elimination scheme $K : X \to 2^N$ such that $K(x) \subset S(x)$ for any $x \in X$.*

Proof. We proceed by induction on the number of alternatives $\mid A \mid$. The case $\mid A \mid = 1$ is obvious. For convenience, we partition the proof into several steps.

Step 1. For $i \in N$, let us denote $X_i = \{x \in X, i \in S(x)\}$. Now define the new profile R'_N by setting $R'_i = (R_i \mid \bar{\bar{X}}_i, R_i \mid X_i)$, i.e. where the sets X_i are being propped down for the preferences R_i. As before the coalitions $S(x)$

reject x at the profile R'_N, since $L(x, R'_{S(x)}) \subset L(x, R_{S(x)})$. In replacing R_N by R'_N we may assume that $\min R_i \in X_i$ for every $i \in S(x)$.

Step 2. If $X = \emptyset$, the assertion in the theorem is trivial. Otherwise, take some $x \in X$ and for every $y \in A$ define the following coalitions

$$\Pi(y) = \{i \in S(x), \min R_i = y\}.$$

Let $Y = \{y \in A, \Pi(y) \neq \emptyset\}$; according to step 1, $Y \subset X$. Clearly $Y \subset L(x, R_{S(x)})$, therefore coalition $S(x) = \cup_{y \in Y} \Pi(y)$ blocks Y. Since the coalitions $\Pi(y)$ are disjoint and B is almost-additive, we conclude that $\Pi(y) B\{y\}$ for some $y \in Y$.

Step 3. Take an alternative $y \in Y$ such that $\Pi(y) B\{y\}$. By step 1, $\Pi(y) \subset S(y)$. We denote by K_0 the minimal subcoalition in $\Pi(y)$, which still blocks y. We may assume that $S(y) = K_0$ shrinking $S(y)$ if needed.

As for the remaining coalitions $S(x)$, we can assume them to be minimal with respect to the fulfilling of the property $S(x) BL(x, R_{S(x)})$ shrinking them (if needed). Then the following "absorption" property obtains:

$$\text{if } K_0 \cap S(x) \neq \emptyset \text{ then } K_0 \subset S(x).$$

Indeed, let $K' = K_0 \cap S(x)$, $S' = S(x) \backslash K_0$. Then $L(x, R_{S(x)}) = L(x, R_{K'}) \cup L(x, R_{S'}) \supset \{y\} \cup L(x, R_{S'})$.By step 1, $y \notin L(x, R_{S'})$. Therefore the coalition $S(x) = K' \cup S'$ blocks $\{y\} \cup L(x, R_{S'})$. By minimality of $S(x)$, coalition S' does not block $L(x, R_{S'})$. Thus due to almost-additivity, we conclude that K' blocks y, that is $K' = K_0$ and hence $K_0 \subset S(x)$.

Step 4. We can assume that the following property holds: the preferences of all agents in K_0 are identical and equal to some preference R_0. Indeed, let R_0 be the preference of some agent in K_0. Let us denote by R'_N the new profile, in which we set the preferences of all agents in K_0 equal to R_0. We affirm that $S(x) BL(x, R'_{S(x)})$ as well, for any $x \in X$. This is trivial when $S(x)$ and K_0 do not intersect. If $K_0 \cap S(x) \neq \emptyset$, then $K_0 \subset S(x)$ by step 3.

Step 5. We can consider reducing coalition K_0 to an individual agent. And we shall assume that coalition K_0 consists of the single agent 0.

Step 6. Now we replace A by $A' = A \backslash \{y\}$ (reminding that $y = \min R_0$). The new blocking B' on the pair (N, A') is given by the following formula where $S \subset N$ and $Z \subset A'$:

$$SB'Z \Leftrightarrow \begin{cases} SBZ & , \text{if} \quad 0 \notin S, \\ SB(Z \cup \{y\}) & , \text{if} \quad 0 \in S. \end{cases}$$

It is easy to check that B' is an almost-additive blocking. A' is smaller than A, therefore, by induction statment applyed to $X' = X \backslash \{y\} \subset A'$, there exists an elimination scheme $K : X' \to 2^N$ such that $K(x') \subset S(x')$ for any $x' \in X'$. Clearly we obtain an elimination scheme for $X = X' \cup \{y\}$ provided we add $K(y) = K_0$ to the previous scheme. ■

This Theorem yields another yet proof of the stability of almost-additive blockings. Indeed, let us suppose that the core of an almost-additive blocking is empty for some profile. Then the above Theorem states that there exists an elimination scheme for the whole set A; however by lemma (4.4.6) some coalition blocks the whole set A, which contradicts axiom B3. This constructive proof of the theorem can be transformed in a procedure aimed at finding an element from the core (see (5.5.2)).

(4.4.8) Corollary. *Let B be an almost-additive blocking, $R_N \in \mathbf{L}^N$, and $X = A \backslash C(B, R_N)$. Then there exists an elimination scheme $K : X \to 2^N$ consistent with the profile R_N in the following sense: for every $x \in X$ one can exhibit an alternative $a_x \in C(B, R_N)$ such that $a_x R_i x$ for every $i \in K(x)$.*

Proof. We showed in Lemma (4.4.3) that B is convex. Hence according to Demange's Theorem (4.3.8) for every alternative $x \in X$, there exists a coalition $S(x)$ and an alternative $a_x \in C(B, R_N)$ such that $S(x)$ rejects x and $a_x R_i x$, for all $i \in S(x)$. It then suffices to extract the scheme K, which exists as a sequel of the Theorem seen above.■

4.5 Necessary Stability Conditions

(4.5.1) In the three previous sections, we established sufficient conditions for stability. And in fact, we found three major classes of stable blockings:

$$\{\text{additive}\} \subset \{\text{almost additive}\} \subset \{\text{convex}\}.$$

Now we discuss a few necessary conditions. In essence, these conditions state that a blocking is not stable if there are "too many" powerful coalitions. In such cases, non-stability is established by exhibiting a profile R_N for which the core $C(B, R_N)$ is empty. Let us return to Example (4.1.3), in which we described a cyclic preference profile leading to an empty core. We can generalize this example. Assume the following partitions of $N = K_1 \sqcup K_2 \sqcup K_3$ and of $A = X_1 \sqcup X_2 \sqcup X_3$, and assume that $K_1 \sqcup K_2$ blocks $X_1 \sqcup X_2$, $K_2 \sqcup K_3$ blocks $X_2 \sqcup X_3$ and $K_1 \sqcup K_3$ blocks $X_1 \sqcup X_3$. Consider the following "cyclic" profile R_N :

X_2	X_3	X_1
X_3	X_1	X_2
X_1	X_2	X_3
K_1	K_2	K_3

Its core is empty. Indeed, the coalition $K_1 \sqcup K_2$ rejects any outcome from X_1, $K_2 \sqcup K_3$ rejects X_2, and $K_1 \sqcup K_3$ rejects X_3.

We call the corresponding necessary stability condition, the *3-cycle condition.* The previous argument can be generalized.

(4.5.2) Proposition. *Let B be a stable blocking. Suppose that we have an integer $m > 0$, some partitions $N = K_1 \sqcup ... \sqcup K_m$, $A = X_1 \sqcup ... \sqcup X_m$, and a mapping $r : \{1, ..., m\} \longrightarrow \{0, ..., m-1\}$. Then for some j the coalition $K_j \sqcup ... \sqcup K_{j+r(j)}$ does not block the set $X_j \sqcup ... \sqcup X_{j+r(j)}$ (where $j + k$ is computed modulo m).*

Proof. Suppose the converse, that is every coalition $K_j \sqcup ... \sqcup K_{j+r(j)}$ blocks the set $X_j \sqcup ... \sqcup X_{j+r(j)}$. Consider the following profile,

$$\begin{array}{|c|c|c|c|}
X_2 & X_3 & \cdots & X_1 \\
\vdots & \vdots & & \vdots \\
X_m & X_1 & \cdots & X_{m-1} \\
X_1 & X_2 & \cdots & X_m \\
\hline
K_1 & K_2 & \cdots & K_m
\end{array}.$$

Its core is empty. Indeed, let us take an alternative x; it belongs to some X_j. Then the coalition $K'_j = K_j \sqcup ... \sqcup K_{j+r(j)}$ rejects x, since $L(x, R_{K'_j}) \subset X_j \sqcup ... \sqcup X_{j+r(j)}$.■

Note that for $r \equiv 0$ the condition, mentioned in Proposition (4.5.2), reduces to the basic property of blockings seen in (1.5.3). Thus the complete set of conditions in Proposition (4.5.2) can be viewed as a generalization of this basic property. However, it might not suffice to warrant stability.

One can establish in a similar way the second assertion in Moulin's theorem (4.2.12) about neutral anonymous blockings, namely if v is not less than v^* then v is unstable. Let us enumerate the alternatives in $A = \{x_1, ..., x_m\}$. Now consider a "cyclic" profile, in which agent i's least preferred alternative is labelled $[im/n]$ and (generally agent i's r-th less preferred alternative is labelled $[im/n] - r + 1$). For example, for $n = 6$,and $m = 4$, this profile has the following form,

$$\begin{array}{|c|c|c|c|c|c|}
x_2 & x_3 & x_3 & x_4 & x_1 & x_1 \\
x_3 & x_4 & x_4 & x_1 & x_2 & x_2 \\
x_4 & x_1 & x_1 & x_2 & x_3 & x_3 \\
x_1 & x_2 & x_2 & x_3 & x_4 & x_4 \\
\hline
1 & 2 & 3 & 4 & 5 & 6
\end{array}.$$

Suppose now that for some $k, 1 \leq k \leq n$, we have $v(k) > v^*(k) = [km/m] - 1$, i.e. $v(k) \geq [km/n]$. Then each alternative is rejected by a suitable coalition composed with k successive agents. In the previous example, suppose that $v(4) \geq [4 \times 4/6] = 3$, i.e., each four agent coalition blocks at most three alternatives. Then x_1 is rejected by the coalition $\{1, 2, 3, 4\}$, x_2 is rejected by the coalition $\{2, 3, 4, 5\}$, x_3 is rejected by the coalition $\{4, 5, 6, 1\}$ and x_4 - by the coalition $\{5, 6, 1, 2\}$.

For a generalisation of this construction procedure of "cyclic" profiles brought by a notion of *acyclic blocking*, (see Danilov and Sotskov (1987a), Keiding (1985), Abdou (1987)). However we shall not detail this issue here and consider only a particular case, in which we generate a blocking from a simple game.

(4.5.3) Simple Games. In (1.5.7), we constructed a blocking B_W from a given simple game $W \subset 2^N$. A numerical characteristic of W, namely the Nakamura number, plays an important role with respect to the stability of B_W. We now give a definition of this number, which differs slightly (by one unit) from the usual one.

(4.5.4) Definition. The *Nakamura number* of a simple game W is the greatest integer $\nu(W)$ (or infinity) such that any number $\nu(W)$ of "winning" coalitions have a non-empty intersection.

(4.5.5) Proposition. *A blocking B_W is stable if and only if $\nu(W) \geq \mid A \mid$.*

Proof. *Necessity.* Let us enumerate the alternatives in $A = \{x_1, ..., x_m\}$. Assume that the m winning coalitions $S_1, ..., S_m$ have an empty intersection. For every agent $i \in N$, let $x_{j+1} P_i x_j$ if $i \in S_j$. We claim that the relations P_i are acyclic. Indeed, if a cycle is to arise here, it should look like

$$x_m P_i x_{m-1} P_i ... P_i x_1 P_i x_m.$$

However this would mean that agent i belongs to every S_j, which contradicts the assumption $S_1 \cap ... \cap S_m = \emptyset$. Thus P_i are acyclic. Now extend the relations P_i to linear orders R_i. We affirm that the core $C(B, R_N)$ is empty. Indeed, every alternative x_j is rejected by the coalition S_j since x_{j+1} is preferred to x_j for all members of S_j and S_j is a winning coalition.

Sufficiency. Let $\nu(W) \geq \mid A \mid$. If the core is empty, then for every alternative $x \in A$ there is an alternative x' and a coalition $K(x) \in W$, such that $x' \succ_i x$ for all $i \in K(x)$. But $\cap_x K(x)$ is non empty by assumption. Let agent i belong to every $K(x)$ and $x = \max R_i$. Then no alternative x' is strictly preferred to x for agent i. The Proposition is thus proved by contradiction.■

We now state a few consequences of this proposition.

(4.5.6) Corollary. *If a blocking B_W is stable for any number of alternatives then $\nu(W) = \infty$.*

In other words, the simple game W is weaker than some oligarchy.■

(4.5.7) Corollary. *Let $|A| \geq 3$, and assume that the blocking B_W is stable and maximal. Then the game W is dictatorial.*

Indeed, in this case $\nu(W) \geq |A| \geq 3$. Therefore any three winning coalitions intersect. So if $K, K' \in W$ then $\overline{K \cap K'} \notin W$ and from maximality $K \cap K' \in W$. To conclude apply lemma (1.3.5). ■

(4.5.8) Corollary. *Let π be an acceptable mechanism (see (2.6.1)), and $\mid A \mid \geq 3$. If the blocking B_π is stable and maximal, then the mechanism π is dictatorial.*

Proof. According to Dutta's corollary (2.6.4), B_π is given by a simple game. Thus as a consequence of the previous corollary, there exists a dictator. ■

(4.5.9) Maximal Stable Blockings. Neither convexity nor almost-additivity are necessary for stability. However for maximal blockings these properties are equivalent to the stability. We present here several different equivalent descriptions of maximal and stable blockings as these types of blockings are important (and will appear so in the Chapter 5).

(4.5.10) Definition. A blocking B is *supermaximal* if for any partitions $N = K_1 \sqcup ... \sqcup K_m$, $A = X_1 \sqcup ... \sqcup X_m$, where $m \geq 2$, there exists $j \in \{1, ..., m\}$ such that $K_j B X_j$.

For $m = 2$ this definition merely expresses the maximality condition.

(4.5.11) Theorem. *Given a blocking B, the following conditions are equivalent:*

1) B is maximal and stable;

2) B is maximal and satisfies the 3-cycle condition;

3) B is maximal and almost-additive,

4) B is maximal and convex;

5) B is supermaximal;

6) for any profile there exists a supporting scheme for some alternative.

In (5.3.8), we will show that these conditions are also equivalent to the following: the blocking B is generated by a strongly consistent mechanism.

Proof. 1) $\Rightarrow$ 2). As we saw in (4.5.1), the 3-cycle condition is a necessary condition for stability.

2) $\Rightarrow$ 3). Let $K_1 \overline{B} X_1, K_2 \overline{B} X_2$ and $X_1 \cap X_2 = \emptyset$. One must check that $(K_1 \cup K_2)\overline{B}(X_1 \cup X_2)$. We can assume that $K_1 \cap K_2 = \emptyset$. We form $K_3 = \overline{K_1 \cup K_2}$ and $X_3 = \overline{X_1 \cap X_2}$. Then $N = K_1 \sqcup K_2 \sqcup K_3$, $A = X_1 \sqcup X_2 \sqcup X_3$. By maximality $\overline{K_1} = K_2 \sqcup K_3$ blocks $\overline{X_1} = X_2 \sqcup X_3$, analogously $K_1 \sqcup K_3$ blocks $X_1 \sqcup X_3$. But then, by the 3-cycle condition, $K_1 \sqcup K_2$ does not block $X_1 \sqcup X_2$.

3) $\Rightarrow$ 4) is established in Lemma (4.4.3).

4) $\Rightarrow$ 1) is true according to Theorem (4.3.4).

3) $\Rightarrow$ 5). Let $N = K_1 \sqcup ... \sqcup K_m$, $A = X_1 \sqcup ... \sqcup X_m$ be partitions of N and A, $m \geq 2$. Suppose that $K_j \overline{B} X_j$ for $j = 1, ..., m-1$. Then due to almost-additivity $K_1 \sqcup ... \sqcup K_{m-1}$ does not block $X_1 \sqcup ... \sqcup X_{m-1}$. K_m blocks X_m, for the blocking is maximal.

5) $\Rightarrow$ 3). Maximality is a particular case of supermaximality for $m = 2$. Almost-additivity is a particular case of supermaximality for $m = 3$. Indeed, if $K_1 \overline{B} X_1$ and $K_2 \overline{B} X_2$ then $\overline{K_1 \cup K_2}$ blocks $\overline{X_1 \cup X_2}$ whence $(K_1 \cup K_2)\overline{B}(X_1 \cup X_2)$.

6) $\Rightarrow$ 1) by Lemma (4.1.7). ■

The previous theorem enables us to state that there exist strong coalitions. For example, consider the following assertion.

(4.5.12) Corollary. *Let B be supermaximal and $\mid N \mid \leq \mid A \mid$. Then there exists at least one agent which can block some alternative.*

Actually, we can even prove a stronger assertion: *there exist a coalition K and a set $X \subset A$ such that $\mid K \mid + \mid X \mid > \mid A \mid$ and every agent in K blocks any single alternative picked in X.*

Proof. Assume the converse: that is, assume that for any coalition K we can exhibit a set X, such that $\mid X \mid \geq \mid K \mid$, and for any alternative $x \in X$, there exists an agent $i \in K$, who is unable to block x. Then by the Marriage Lemma (4.A3) there is an injective mapping $\varphi : N \to A$ such that no agent i blocks the alternative $\varphi(i)$. But this contradicts supermaximality. ■

In Maskin (1979), the assertion of this Corollary was formulated without the assumption $|N| \leq |A|$. But the assertion does not hold unless the assumption is fulfilled.

Using similar arguments, we can show that if $|A| \geq m|N|$ then some agent will be able to block sets of size m.

4.6 Veto as a Decision-making Procedure

(4.6.1) Veto-procedures. How can agents select "nice" outcomes by working out the possibilities offered by a given blocking? The idea of veto-procedures expresses the following type of group behaviour. Take a group of agents N, assume a given blocking B and a preference profile R_N. In a veto-procedure, the first task of participants, as rational agents, is to block their worse alternatives. To this end, they may join in coalitions. Indeed, by joining a coalition, agents might benefit from a more decisive blocking power than by staying alone and may be able to eliminate their "worse" alternatives. (A discarded alternative does not receive further consideration.) Suppose that at some point in time, every agent $i \in N$ has been able to eliminate (by participating to appropriate blocking coalitions) a set of alternatives M_i. Of course every agent i might want next to eliminate his worst alternative $a_i = \min R_i \mid (A \backslash M_N)$ (here $M_N = \cup_i M_i$) among the alternatives that remain. Then agents will try to form a coalition K, which would be capable of blocking the set $M'_K = \cup_{i \in K}(M_i \cup \{a_i\})$. Every agent in K is interested in blocking this set of alternatives in the following sense: the complementary set of alternatives to M'_K (if it is not empty) is then strictly preferred for every $i \in K$ to the set $M_i \cup \{a_i\}$. Suppose coalitions can be formed to block M'_K. Then take one of them and denote it for simplicity K; it discards its alternatives a_i. The procedure is iterized, starting with the set of eliminated alternatives $M'_i = M_i \cup \{a_i\}$ for $i \in K$ and $M'_i = M_i$ for $i \notin K$. An outcome of this procedure is the set of alternatives that are non discarded when it becomes impossible to form any coalitions K. An outcome, obtained through this successive elimination procedure, starting with empty sets M_i, is a *veto-outcome* or *outcome of the successive veto procedure.*

(4.6.2) Proposition. *Veto-outcomes lie in the core.*

Proof. Take a given B, a profile R_N,and a veto-outcome $C = A \backslash M_N$. Suppose that $a \in C \backslash C(B, R_N)$. Then there exists a coalition $K \subset N$ which blocks the set $L(a, R_K)$. Since $(M_i \cup \{a_i\}) \subset L(a, R_i)$, for any i, (where $a_i = \min R_i \mid (A \backslash M_N)$), then by B1 the coalition K blocks also $M'_K = \cup_{i \in K}(M_i \cup \{a_i\})$. This contradicts the definition of a veto-outcome.■

To warrant the non-emptyness of the veto-outcome set, we have to restrict both the class of blockings we consider and define a stricter elimination procedure. We consider the following veto-procedure.

(4.6.3) *Veto-procedure.* For each agent i, at each step t, define the set M_i^t to be the set of discarded at time t alternatives such that either $M_i^t = M_i^{t-1}$ or $M_i^t = M_i^{t-1} \cup \{a_i^t\}$ where $a_i^t = \min R_i \mid A \backslash M_N^{t-1}$. The following requirements are to be fulfilled in order to expand the sets M_i^{t-1} :

a) the coalition $K_t = \{i \in N,\ M_i^t \neq M_i^{t-1}\} \neq \emptyset$ and blocks the set $M_{K_t}^t = \cup_{i \in K_t} M_i^t$;

b) the coalition K_t is minimal, i.e. $K \subset K_t$, $KBM_K^t \Rightarrow K = K_t$.

If such a sequence of sets $M^t = (M_i^t)$ starting from $M_i^0 = \emptyset$ can not be pursued further after some $t = T$ is reached, then the set $C = A \backslash M_N^T$ is an outcome of the veto-procedure. (Obviously this is a veto-outcome.)

(4.6.4) Example. Let $N = \{1, 2, 3\}$, $A = \{a, b, c\}$. Take the following blocking, where $\{1, 2\}B\{a, b\}$, $\{2, 3\}B\{b, c\}$, $\{2\}B\{b\}$ and all the relations derived from B1 obtain. The profile R_N has the form

c	c	a
b	a	b
a	b	c
1	2	3

.

The veto-procedure yields here a unique sequence $M^t = (M_i^t)$: $M^1 = (\emptyset, \{b\}, \emptyset)$; $M^2 = (\{a\}, \{a, b\}, \emptyset)$; $C = \{c\} = C(B, R_N)$.

(Note here that if we do not satisfy condition b), then by starting, for example, with $\tilde{M}^1 = (\emptyset, \{b\}, \{c\})$ we shall have to pursue with $\tilde{M}^2 = (\{a\}, \{a, b\}, \{c\})$ and $C = \emptyset$).

(4.6.5) Theorem. *The veto-procedure always gives a non-empty outcome for convex blockings.*

Proof. Let B be a convex blocking, R_N be a preference profile, and let the sequences M^t, K_t, $t = 1, ..., T$, be constructed following the above described veto-procedure. We prove by induction that $(\bigcup_{\tau=1}^{t} K_\tau) B M_N^t$ for every t. Suppose that for some s, $1 < s < t$, $(\bigcup_{\tau=s}^{t} K_\tau) B (\bigcup_{\tau=s}^{t} M_{K_\tau}^{\tau})$. We check that

$$(\bigcup_{\tau=s-1}^{t} K_\tau) B (\bigcup_{\tau=s-1}^{t} M_{K_\tau}^{\tau}). \tag{4.1}$$

We denote by $S = K_{s-1} \cap (\bigcup_{\tau=s}^{t} K_\tau)$, $X = M_{K_{s-1}}^{s-1} \cap (\bigcup_{\tau=s}^{t} M_{K_\tau}^{\tau})$. If $S = \emptyset$ then the relation (4.1) is true according to the first step of induction and axiom B2. If $S = K_{s-1}$ then $X = M_{K_{s-1}}^{s-1}$ and the relation (4.1) is fulfilled again due to the first step of induction. Let $S \subset K_{s-1}$, $S \neq K_{s-1}$. Since $X \supset \bigcup_{i \in S} M_i^{s-1}$, then $S\overline{B}X$ by minimality of K_{s-1} and the axiom B1. Then the convexity of B gives the relation (4.1). For s=2 and t=T we have

$$(\bigcup_{\tau=1}^{T} K_\tau) B (\bigcup_{\tau=1}^{T} M_{K_\tau}^{\tau}).$$

By virtue of axioms B1 and B3, $M_N^T = (\bigcup_{\tau=1}^{T} M_{K_\tau}^{\tau}) \neq A$, i.e. $C = A \backslash M_N^T \neq \emptyset$.■

The convexity requirement is essential to this theorem. Indeed, take the blocking of Example (4.6.4) and remove from its definition the relation $\{2\}B\{b\}$. This blocking is already non-convex, but being weaker than the original, it remains stable. However the veto-procedure ends up in discarding A in its entirety, if it starts with $M^1 = (\emptyset, \{b\}, \{c\})$.

(4.6.6) Single-Element Outcomes. The set of veto-outcomes might contain more than one element. This is the case when the blocking is "weak". For example if $B = \emptyset$ then the set of veto-outcomes is equal to A. This is, obviously, not very convenient, because choosing among these elements can bring conflict. In order to prevent multiple outcomes, we will work with maximal blockings. Moreover, we need to modify slightly the elimination procedure described earlier . (Generally speaking, we state our elimination rules according to definite requirements on outcomes: the stricter they are, the greater the coordination required from participants (see procedure $\Pi 2$ in (5.5.2)).) We shall propose here one of the possible variants. Assume that a coalition K_t may eliminate only one alternative a^t among those being proposed $a_i^t, i \in K_t$. (One can imagine, for example, that this coalition chooses a representative, who in turn declares which alternative is to be discarded.)

(4.6.7) The Veto*-procedure is as follows: at every step $t = 1, ..., T$, a minimal coalition K_t discards the "worse" alternative a^t of one of its members; that is

a) K_t is not empty and blocks the set $M_{K_t}^t = \cup_{i \in K_t} M_i^t$, where $M_i^t = M_i^{t-1} \cup \{a^t\}$ if $a^t = \min R_i \mid A \backslash M_N^{t-1}$, and $M_i^t = M_i^{t-1}$ otherwise;

b) K_t is a minimal coalition satisfying a).

If such a sequence M^t (starting from $M_i^0 = \emptyset$) can not be pursued further than some T, then the outcome of the veto*-procedure is the set $C = A \backslash M_N^T$.

As it is proved in Proposition (4.6.2) and Theorem (4.6.5), a veto*-outcome is a subset of the core; it is non-empty for convex blockings.

(4.6.8) Theorem. *The veto*-procedure yields a single-element outcome for every preference profile if and only if the blocking is supermaximal.*

In order to prove sufficiency, we use the following lemma.

(4.6.9) Lemma. *Let B be an almost-additive blocking, and R_N be a profile. Let (M_t, K_t), $t = 1, 2, ...$ be a veto*-procedure sequence for the profile R_N. Then either $K_\tau \subset K_t$, or $K_\tau \cap K_t = \emptyset$ for any $\tau < t$.*

Proof. The proof is by induction on τ. Denote by

$$K_{\tau t} = K_\tau \cap K_t,\ \ K_{t\tau} = K_t \backslash K_\tau,\ \ M_{\tau t} = M^\tau_{K_{\tau t}},\ \ M_{t\tau} = M^t_{K_{t\tau}},\ \ \tau < t.$$

We verify the assertion for $\tau = 1$. Suppose it is not true, i.e. $K_1 \cap K_t \neq \emptyset$ and K_1 is not a subset of K_t. By minimality of the coalitions K_1 and K_t, we have $K_{1t}\overline{B}M_{1t}$, $K_{t1}\overline{B}M_{t1}$. By construction $M_{1t} \cap M_{t1} = \emptyset$ (all agents $i \notin K_1$ stop taking alternative a^1, blocked by K_1, into consideration). Now we make use of the almost-additivity of B. We get $(K_{1t} \cup K_{t1})\overline{B}(M_{1t} \cup M_{t1})$. Here $K_{1t} \cup K_{t1} = K_t$ and $M_{1t} \cup M_{t1} \subset M^t_{K_t}$. Hence, by axiom B1, $K_t\overline{B}M^t_{K_t}$ which contradicts the definition of the procedure. Thus, for $\tau = 1$, the result is true.

Now we verify the assertion for an arbitrary $\tau < t$. Suppose it is not true, i.e. $K_\tau \cap K_t \neq \emptyset$ and K_τ is not a subset of K_t. Again by the minimality of both coalitions K_τ and K_t, we get $K_{\tau t}\overline{B}M_{\tau t}$ and $K_{t\tau}\overline{B}M_{t\tau}$. We assert that $M_{\tau t} \cap M_{t\tau} = \emptyset$. Indeed, suppose that $x \in M_{\tau t} \cap M_{t\tau}$. Then $x = a^s$, for some $s < \tau$ and K_s is not a subset of K_τ, $K_s \cap K_\tau \neq \emptyset$. But this contradicts the inductive assumption. Almost-additivity gives us as previously that $K_t\overline{B}(M_{\tau t} \cup M_{t\tau})$, and consequently $K_t\overline{B}M^t_{K_t}$ which is impossible.

This lemma states that for every preference profile, the veto*-procedure yields an elimination scheme of the set $\{a^1, ..., a^T\}$. Therefore pose $K(a^t) = K_t$; by the previous Lemma, if coalition K_t intersects anyone of the following coalitions $K_1, ..., K_{t-1}$ then it contains it entirely. This feature eases the finding of coalitions K_t. In fact, the procedure can and should be pursued as long as we are left with more than one alternative. The reader can easily check this assertion making use of the almost-additivity and maximality of the blocking. We do not elaborate further; see Sotskov (1988, 1994) for a complete proof of the sufficiency part of the theorem (and also some other rules of elimination).

We proceed to proving the theorem in the opposite direction.

Necessity. Assume now that the veto*-procedure always yields single-element outcomes and the blocking is not supermaximal. Then there exist partitions $N = K_1 \sqcup ... \sqcup K_r$ and $A = X_1 \sqcup ... \sqcup X_r$ such that no K_j blocks

X_j. Consider the following profile R_N, for which the members of each coalition K_j rank the previously ordered alternatives of X_j at the "bottom" of their preferences:

$$\begin{array}{|c|c|c|c|} * & * & \cdots & * \\ X_1 & X_2 & \cdots & X_r \\ \hline K_1 & K_2 & \cdots & K_r \end{array}.$$

The veto*-procedure does not exclude the sets $X_1, ..., X_r$ inside each coalition K_j. So it does not yield one-element outcome for this profile. We have a contradiction.■

(4.6.10). Algorithmization. Finally, we consider another variation on the veto-procedure, which is better in terms of voting procedure, for it yields a quasi-automatic building of blocking coalitions. Roughly, the procedure consists in the following: Every agent (acting by turns) discards as many as possible of his "worse" alternatives among those left after the previous agents performed their discarding tasks. In order to discard his worse alternatives, he can seek the help of the previous agents. However, agent $j < i$ agrees to help i block his worse alternatives if he is warranted that his following by order alternative $a(j)$ will also be discarded. We give a formal description of this sequential veto-procedure which we denote by $\Pi 1$.

(4.6.11) The Sequential Veto-procedure. Let N be the ordered set $\{1, \ldots, n\}$. Consider a blocking B and a profile R_N. Each step of the procedure is identified with the corresponding agent's label. Agent 1 eliminates the largest lower contour of his preferences, which he able to block. Denote this contour by M_1. Let $A_1 = A \backslash M_1$, and $a(1) = \min R_1|A_1$. Agent 2 eliminates the largest lower contour M_2 in $R_2|A_1$ either by himself (say if $\{2\}BM_2$) or together with agent 1, if $a(1) \in M_2$ and $\{1,2\}B(M_1 \cup M_2)$. Let $A_2 = A_1 \backslash M_2$, and $a(2) = \min R_2|A_2$. Now let us describe what occurs at step i. By this time, we have a collection of sets M_j, of alternatives $a(j)$, $j < i$, and the following set of alternatives $A_{i-1} = A \backslash \bigcup_{j=1}^{i-1} M_j$ remaining after sequential discarding. Let M_i be the largest lower contour of agent i, for the linear order $R_i \mid A_{i-1}$, that he is able "to block" in the following sense: there is a coalition $S \subset \{1, ..., i\}$ containing i such that

$$SB(\bigcup_{j \in S} M_j), \text{ and } a(j) \in M_S := (\bigcup_{j \in S} M_j) \text{ for every } j \in S \backslash i. \tag{4.2}$$

Define now $A_i = A_{i-1} \backslash M_i$ and $a(i) = \min R_i|A_i$.

An outcome of the procedure is the set A_n.

We work out the procedure in Example (4.6.4). It yields the following sequence of pairs: $(M_i, a(i))_{i=1}^{i=3} = (\emptyset; a)$, $(\{a, b\}; c)$, $(\emptyset; c)$. The outcome is $A_3 = A \backslash M_{\{1,2,3\}} = \{c\} \in C(B, R_N)$.

(4.6.12) Theorem. *The sequential veto-procedure always has a non-empty outcome, i.e.* $A_n \neq \emptyset$*. If the blocking* B *is almost-additive, then* $A_n \subset C(B, R_N)$*. If additionally* B *is maximal, then* $A_n = \{a(n)\}$*. Conversely, if the procedure always yields single-element outcomes, then the blocking* B *is supermaximal.*

Proof. We outline the reasons for which the outcome is always non-empty. Let B be a blocking, R_N be a preference profile and $(M_i, a(i))$, $i = 1, ...n$, be a veto-sequence. Let S_i be a coalition of agents helping i, such that $S_i B M_{S_i}$ and $a(S \backslash \{i\}) \subset M_S$ (see relation (4.2). By induction and applying axiom B2, we can show that the coalition S_i can be determined as follows,

$$S_i = \{i\} \cup S(M_i). \tag{4.3}$$

Here we use the following general notation: $S(M) = \cup_k S_k$ for k such that $k < i$ and $a(k) \in M$. The reader will fill in the details. It is easy to check as well, that either $S_j \subset S_i$ or $S_j \cap S_i = \emptyset$ holds, for all i, j with $j \leq i \leq n$. From this and from axioms B2 and B3, it follows that the outcome $A_n \neq \emptyset$.

We prove that $A_n \subset C(B, R_N)$ if the blocking B is almost-additive. Let $a \in A_n \backslash C(B, R_N)$. Then there exists a coalition $K \subset N$ which blocks the set $L(a, R_K)$. Obviously, $a(i) \in L(a, R_K)$, $\forall i \in K$. Let i_1 be the agent whose label is maximal in K, and let S^1 be the maximal coalition $S \subset K$, such that $S \ni i_1$ and $a(S_{i_1}) \subset M_S$; denote by $M^1 = M_{S^1} \cup \{a(i_1)\}$. Now, let i_2 be the agent whose label is maximal in $K \backslash S^1$, and let S^2 be the maximal coalition $T \subset K \backslash S^1$, containing i_2 and such that $a(T \backslash i_2) \subset M_T$, $M^2 = M_{S^2} \cup \{a(i_2)\}$ etc. The sets M^1, $M^2, ...$ are disjoint, $\cup_j M^j \subset L(a, R_K)$ and $K = \cup_j S^j$. From this, almost-additivity and axiom B1, we conclude that $S^j B M^j$, for some j. However this contradicts the maximality of M_{i_j}.

We show that if the blocking is supermaximal then $A_n = \{a(n)\}$. To this end, we do as above for $K = N$. This yields pairwise disjoint sets M^1, $M^2, \ldots, M^r$ and a corresponding partition $N = \Pi_{j=1}^r S^j$. If $| A_n | > 1$ then $r > 1$ (otherwise by the maximality of B, $a(n)$ could be discarded). But if $r > 1$, then by supermaximality of B, there would exist some j such that $S^j B M^j$, which is impossible.

The necessity part of the argument is analogous to that appearing in Theorem (4.6.8). ■

The inductive formula (4.3) for the "team" formed by agent i and his helpers can be incorporated into the procedure, and conveniently solves the issue of organizing blocking coalitions for this veto-procedure. Since no outcome of the procedure is empty whatever the blocking considered (even if unstable), this procedure can be thought of as providing a "universal" social decision. Theorem (4.6.12) states some good properties of this social decision provided the blocking has "nice" properties as well.

4.A1 Balanced Blockings

(4.A1.1) In addition to the class of convex blockings, the class of balanced blockings constitutes another case of stable blockings. The proof of the non-emptyness of cores, in this case, is based on a nontrivial theorem on balanced games by Scarf (1967) (which is in fact equivalent to the Brouwer fixed point theorem, see Danilov and Sotskov (1987b)).

Let us introduce a few useful notions. A bundle of coalitions $\Xi = \{K\}$ forms *a balanced covering,* if there exist real numbers $\lambda_K \geq 0, K \in \Xi$, such that for any $i \in N$, $\Sigma_{i \in K} \lambda_K = 1$. It is as if every agent i would participate in a coalition $K \ni i$ with intensity (or share) λ_K; as a result, the coalition "functions" with intensity λ_K.

The following *Scarf condition* (SC_Ξ) on a blocking B can be imposed on a balanced covering Ξ.

(SC_Ξ) Take an arbitrary family of sets $X(i) \subset A$, $i \in N$. If every coalition $K \in \Xi$ blocks the set $X(K) = \cup_{i \in K} X(i)$, then N blocks $X(N) = \cup_{i \in N} X(i)$.

We can think of this condition to be a generalization of the super-additivity axiom B2. The sets $X(i)$ may be assumed disjoint for different i.

(4.A1.2) Definition. A blocking B is *balanced* if the Scarf condition (SC_Ξ) is fulfilled for every balanced covering Ξ.

For example, additive blockings are balanced. In general, note that balancedness is closely related to additivity (see Danilov and Sotskov(1987a)).

(4.A1.3) Theorem. *Balanced blockings are stable.*

Proof. We establish the point using games without side payments. Let $R_N = (R_i)_{i \in N}$ be an arbitrary preference profile. A preference R_i is then represented by a utility function $u_i : A \to \mathbf{R}$. Given a coalition K and a set $X \subset A$, we denote by $u_K(X)$ the vector in $\mathbf{R}^K$ whose $i-$th coordinate is $\min_{x \in X} u_i(x)$.

Given a blocking B, we form the game without side payments $G(B, u_N) = V = (V_K)_{K \subset N}$. Here the set $V_K \subset \mathbf{R}^K$ is the set of vectors, less or equal to $u_K(X)$ for some X, such that K enforces X.

We claim that the resulting game V is balanced in the sense of Scarf. That is, a vector $v \in \mathbf{R}^N$ belongs to V_N, as soon as its projections v_K on the spaces $\mathbf{R}^K$ belong to V_K for all coalitions K from some balanced covering Ξ. To check this, we form the following sets $X_i = \{x \in A, u_i(x) \geq v_i\}$. Since $v_K \in V_K$, then there exists a set $Y_K \subset \cap_{i \in K} X_i$ which the coalition K enforces. By monotonicity, K blocks $A \setminus \cap_{i \in K} X_i = \cup_{i \in K} \overline{X_i}$. Applying the condition (SC_Ξ) to the family of sets $(\overline{X_i})$ we conclude that N blocks $\cup_{i \in N} \overline{X_i}$, i.e. that N enforces $\cap_{i \in N} X_i$. This means simply that $v \in V_N$, which establishes the balancedness of V .

Now by Scarf's theorem (1967) (see also Danilov, Sotskov (1987b)), the core of the game V is not empty; it contains some vector v. Let x be an alternative for which $v \leq u_N(x)$. It is clear that x belongs to $C(B, R_N)$. Indeed, if x could be rejected by some coalition K, then K would enforce the set $Y = A \backslash L(a, R_K)$. And since $u_K(x) < u_K(Y)$ then $v_K < u_K(Y) \in V_K$, which contradicts the assumption that v belongs to the core of the game V. ■

Once more, the following corollary obtains: *additive blockings are stable.*

We have encountered up to now three sufficient conditions for stability, namely convexity, balancedness and almost-additivity. In general, none of them is necessary. The three agents case constitutes an exception to this rule. Indeed, in that case, the only non-trivial balanced covering is ({1,2},{2,3},{3,1}). And the corresponding Scarf condition coincides with 3-cycle condition, which is necessary for the stability (4.5.1). Henceforth we have the following corollary.

(4.A1.4) Corollary. *A blocking with three agents is stable if and only if the 3-cycle condition is fulfilled.*

We may seek for a somewhat more refined formulation of necessary conditions of stability. Is it true, for example, that for any stable blocking B there is a convex blocking B' such that $B \subset B'$? In Moulin (1983) we find the following claim: for any stable blocking B there is a maximal and stable blocking $B' \supset B$. Unfortunately this assertion is not true as evidenced by the counter-example in Gurvich and Menshikov (1989).

(4.A1.5) Example. Suppose we have three agents and six alternatives. Alternatives are divided in two groups of equal size, for example three 'black' and three 'white' alternatives. The blocking B is structured as follows: every agent enforces any group of five alternatives, every coalition of two agents enforces any pair of alternatives of identical colour (and of course any larger set of alternatives), the coalition of three agents enforces any alternative. The reader will be easily convinced that the 3-cycle condition is fulfilled for this blocking, thus B is stable. We show that it is not possible to extend B to a maximal and stable blocking B'. First, note that it is impossible to strengthen a two-agents coalition without loosing stability. Indeed, strengthening here would mean that some two-agents coalition is able to enforce a bicolour pair; but then the 3-cycle condition and, consequently stability, would not be fulfilled.

Then, by maximality of B', every agent would have to block any bicolour pair of alternatives. By superadditivity, any pair of agents would enforce (with respect to B') any pair of alternatives, which again would contradict the stability of B'.

4.A2 Blockings with Infinite Number of Alternatives

(4.A2.1) We have worked till now with the assumption that the number of alternatives is finite. We made this choice in order to avoid obscuring the matters by dealing with secondary issues. Here, we say a few words about the modifications needed in order to handle the case of an infinite number of alternatives.

Typical examples of infinite alternatives sets are sets of tax levels, sanitary norms levels, security levels etc. One could give many other examples such as exchange, joint production, etc. All these social choice objects vary continuously. An infinite set of alternatives usually has some natural topology. Thus we have to account for this topology in our definitions.

First, preferences should be continuous, i.e., closed (complete and transitive) binary relations $R \subset A \times A$. In this set-up, the earlier assumption of linearity of preferences is both restrictive and unnatural. Besides one should assume that coalitions enforce *closed sets*, i.e. block *open* subsets of A. Denote by Ω the class of open subsets of A. Then a blocking is a relation between 2^N and Ω, which satisfies axioms B1-B3. (Sometimes we shall consider that a non-open set X might be blocked by a coalition if this coalition blocks some open set $\widetilde{X} \supset X$.) For example, if the strategy sets of a mechanism $\pi : S_N \to A$ are compact and the mapping π is continuous, then the generated blocking B_π satisfies these requirements.

As in the finite set-up, we can define the notions of convex, almost-additive and maximal blockings considering only open sets or their complements when needed. The definition of the core is unchanged; it is a closed subset in A. Additive blockings are generated by probabilistic measures μ on N and β on A by the formula: KBX if and only if $\mu(K) > \beta(X)$.

We assume also the following *continuity* axiom:

B5. Let (X_α) be a directed family of open sets in A. If KBX_α for all α then $KB(\cup_\alpha X_\alpha)$.

It is easy to check that a maximal blocking is continuous, that a continuous almost-additive blocking is convex etc. as in the finite case. Given a continuous blocking, one can construct a mechanism which generates it (see (1.5.10)). However in general such a mechanism might be discontinuous. The following example illustrates this point.

(4.A2.2) Example. Let x_0 be a fixed alternative in A; let the strategies of agents be subsets $X_i \subset A$. Let c be an arbitrary choice rule on compact subsets of A. We define the following mechanism:

$$\pi(X_1, ..., X_n) = \begin{cases} c(A \setminus \cup_i X_i) & \text{, if } \cup_i X_i \neq A, \\ x_0 & \text{otherwise.} \end{cases}$$

This mechanism is clearly discontinuous for any reasonable topology, but the generated blocking B_π (a Maskin blocking) is continuous.

What we do not know, however, is whether a continuous blocking may be generated by a continuous mechanism.

(4.A2.3) Veto-procedures should also be somewhat modified. We consider more in detail the case of the sequential procedure, and as it happens it clarifies the finite number of alternatives and weak order preferences case. At step i of the procedure (recall that steps are agents' labels) we have: a) a closed remaining set of alternatives $A_i \subset A$, and b) a family of alternative sets $M_k,\ k < i$, eliminated by agent k. If $A_i \neq \emptyset$, then the procedure is carried one step further. In this step agent i eliminates the maximal by inclusion open set $M_i \subset A_i$ among the sets $M(y) = \{x \in A_i,\ \ x \prec_i y\}$ which he "blocks" with the help of a group of agents. This means that there exists a coalition of "assistants" $S \subset \{1, ..., i-1\}$ such that

$$(\{i\} \cup S)B(M_i \cup M_S^+),$$

where $M_S^+ = \cup_{k \in S} M_k^+$, and M_k^+ is the maximal lower contour in $R_k|A_k$ (i.e. M_k^+ has the form $\{x \in A_k,\ x \preceq_k y\}$), which is not "blocked" by agent k. A new closed set $A_{i+1} = A_i \backslash M_i$ is thus determined and we add the set M_i to the list of eliminated sets. The set A_{n+1} is the outcome of the procedure (it is an empty set if the procedure ends before n).

The non-emptyness of outcome result proven earlier remains true for convex and continuous blockings. And if additionally the blocking is maximal, then the last agent n considers all alternatives from an outcome as equivalent.

In short, the infinite alternatives case brings only technical difficulties; moreover, as Abdou (1987) showed, a lot is carried over to the infinite agents case. Actually, the main idea here is that the intrinsic difficulties of social choice theory bear little on the issue of finiteness or infiniteness of the set of alternatives.

4.A3 The Harems Lemma

Let there be finite sets X and Y, a correspondence $S : X \Longrightarrow Y$, and a function $b : X \longrightarrow \mathbf{N}$. In this set-up a family of "harems" is a family $(K(x),\ x \in X)$ of pairwise disjoint subsets $K(x) \subset S(x)$ of size $b(x)$. The Harems Lemma (see Wilson (1972)) states that a harem family exists if and only if $|S(X')| \geq b(X')$ for any $X' \subset X$.

We now sketch the proof. For convenience we assume that $b(x) = 1$ for every $x \in X$. This monogame variant constitutes the Marriage lemma. The general (polygame) case can be formally constructed from this one by replicating x, $b(x)$ times.

Suppose X represents a group of spectators and Y a set of seats. For every spectator $x \in X$, let $S(x)$ be the set of seats which he may have access to. In general, these sets of seats may intersect for different spectators. The question is: can all spectators take possession of one of their feasible seats? Of course, we have to assume that every group of spectators $X' \subset X$ has access on a whole to a sufficient number seats in the sense that

$$|S(X')| \geq |X'|. \tag{4.4}$$

Let the spectators enter the hall one by one and sit in one of their feasible seats. Assume that $U \subset Y$ is the subset of already occupied seats. The next spectator x enters the hall. If there is a remaining free seat among his feasible seats, he takes it; then the next spectator enters. Suppose on the contrary, that all the feasible seats of x have already been taken by spectators $Z \subset X$, i.e. $S(x) \subset U$ and $|S(x)| = |Z|$. Then the assumption (4.4) for $X' = Z \cup \{x\}$ implies that the group of spectators Z should have a feasible seat which does not lie in $S(x)$, i.e. $S(Z) \supset S(x)$ and is not equal to it. If there is such a free feasible seat (i.e. outside of U), then the corresponding spectator releases his seat for x and takes another seat outside of U. If there is no feasible free seat, then we consider the set Z' of spectators who occupy the seats $S(Z)$. Again, by condition (4.4), $S(Z')$ is bigger than $S(Z)$. We generate in this way an increasing family of feasible sets of seats until one of these overlaps the limits of U. Then, the corresponding chain of spectators change their seats, occupying some other accessible them seat while releasing their present seat for spectator x. Spectator x sits and the next spectator enters. This goes on until all the spectators are seated.

Bibliographic Comments

The first appearance of the notions of stable outcomes and of core can be traced back to the theory of cooperative games (see Luce and Raiffa (1957)). Later these notions were transposed in the context of games without side payments. Non-emptyness of the core and balancedness of a game are closely related (Bondareva, Shapley, and Scarf). Moulin and Peleg (1982) applied these notions in the context of blockings, although Oren (1981) had already discussed a few particular cases. Moulin and Peleg (1982) proved the stability of additive blockings (Theorem (4.2.2)). Proposition (4.2.7) was established by Oren (1981) and Proposition (4.2.10) by Moulin (1983). Moulin (1981a) proposed the proportional veto concept. He showed its importance in the checking of stability in the anonymous and neutral case and additionnally exhibited the veto-function for Borda rules.

The theorem about equivalence of the stability and acyclicity (see (4.5.2)) were obtained by Keiding (1985) and independently by Danilov and Sotskov (1987a). Nakamura (1979) investigated the issue of stability of simple games (see also Moulin (1985)). Cycle conditions are discussed in Danilov and Sotskov (1987a).

The notion of balanced blockings was elaborated by Danilov and Sotskov (1987b). Balanced blockings are stable (see 4.A1). Recently Boros and Gurvich (2000) proposed an interesting construction of balanced blockings by means of perfect graphs.

The notion of almost-additive blocking was introduced by Danilov and Sotskov (1987a), though it draws its sources from Oren (1981). The notion of convex blocking, generalizing the analogous notion of convex game (see Vilkov (1977)), appeared in Peleg (1984). He proved the stability of convex blockings. The direct proof given here (4.3.4), based on the sequential elimination of agents, is proposed in Demange (1987) (see also Greenberg (1985)). Lemma (4.4.3) about stability of almost-additive blockings was proved by Danilov and Sotskov (1985, 1987a). Their proof yields not only the non-emptyness of the core, but also succeeds in effectively constructing an elimination scheme.

The notion of veto-procedure from Section 6 is innovative. It did not arise from a tabula rasa, but emerged after thoroughly reading of Demange (1987). It generalizes the previously devised elimination procedures of Peleg (1978), Moulin (1983) and Holzman (1986) (see also (5.5.2) and Sotskov (1988)).

Extensions of the notion of blocking and related notions of stability, acyclicity and convexity for infinite sets of alternatives or agents can be found in Keiding (1985), Abdou (1987), Abdou and Keiding (1991).

5. Strongly Consistent Mechanisms

In this chapter, we examine mechanisms that have strong Nash equilibria for any preference profile. We start with some examples (Section 5.1), then investigate in more detail a strongly consistent mechanism with tokens (Section 5.2). Section 5.3 addresses the issue somewhat more theoretically. In particular, we show that for any maximal blocking B, there exists a mechanism whose set of equilibrium outcomes coincides with the core of the blocking B. Further (Section 5.4) we consider direct core mechanisms, that is SCFs whose outcomes are in the core. The existence of a strongly consistent selector from the core depends on a property of the underlying blocking, namely "laminability". In Section 5, we introduce several equivalent characterizations of laminable blockings, in particular an elimination procedure for finding strong Nash equilibria. Then (Section 5.6) we provide examples of laminable blockings and in Section 7, formulate a necessary and sufficient condition for laminability in terms of the blocking relation itself. In Section 8, we tackle the case of neutral laminable blockings. The Appendix provides insights on the strong implementation issue.

5.1 Definitions and Examples

(5.1.1) In this chapter as in the previous one we pursue our examination of the cooperative aspects of social choice mechanisms. As before we assume that agents have complete freedom of communication and may join in coalitions and coordinate their actions in order to improve the outcome. The main difference with respect to Chapter 4 consists in the manner by which we shall warrant the stability of an outcome. In (4.1.1) "threats" of counteracting coalitions were essential to warrant stability. In a stable state, every attempt of a coalition K at improving its well-being is prevented by an opposing coalition threatening to apply a strategy which renders the final outcome unprofitable for K. Here we go a little further along this path and ask: does there exist an array of fixed actions of agents such that *these actions per se* constitute a threat to any deviating coalition. Thus the issue investigated in the previous chapter, namely finding a stable state is now replaced by that of finding "equilibrium" strategies. Formally we will be interested in constructing mechanisms which have strong equilibria for every preference profile. Moreover we will investigate the means by which agents might come to an equilibrium. In connection with this, we devote some time to the construction of "simple" mechanisms and procedures for finding equilibria.

We now give more precise definitions. Let $\pi : S_N = \prod_{i \in N} S_i \to A$ be a social choice mechanism, and R_N be a preference profile.

(5.1.2) Definition. A *strong equilibrium* (or strong Nash, or coalition equilibrium) in a game $G(\pi, R_N)$ is a strategy profile $s_N^* = (s_i^*)$ such that, for every coalition $K \subset N$ and every coalitional strategy $s_K \in S_K$, there exists an agent $i \in K$ for which

$$\pi(s_N^*) R_i \pi(s_K, s_{\overline{K}}^*).$$

Thus a strong equilibrium (in what follows we will often omit the qualifier "strong") has the property that *no* coalition can strictly improve the outcome for *all* its members given that *the complementary coalition does not change its strategy.* A strong equilibrium is a Nash equilibrium; it belongs to the core (see lemma (5.1.4)). Thus it is very stable. In particular, a strong equilibrium outcome is Pareto optimal in the set $\pi(S_N)$; we will incidentally deal with mechanisms for which $\pi(S_N) = A$. Thus the strong equilibrium concept is a very interesting solution concept.

The set of strong equilibria in a game $G(\pi, R_N)$ is denoted by $SE(\pi, R_N)$. Unfortunately this set might happen to be empty for some mechanisms that otherwise seem very natural.

(5.1.3) Example. *Simple majority.* Let there be three participants and three alternatives. Every agent names an alternative; an outcome is an alternative supported by a majority of agents. When all three alternatives win the same number of votes, we select the alternative named by agent 2 (though the choice of a tie-breaking rule does not play a big role here). Assume the following preference profile R_N :

x	y	z
y	z	x
z	x	y
1	2	3

There is no equilibrium in this case, obviously, and by the way the core $C(B_\pi, R_N)$ is empty. In fact, we have the following important property.

(5.1.4) Lemma. *Let* π *be a mechanism,* B_π *be the blocking generated by* π (see Chapter 1, Section 5). *Then for any profile* R_N

$$\pi(SE(\pi(R_N)) \subset C(B, R_N).$$

In other words, equilibrium outcomes belong to the core. Indeed suppose that outcome $a = \pi(s_N^*)$ is rejected by the coalition K. This means that the coalition K blocks the set $L(a, R_K)$, i.e., it has a strategy s_K such that $\pi(s_K, s_{\overline{K}}) \notin L(a, R_K)$ (for any strategy $s_{\overline{K}}$ of the complementary coalition $\overline{K}$, in particular $s_{\overline{K}}^*$). Thus the alternative $\pi(s_K, s_{\overline{K}}^*)$ is strictly preferred to $a = \pi(s_N^*)$ for all members of the coalition K and s_N^* is not an equilibrium.■

We give several more elementary mechanisms of the same type generating non-stable blockings.

(5.1.5) Example. *Decisive participant rules.* ($|N| = n \geq 3,\ |A| \geq 3$).

a) Every agent names some alternative $a_i \in A, i \in N$. If the first $n-1$ agents named the same alternative, it is chosen; if not, then the alternative a_n named by agent n is chosen:

$$\pi(a_1, ..., a_n) = \begin{cases} a_1 = ... = a_{n-1}, & \text{in case of equality} \\ a_n, & \text{otherwise.} \end{cases}$$

b) Agents $1, ..., n-1$ name the alternatives $a_1, ..., a_{n-1}$, and agent n chooses among them.

c) $\pi(R_N)$ is an arbitrary Pareto alternative for the coalition $N \backslash \{n\}$ which is preferred to $a_n = \max R_n | A$ for all members of the coalition $N \backslash \{n\}$.

d) The "kingmaker" rule where the kingmaker is agent n (Example (2.1.3)).

All these mechanisms generate a blocking B_π of a simple game with winning coalitions $\{1, ..., n-1\}$ and $\{i, n\},\ i \in N$. The Nakamura number here is equal to 2 and since the number of alternatives ≥ 3, then by Proposition (4.5.5) the blocking B_π is not stable. Therefore due to lemma (5.1.4), $SE(\pi, R_N) = \emptyset$ for the corresponding profile R_N. Of course, equilibria may exist for different profiles.

Recalling the general definition of consistency given in (1.4.3), we introduce the following notion.

(5.1.6) Definition. A mechanism π is *strongly consistent* (or *SC-mechanism*) if for any preference profile $R_N \in \mathbf{L}(A)^N$ there exists a strong equilibrium in the game $G(\pi, R_N)$, i.e. $SE(\pi, R_N) \neq \emptyset$.

A legitimate question arises here: do SC-mechanisms exist? Are some of them particularly attractive, and in general how many of them are there? Below we give several examples of such mechanisms beginning with the most simple ones.

(5.1.7) Example. *A constant mechanism.* Whatever the strategies of agents, its outcome is a fixed in advance alternative $a \in A$. This is a SC-mechanism, though a rather dull one.

(5.1.8) Example. *A dictatorial mechanism.* A "dictator" is chosen among the agents. The outcome is the alternative selected by this dictator. This is also a trivial SC-mechanism.

(5.1.9) Example. There are a lot of non-constant and non-dictatorial SC-mechanisms in the two alternatives case (see (1.3.10), and (3.2.2-4)). For that one should construct a maximal simple game (N, W). Every participant calls an alternative. The winning alternative is the alternative for which the winning coalition in W has voted. Here there is a strong equilibrium in dominant strategies.

In particular, the simple majority rule is an anonymous and neutral SC-mechanism when there is an odd number of agents. Unfortunately when $|A| > 2$, this rule loses the coalitional stability property. We need more refined rules. We now give two less elementary examples.

(5.1.10) Example. *Maskin mechanism.* We fix an alternative a. Every agent i calls an alternative x_i. If each agent voted for the same alternative, we have an outcome; otherwise we set the outcome to be a.

This is a twofold mechanism. Participants, above all, try to achieve a consensus, but if this fails, a constant mechanism (with outcome a) is triggered. This is incidentally a particular case of the more general composite mechanisms mentioned in (1.5.10).

We check that this is a SC-mechanism. Let R_N be a preference profile and $M(R_N)$ be the set of Pareto optimal alternatives which dominate a for every agent (cf. Example (1.3.4)). Let all agents unanimously vote for the alternative x from $M(R_N)$. We affirm that this is a strong equilibrium. Indeed the coalition N is not interested in changing x which is Pareto optimal. Other coalitions can only force the outcome a. However, no agent would agree with this since xR_ia.

We would like to point out that there is a slight additional difficulty. It is connected with the computation of equilibria issues. Indeed, though the existence of equilibrium is not at stake, it is not altogether clear how participants may be able to reach an equilibrium without some outside help. They are supposed to find out, somehow, the preferences of other participants in order to build the set $M(R_N)$ and to choose an element in it. This is merely a social choice problem. The Maskin mechanism (and generally SC-mechanisms) can not guide agents in their search for an equilibrium; it only can stabilize those compromises that have been already agreed upon. We still need to devise a procedure by which agents are to reach an equilibrium. This can be done through the use of veto-procedures akin to the procedure $\Pi 1$ described in (4.6.11), with the difference that we want it to lead to an equilibrium (see procedure $\Pi 2$ in (5.5.2)). (Incidently this procedure ($\Pi 1$) brings agents to an equilibrium in the case of a Maskin blocking.) The procedures we think work are as follows: they fragment the act of choice into a sequence of elementary steps. At each step, the agents actions possess a "natural" and almost automatic character. Of course this simplicity is obtained at the expense of strong requirements on the blocking. As an example of such a procedure, we consider a simplified variant of the SC-mechanism proposed by Peleg (1978).

The understanding of the logic of this mechanism will greatly facilitate the reading of the remainder of this chapter.

(5.1.11) Example. *The Peleg mechanism.* To each alternative x associate an integer $\beta(x) \geq 0$. Let $\sum_{x \in A} \beta(x) = n + 1$, where $n = |N|$ is the number of participants. An agent is asked to designate an alternative that he wishes to block. If a number of agents $\geq \beta(x)$ designates the alternative x, then it is discarded. The outcome is defined to be any alternative among those that have not been discarded (which one exactly is of no importance here because, owing to condition $\sum \beta(x) > n$, we know that some alternatives will not be discarded). Let us check that this mechanism is strongly consistent. For simplicity, we consider the particular case for which every $\beta(x) \equiv 1$ and thus $|A| = |N| + 1$. The general case will be discussed in Section 2. Let agent 1 designate his worst alternative x_1, the second agent designates the alternative x_2, his worst among the set $A \backslash \{x_1\}$, then the third agent designates his worst alternative among the set $A \backslash \{x_1, x_2\}$ and so on. Finally n alternatives $x_1, x_2, ..., x_n$ will have been designated and discarded. We denote by a the remaining alternative. This profile of messages is a strong equilibrium. Indeed, suppose that a coalition K might be able to move the outcome to, say, x_i (another alternative) . Then K has to include agent i, because his mere sending a message discarded x_i. However for agent i alternative a is better than x_i since by construction $x_1, ..., x_{i-1}$ are the only alternatives worse than x_i for him.

The behavior of agents in this procedure is both very natural and "almost dominant". Every agent designates his worst alternative among those not yet discarded by predecessors. Of course, this strategy is not dominant since it depends on the strategies of predecessors. But as we have already seen, there are not too many possible mechanisms in pure dominant strategy, thus we are constrained to relax this concept somewhat.

5.2 A Tokens Mechanism (or Veto-mechanism)

(5.2.1) We now consider mechanisms which are both more general and more interesting from a practical viewpoint than that of Example (5.1.11). Given its importance, we devote a separate section to it.

The idea underlying the Peleg mechanism can be extended to the whole class of additive maximal blockings ((4.2.1)). We construct a mechanism by starting to define two families of non-negative integers:

$$\mu = (\mu(i), \, i \in N), \; \beta = (\beta(x), \, x \in A),$$

connected by the relation

$$\mu(N) + 1 = \beta(A).$$

We will view μ and β as measures on N and A so that $\mu(K) = \sum_{i \in K} \mu(i)$ and $\beta(X)$ is defined analogously.

We start by an informal description of the mechanism. Every agent i receives $\mu(i)$ tokens, which he is asked to lay on alternatives. Alternative x is eliminated if the number of tokens piled on it is greater or equal to $\beta(x)$. An outcome is one of the not yet discarded alternatives and we are assured of its existence by virtue of the inequality $\beta(A) > \mu(N)$.

More formally: a strategy of agent i is a function $s_i : A \to \mathbf{Z}_+$ such that $s_i(A) = \sum_x s_i(x) = \mu(i)$. Suppose that each agent chose a strategy s_i. We form a set

$$\Pi(s_N) = \{a \in A,\ \sum_i s_i(a) < \beta(a)\},$$

which is non-empty for every $s_N \in S_N$. Let π be an arbitrary selector from the correspondence Π, such that $\pi(s_N) \in \Pi(s_N)$.

The structure of the blocking B_π is simple to figure out. A coalition K blocks a set X if and only if $\mu(K) \geq \beta(X)$. This blocking is maximal and additive. We easily check that it is super-maximal as well, thus stable (Theorems (4.2.2) and (4.5.11)). Indeed, we derive it now from strong consistency.

(5.2.2) Theorem. *The mechanism π is strongly consisten*

Proof. Fix a preference profile R_N. We now proceed to construct an equilibrium in a mechanical way. We start by fixing an order by which agents lay out their tokens, i.e. a mapping

$$\sigma : \{1, ..., \mu(N)\} \to N$$

such that $|\sigma^{-1}(i)| = \mu(i)$ for any $i \in N$. In other words, the process of laying out tokens is divided in $\mu(N)$ (the number of tokens) step. At the k-th step, agent $\sigma(k)$ lays out a token. The behavior of an agent here is both sincere and prudent. Namely, at every step, an agent lays his token on the worst (from the point of view of R_i) alternative among those not yet discarded in the previous steps.

More formally, we define a sequence of alternatives x_k $(1 \leq k \leq \mu(N))$ and a sequence of eliminated sets Z_k $(o \leq k \leq \mu(N))$ as follows,

$Z_0 = \{x \in A,\ \beta(x) = 0\}$,
$x_k = \min(R_{\sigma(k)} | A \backslash Z_{k-1})$,
$Z^k = \{x \in A,\ x \text{ meets } \beta(x) \text{ times in the sequence } (x_1, ..., x_k)\}$.

Clearly, Z_k is equal to either Z_{k-1} or $Z_{k-1} \cup \{x_k\}$. By the end of the procedure, the set $Z_{\mu(N)}$ consists of all alternatives, but one, which we denote by $a = a(\sigma, R_N)$.

Note that at each step k agent $\sigma(k)$ never puts his token "higher" than a, i.e. on an alternative from $L(a, R_{\sigma(k)})$. We affirm that the resulting allocation of tokens s_N^* is a strong equilibrium. Let the coalition $K(x)$, for $x \neq a$, gather those agents which put some tokens on x. The coalition $K(x)$ consists of adversaries of x. Besides $\sum_{i\in K(x)} s_i^*(x) = \beta(x)$, so the coalition $K(x)$ eliminates x using a strategy $s^*_{K(x)}$. Hence as long as $K(x)$ does not deviate from the strategy $s^*_{K(x)}$, those agents for which $x \succ a$ are not able to see x realize as an outcome. Thus s^*_N is an equilibrium. ■

One of the nice features of this mechanism as an equilibrium revealing procedure is that it requires minimal information from every agent. The only information that an agent needs to know, at every step k, is the set Z_{k-1} of alternatives rejected up to this step. He does not need to know the preferences of other agents. And in fact, it is of little importance to him to know which agent put what token where. Another nice feature of this mechanism is that the strategy "to put a token on one's worst alternative among $A\backslash Z_{k-1}$" is a very natural strategy, especially if he has no idea about the preferences of the subsequent agents.

Remark. Any allocation of tokens $s_N = (s_i)$ such that:
1) $\pi(s_N) = a$,
2) $s_i(x) > 0 \Rightarrow x \in L(a, R_i)$, and
3) no alternative is "overloaded" (i.e. $\sum_i s_i(x) \leq \beta(x)$)
is a strong equilibrium.

In the following example, we show how to find an equilibrium using this procedure.

(5.2.3) Example. Consider that $N = \{1,2,3,4\}$, $A = \{x, y, z\}$, $\mu(i) = 2$ for any i, $\beta(x) = \beta(y) = \beta(z) = 3;$ and assume that π is an arbitrary mechanism of choice on the set of non-discarded alternatives (as in Theorem (5.2.2)). Take the following preference profile R_N :

x	z	y	x
z	y	z	y
y	x	x	z
1	2	3	4

The first agent puts both his tokens on y, the second agent puts both his tokens on x, the third agent lays one token on x, which as a consequence is discarded, then he puts his second token on z, the fourth agent puts both his tokens on z, which is therefore discarded as well. y is the only remaining alternative, so $\pi(R_N) = y$. This is an equilibrium according to Theorem (5.2.2).

(5.2.4) Another nice feature of the tokens mechanism is that the agents reach a strong equilibrium through *individual* actions. Roughly speaking "every Nash equilibrium is also a strong equilibrium". Of course, this is not true to the letter. Not every Nash equilibrium is a strong equilibrium (nor is it always efficient). Indeed, suppose that every Nash equilibrium be efficient, then by Dutta's Corollary (2.6.4) the blocking is generated by a simple game. Since this blocking is stable and maximal, it is dictatorial (see (4.5.7)). In order to give a meaning to the sentence in quotes, we should exclude silly (abnormal) equilibria. It is easily done here. We shall say that strategy profile $s_N = (s_i)$ is *normal* if no alternative is overloaded (see Remark above). This requirement is very natural if tokens are taken out successively and if every agent is able to see his predecessors' tokens' lay-out before he moves. Thus the following proposition is true.

(5.2.5) Proposition. *Let s_N^* be a normal Nash equilibrium in the game $G(\pi, R_N)$. Then s_N^* is a strong equilibrium.*

Proof. Let s_N^* be a normal Nash equilibrium. Then the strategy s_i^* of agent i is such that no token is put on alternatives ranking higher than $a = \pi(s_N^*)$, i.e. all tokens piled on alternatives in $L(a, R_i)$. Normality implies that all alternatives differing from a are discarded and that $\beta(a) - 1$ tokens lay on a. Suppose that agent i had put a token on an alternative b ranking higher than a according to him. Now if he would transfer this token to a , the outcome would be b. This contradicts the assumption that a is a Nash equilibrium outcome.

The remainder of the proof goes as in Theorem (5.2.2). ■

(5.2.6) Let us return to the issue of the properties of the equilibria constructed in the proof of Theorem (5.2.2). The outcome $a = a(\sigma, R_N)$ clearly depends on the sequence of moves σ. More, one can show (see Moulin (1983)) that any outcome from the core $C(B, R_N)$ can be obtained by the choice of a suitable σ. Agents, however, are not indifferent to the choice of σ. The later an agent lays his tokens, the better; for, in effect, some other agent might have contributed to discard his worse alternatives. Thus to respect fairness, the allocation of moves of agents on the segment $\{1, ..., \mu(N)\}$ should be in some way "uniform".

Using suitable weights μ and β, one can construct "almost" anonymous and neutral SC-mechanisms. More exactly, define $n = |N|$ and $m = |A|$. If n and m are relatively prime, then there exist some natural integers μ_0 and β_0 such that $\mu_0 n + 1 = \beta_0 m$. Then pose $\mu(i) \equiv \mu_0$, $\beta(x) \equiv \beta_0$. This yields an anonymous and neutral blocking B.[1] When n and m are not relatively prime, there does not exist even an anonymous and neutral (maximal and stable) blocking. However, in this case, one can allocate the same number of tokens μ_0 to every agent, then find a solution of the equation

$$\mu_0 n + 1 = \beta_0 m + r,$$

where $0 < r < m$, and define $\beta(x)$ to be equal to β_0 or to $\beta_0 + 1$. Although alternatives will not have exactly the same weight, this imbalance is negligible (especially for a large value of μ_0). In an analogous way, we can give

[1] Nevertheless the corresponding mechanism $\pi = \pi_{\mu,\beta}$ may not be anonymous or neutral. This is due to the fact that the choice of a selector from Π is arbitrary, though this is somewhat inessential.

alternatives an equal weight and allocate almost the same number of tokens to the agents, which makes them almost equal.

5.3 Blockings Generated by SC-mechanisms

(5.3.1) We undertake now a more theoretical investigation of SC-mechanisms and their construction. Allocations of power among coalitions, that is blockings, play the major role in the theory of SC-mechanisms. The importance of the blocking has been felt in both the cases of the Peleg mechanism and the tokens mechanism seen above. These SC-mechanisms are based on maximal additive blockings. What are the properties of blockings generated by SC-mechanisms? Let us answer right away: maximality (1.5.12) and stability (4.1.8); see also Theorem (4.5.11), where numerous characterizations of stable maximal blockings are given. We give more precise statements below.

(5.3.2) Proposition. *If π is a SC-mechanism, then the generated blocking B_π is maximal and stable.*

Proof. The stability follows from the definitions and lemma (5.1.4). We check the maximality property. The reasoning is analogous to the case of two agents seen in Proposition (2.2.2). Suppose the coalition K does not block the set $X \subset A$ and suppose that $\overline{K}$ does not block the set $\overline{X}$. We consider the following preference profile R_N:

$$\begin{array}{|c|c|} \overline{X} & X \\ X & \overline{X} \\ \hline K & \overline{K} \end{array}\,.$$

Let s_N^* be a strong equilibrium in a game $G(\pi, R_N)$. Then $\pi(s_N^*)$ belongs either to X or to $\overline{X}$; suppose that $\pi(s_N^*) \in \overline{X}$. The coalition K does not block X, thus the complementary coalition $\overline{K}$ is able by strategy $s_{\overline{K}}$ to transfer the outcome to X, i.e. $\pi(s_K^*, s_{\overline{K}}) \in X$. This outcome $\pi(s_K^*, s_{\overline{K}})$ is strictly preferred to $\pi(s_N^*)$ (which belongs to $\overline{X}$) for all members of $\overline{K}$; this is in contradiction to the assumption that s_N^* is an equilibrium. And analogously for $\pi(s_N^*) \in X$. ■

The converse is not true: a mechanism can generate a maximal stable blocking, while failing to be strongly consistent (compare with Theorem (2.2.3)).

(5.3.3) Example. There are two agents. Each agent has two strategies. The mechanism $\pi : S_1 \times S_2 \to A = \{x, y, a\}$ is given by the table

	s_2	s_2'
s_1	a	a
s_1'	x	y

One can see that the first agent blocks sets $\{a\}$ and $\{x, y\}$, the second blocks $\{x\}$ and $\{y\}$. Thus the blocking is maximal. Stability is obvious. However at the profile

y	x
a	y
x	a
1	2

there is a unique Nash equilibrium (s_1, s_2) with outcome a, which is not Pareto optimal.

In order to sense the difference between strong consistency of a mechanism and stability and maximality of its associated blocking, we give the following necessary and sufficient condition ensuring that a strategy profile be a strong equilibrium.

(5.3.4) Proposition. *A strategy profile $s_N^* \in S_N$ is a strong equilibrium in a game $G(\pi, R_N)$ if and only if for every alternative x, not equal to $a = \pi(s_N^*)$, there exists a coalition $K(x) \subset N$ satisfying the two properties:*

1) *the relation aR_ix holds for every $i \in K(x)$;*

2) *$K(x)$ blocks x using the strategy $s^*_{K(x)}$.*

Proof. *Sufficiency.* Suppose the coalition K has a strategy s_K such that the outcome $\pi(s_K, s^*_{\overline{K}}) = x \succ_K a$. According to 1) $K \cap K(x) = \emptyset$, so $K(X) \subset \overline{K}$. Then by 2) $\pi(s_K, s^*_{\overline{K}}) \neq x$. We have got a contradiction.

Necessity. Define $K(x) = \{i \in N,\ aR_ix\}$ and suppose that $K(x)$ does not block x using the strategy $s^*_{K(x)}$. This means that the coalition $\overline{K(x)}$ has a strategy $s_{\overline{K(x)}}$ such that $\pi(s^*_{K(x)}, s_{\overline{K(x)}}) = x$. Now the fact that $x \succ a$ for members of the coalition $\overline{K(x)}$ contradicts the assumption that s_N^* is a strong equilibrium.■

Remark. In Proposition (5.3.4), the property 2) could be replaced by the following property

2′) for any $X \subset A$, the coalition $K(X) = \cup_{x \in X} K(x)$ blocks X using the strategy $s^*_{K(x)}$.

(5.3.5) The family of coalitions $(K(x), x \in A \backslash \{a\})$ in Proposition (5.3.4) is a supporting scheme for the equilibrium outcome a (see Lemmas (4.4.6) and (4.4.7)). There, and here as well, we note that a stable outcome a requires a supporting scheme (or a structure which enables us to suppress all other alternatives). Here this scheme is implemented by some fixed strategy profile s_N^*.

Of course, the existence of such strategies s_N^* requires that the stock of strategies at our disposal be sufficiently rich. There are standard ways to enrich the set of strategies and to increase the stability of outcomes, for instance, by considering repeated game set-ups or informational extensions.

We now say a few words on informational extensions in order to acquaint the reader with what follows. By informational extension of a mechanism we mean that participants are allowed to communicate (threats etc.) with other participants while keeping their initial strategies sets fixed. We illustrate this idea returning to Example (5.3.3). The "good" outcome y is obtained through the use of the strategy profile (s_1', s_2'). However, the second agent has an incentive to switch over to s_2, because (assuming that the first agent remains passive) he prefers the outcome x. But if agent 1 can threaten him to switch from s_1' to s_1 should he consider the move to s_2, then he will refrain from doing so. The transition to informational extensions increases the stability (see, for example, Kukushkin and Morozov (1984)). Incidentally, the following theorem uses in its proof of the idea of a "universal" informational extension for which the difference between core and strong equilibria outcomes vanishes.

(5.3.6) Theorem. *Let B be a maximal blocking. Then there exists a mechanism π such that $B = B_\pi$ and*

$$\pi(SE(\pi, R_N)) = C(B, R_N)$$

for any preference profile R_N.

For convenience we divide the proof in three steps.

Proof. a) *Construction of the mechanism.* An agent's message consists in a pair (R_N, x), where the profile $R_N \in \mathbf{L}(A)^N$ and the alternative $x \in C(B, R_N)$. In short, every agent tries to guess the group's preference profile and proposes some stable alternative as a solution to the social choice problem. Thus the set of strategies S_i of an agent i, is the graph of the core correspondence $C(B, \cdot)$.

We now explain how the outcomes form. Suppose that the agents messages are $s_i = (R_N^i, x^i)$. Those agents, whose messages coincide, join in coalitions K_j of "like-minded individuals". This yields a partition of agents $N = K_1 \sqcup ... \sqcup K_s$, where the message of members of K_j is (R_N^j, x^j). Now we form the sets

$$X_j = L(x^j, R^j_{\overline{K}_j}).$$

Since x^j belongs to the core at the profile R_N^j , the coalition $\overline{K}_j$does not block the set $X_j = L(x^j, R^j_{\overline{K}_j})$. By maximality of the blocking B, the coalition K_j blocks $\overline{X}_j$. So by virtue of the basic property of blockings, $\overline{X} \cup ... \cup \overline{X}_s \neq A$, i.e. $X_1 \cap ... \cap X_s \neq \emptyset$. We then pick an arbitrary element from $X_1 \cap ... \cap X_s$ as an outcome $\pi(s_N)$. This completes the construction of the mechanism π. Note that π is almost uniquely determined by the blocking B, if we except some arbitrariness in the choice from $X_1 \cap ... \cap X_s$.

The message (R_N, x) implicitly urges the coalition K to block the complement to the set $L(x, R_{\overline{K}})$. That is every agent is more preoccupied in punishing those individuals which do not share his views ("dissidents"), than in improving upon certain outcome for himself. The efficiency of this punishment will depend directly on the number of agents in the coalition to which he belongs. Therefore one can expect that when all agents send the same message $s^* = (R_N, a)$, where R_N is the true profile of preferences, we have an equilibrium for the game $G(\pi, R_N)$.

b) *Coincidence of the core and of the equilibrium outcomes set.* We claim that for any $a \in C(B, R_N)$ the messages $s_i^* = (R_N, a)$ form a strong equilibrium for the game $G(\pi, R_N)$.

We start by checking that $\pi(s_N^*) = a$. Since all agents send the same message, all of them are like-minded and the outcome $\pi(s_N^*)$ belongs to $L(a, R_\emptyset)$ which by definition is equal to $\{a\}$. Let us see now what would happen if some coalition K deviated from the strategy s^*. In this case, one of the "like-minded" coalitions consists of $\overline{K}$. So $\pi(s_K, s_{\overline{K}}^*) \in L(a, R_K)$. But this means that the new outcome is no better than a, for some member of K.

Thus we checked the inclusion

$$C(B, R_N) \subset \pi(SE(\pi, R_N)).$$

The reverse inclusion is true if $B = B_\pi$ (Lemma (5.1.4)). This equality will be established in the step c). The second assertion of the theorem is proved.

c) *Coincidence of the blockings B and B_π.* Due to maximality of B it suffices to check that $B \subset B_\pi$. Let K be a coalition, $X \subset A$ be a set, and KBX. We need to show that $KB_\pi X$. For that we take a profile R_N:

$$\left|\begin{array}{c|c} \overline{X} & X \\ X & \overline{X} \\ \hline K & \overline{K} \end{array}\right|.$$

At such a profile, the coalition K rejects any alternative from X so that $C(B, R_N) \subset \overline{X}$. We will check below that the core $C(B, R_N)$ is non-empty, but for the time being we pursue the line of reasoning. We take an arbitrary element $a \in C(B, R_N)$ and suppose that all members of the coalition K send the same message (R_N, a). Then the outcome will fall into the set $L(a, R_{\overline{K}})$, which belongs also to $\overline{X}$ as $a \in \overline{X}$ whatever be the messages sent by other agents. This just means that $KB_\pi X$.

Lastly we need to establish the following fact. Let the group N be divided in two sub-groups K and $\overline{K}$. Denote by R the preference common to all members of the coalition K and by ${}^tR$ the opposite preference common to all members of $\overline{K}$. This we call an *antagonistic preference profile.*

(5.3.7) Lemma. *Let there be given an antagonistic preference profile R_N and a mechanism π with maximal blocking B_π. Then there exists a unique equilibrium outcome.*

In particular, the core $C(B, R_N)$ is a singleton.

Proof. We view K and $\overline{K}$ as two agents whose strategies are $S_K, S_{\overline{K}}$ and whose preferences are respectively R, ${}^tR$. The blocking being maximal, there exists a Nash equilibrium for this game. We can easily exhibit its associated outcome a. It is such that K blocks the set $L(a,R)\backslash\{a\}$ but does not block $L(a,R)$. Let K block $L(a,R)\backslash\{a\}$ with the help of a strategy s^*_K , thus due to the maximality of B_π, coalition $\overline{K}$ blocks $L(a,^t R)\backslash\{a\}$ using a strategy $s^*_{\overline{K}}$. Clearly the strategy bundle $(s^*_K, s^*_{\overline{K}})$ is a strong equilibrium whose outcome is a. And moreover a is the unique outcome from the core $C(B,R_N)$. This completes the proof of both the lemma and the theorem. ■

(5.3.8) Corollary (Moulin, Peleg). *A blocking B is stable and maximal if and only if there exists a SC-mechanism π with $B_\pi = B$.*

(5.3.9) We now have one more characterization of maximal and stable blockings, namely they appear as blockings generated by strongly consistent mechanisms.

There is something unsatisfactory in the mechanism constructed above. In effect, in spite of its strong consistency, it has the same kind of drawback we noted in the case of the Maskin mechanism discussed in Example (5.1.10). In order to come to an equilibrium, agents have to accomplish some kind of a miracle, namely they have to guess first what the preferences of other agents are and then to agree to the choice of an element from $C(B,R_N)$. It is not clear how they can manage to do so if the best they can do is to have vague ideas about what the preferences of their partners might be.

To some extent, considering a reasonably organized social choice procedure might save the case. We note that it follows from both the definition of the mechanism and the structure of the equilibrium messages $s^*_i = (R_N, a)$, that every message of the type $s_i = (R'_N, a)$, where $L(a,R'_i) \subset L(a,R_i)$, $i \in N$, $a \in C(B,R'_N)$, is also an equilibrium message. Thus, in order to reach an equilibrium, it is crucial that agents identify the sets of alternatives which they are ready to prop down. And the agents can determine these sets, going through the sequential procedure (or procedure $\Pi 1$). These sets M_i , $a(i)$, $i \in N$, and the outcome of the procedure $A_{n+1} = \{a\}$. We denote by

$$R^*_i = (* > a > a(i) > M(i)).$$

(5.3.10) Proposition. *Let B be a maximal stable blocking. Let $N = \{1,2,...,n\}$ the set of agents and R_N be a preference profile. Let a be an outcome of the sequential procedure. Then the messages $s^*_i = (R^*_N, a)$ are equilibrium strategies and a is an equilibrium outcome for the game $G(\pi, R_N)$.*

Proof. Since a is an outcome of the procedure, $a \in C(B,R_N)$ (Theorem (4.6.12)). Since $L(a,R^*_i) \subset L(a,R_i)$ for any $i \in N$, then by monotonicity of the core correspondence ((4.1.12)) $a \in C(B,R^*_N)$. Thus (s^*_i) are messages. By definition of the mechanism π, the equality $\pi(s^*_N) = a$ is true. We check now that s^*_N are equilibrium strategies. Let some coalition K implement a strategy s'_K and transfer the outcome to a point x, while the coalition $\overline{K}$ of "like-minded" agents and opponents to those agents in K, sends the same messages (s_i^*). In this case according to the definition of the mechanism π, the outcome falls into one of the sets $X_j = M_j \cup \{a(j)\}$, $j \in K$. Hence if the coalition K deviates from the strategy s_K^*, it can not warrant improving strictly the gains to all its members. Thus s_N^* is an equilibrium strategy profile and a is an equilibrium outcome.■

In the next sections, we propose alternative constructions of SC-mechanisms, for which the process of reaching equilibrium and stabilizing outcomes is more natural than here.

5.4 Direct Core Mechanisms

(5.4.1) The universal SC-mechanism, constructed in (5.3.6), was somewhat unsatisfactory for two reasons. The first one was that it required that agents send rather intricate and ingenious messages. The second reason was that it was rather difficult to figure out how agents could really come to an equilibrium. We propose here to minimize these difficulties as much as possible.

In some sense, information on preferences is the most natural information required for decision making within a group. Therefore it is natural to investigate the properties of direct mechanisms, i.e. those mechanisms $\pi : S_N \to A$ for which the strategy sets consist in linear orders, $S_i = \mathbf{L}(A)$. However this restriction alone is too weak because it says nothing more than the set S_i consists of $|A|!$ elements. We should link in a reasonable way the outcomes $\pi(R_N)$ with the profiles R_N, and indeed account for the rich structure of the set $\mathbf{L}(A)$. This can be done in various ways, but we shall consider only one. We would have ideal mechanisms π, if any true preference profile R_N were already a strong equilibrium in the game $G(\pi, R_N)$. But, as we know from Gibbard's theorem, this is impossible with the exception of a few degenerate cases (however, see Theorem (5.4.6) below). Nevertheless, one might try to ensure that the profile R_N be an "almost equilibrium" or at least close enough to an equilibrium, in order that the mechanism π does the job of finding the equilibria. We know from Lemma (5.1.4) that equilibrium outcomes in the game $G(\pi, R_N)$ belongs to the core $C(B_\pi, R_N)$. So it seems reasonable to require that the outcome $\pi(R_N)$ belong to $C(B_\pi, R_N)$, for any profile $R_N \in \mathbf{L}^N$. Though such an alternative is not necessarily an equilibrium outcome, it is already good enough to prevent agents, to some extent, wishing to distort their preferences.

(5.4.2) Definition. A direct mechanism $\pi : \mathbf{L}^N \to A$ is called a *core mechanism* if $\pi(R_N) \in C(B_\pi, R_N)$ for any $R_N \in \mathbf{L}^N$.

In the sequel, we shall examine core mechanisms only. We shall start with a blocking B (viewed as a "pre-mechanism"), which we shall generally assume maximal and stable (see the Moulin-Peleg theorem (5.3.8)). The following question arises: is it possible to choose a selector $\pi(\cdot) \in C(B_\pi, \cdot)$ of the core correspondence of the blocking B so that the resulting direct mechanism $\pi : \mathbf{L}^N \to A$ be strongly consistent and how should one do this? We give an initial and somewhat approximate answer. In many cases, that is for many blockings, we can not find a strongly consistent selector. But, when we can find a strongly consistent selector, then all selectors will turn out to be strongly consistent as well. Thus the issue does not consist in an ingenious choice of a selector from $C(B, \cdot)$, but in the proper choice of the pre-mechanism B. The strong consistency property of the selector requires that we pick a narrower class of blockings in the class of all supermaximal blockings. We call this class, the class of laminar blockings.

Let us start with an example of blocking, whose associated and arbitrary core mechanism is not strongly consistent.

(5.4.3) Example. $N = \{1,2,3,4\}$, $A = \{x,y,z\}$. Assume that the blocking B is additive, that the weight of each agent is equal to 2, and that the weight of each alternative is equal to 3. In other words, any number of agents r, where $r < 4$, blocks any $r-1$ alternatives. The blocking B is stable and maximal.

Let now $\pi : \mathbf{L}^N \to A$ be an arbitrary selector of the core correspondence $C(B, \cdot)$. We affirm that for the following profile R_N,

$$R_N = \begin{array}{|c|c|c|c|} x & z & y & x \\ z & y & z & y \\ y & x & x & z \\ \hline 1 & 2 & 3 & 4 \end{array}$$

there is no equilibrium in the game $G(\pi, R_N)$. Indeed, let $R_N^* \in SE(\pi, R_N)$. Since $x \notin C(B, R_N)$ then $\pi(R_N^*) \neq x$. We start by checking that the condition $\min R_2^* = \min R_3^* = x$ is fulfilled for an equilibrium profile R_N^*, as it is for the original profile R_N. In fact, suppose $\min R_2^* = y$. Then the coalition $\{1,4\}$, using a strategy $R_N^{'}$:

$$R'_N = \begin{array}{|c|c|c|c|} x & * & * & x \\ y & * & * & y \\ z & y & * & z \\ \hline 1 & 2 & 3 & 4 \end{array}$$

can make the outcome equal to x. Indeed, at such a profile, alternative z is rejected by the coalition {1,4}, while alternative y is rejected by the coalition $\{1,2,4\}$, so $C(B, R'_N) = \{x\}$, and $\pi(R'_N) = x$. However, alternative x is preferred to $\pi(R_N^*) \neq x$ for the agents 1 and 4.

Thus $\min R_2^* = \min R_3^* = x$. Let $\pi(R_N^*) = y$ (the case $\pi(R_N^*) = z$ is symmetrical). In this case, the coalition {1,2} can benefit from transferring the outcome to z, using a strategy,

z	z	$*$	$*$
x	x	$*$	$*$
y	y	x	$*$
1	2	3	4

.

This contradicts the assumption of R_N^* being an equilibrium profile.

(5.4.4) Singular Profiles. In the previous example, the core $C(B, R_N)$ consisted of two alternatives y and z. And this was not just by chance. As we shall see, if at a profile R_N the core $C(B, R_N)$ consists of unique alternative, then equilibria exist. Moreover, for any core mechanism π the profile R_N itself will be an equilibrium in the game $G(\pi, R_N)$.

(5.4.5) Definition. A profile R_N is called *a-singular* if $C(B, R_N) = \{a\}$. A profile R_N is *singular* if $C(B, R_N)$ consists of a single alternative.

(5.4.6) Theorem. *Let B be a convex blocking* (see (4.3.1))*, and R_N be a singular profile. Then for any selector $\pi(\cdot) \in C(B_\pi, \cdot)$, $R_N \in SE(\pi, R_N)$.*

Proof. Let $C(B, R_N) = \{a\}$. According to Demange's theorem ((4.3.7)), for any alternative $x \neq a$, there exists a coalition of "opponents" to x, $K(x) \subset N$, such that

a) aR_ix , for all $i \in K(x)$,
b) $K(x)BL(x, R_{K(x)})$.

By definition of the mechanism π and property b) the coalition $K(x)$ blocks x by the strategy $R_{K(x)}$. We only need now to apply Proposition (5.3.4). ■

We see that the property of strong consistency of core mechanisms is closely connected with the presence of singular profiles. More exactly, we can prove the following corollary.

(5.4.7) Corollary. *Let B be a convex blocking, and a profile R_N have the following property: there exists a-singular profile R_N^* such that $R_N^* \preceq_a R_N$. Then $R_N^* \in SE(\pi, R_N)$ for any core mechanism π.*

Proof is similar to that presented above.■

(5.4.8) We will call *regular* those equilibria whose existence is established in Corollary (5.4.7). Regular equilibria have two additional nice properties. The first is that $C(B, R_N^*) = \{a\}$, which determines uniquely the value of the core mechanism π at the point R_N^*. The second is that the profile R_N^* is "similar" to the true profile R_N, in the sense that $R_N^* \preceq_a R_N$. Denote by $RE(B, R_N)$ the set of regular equilibrium outcomes of a blocking B at the preference profile R_N.

Note that the profile R_N^* can be assumed a-equivalent to the profile R_N. For that it suffices to define, for agent $i \in N$,

$$R_i^{**} = (R_i|\overline{X}_i,\ R_i^*|X_i),$$

where $X_i = L(a, R_i^*)\backslash\{a\}$. Then for $i \in K(x)$, we have $x \in X_i$, so that $L(x, R_i^{**}) = L(x, R_i^*)$ and $K(x)BL(x, R_{K(x)}^{**})$. Analogously we check that $R_N^{**} \approx_a R_N$.

This shows that a regular equilibrium (provided it exists, as we know it might not always exist, see Example (5.4.3)) is obtained from the true preference profile R_N by shuffling some alternatives lying "under" a. The aim of such shuffling is to reject all alternatives differing from a. This hints at how to construct regular equilibria. We discuss this idea more in detail in the next Section.

Since the notion of regular equilibrium does not depend on the choice of a selector in the core correspondence, but only on the blocking itself, it is justified to introduce the following core equilibrium notion.

(5.4.9) Definition. Let B be a blocking and R_N be a preference profile. We say that a profile R_N^* is a *core equilibrium* for R_N if

CE1. $C(B, R_N^*) = \{a\}$,

CE2. $C(B, (R_K^{'}, R_{\overline{K}}^*)) \subset L(a, R_K)$ for any $K \subset N$ and any $R_K' \in \mathbf{L}^K$.

The alternative a is *a core equilibrium outcome (*or *CE-outcome*) for the preference profile R_N. Let $CE(B,R_N)$ denote the set of CE-outcomes at R_N. In the sequel, we shall drop the symbol B in expressions like $C(B, .)$ for brevity.

It is easy to see that $CE(R_N) \subset C(R_N)$. Assume that $a \notin C(R_N)$ and that the coalition K blocks the set $X = L(a, R_N)$. Let R' be a preference order of the form $(\overline{X} \succ X)$. Then $C(R_K', R_K^*) \subset \overline{X}$ and hence $a \notin CE(R_N)$.

(5.4.10) Proposition. *Let B be a stable blocking. A pair (R_N^*, a) is a core equilibrium if and only if for any alternative $x \neq a$ there exists a coalition $K(x) \subset N$ such that*

a) *aR_jx for all $j \in K(x)$,*

b) *$K(x)BL(x, R_{K(x)}^*)$.*

Proof. Let (R_N^*, a) be a core equilibrium for a blocking B. Suppose that, for some alternative $x \neq a$, there exists no coalition $K(x)$, satisfying the conditions a) and b). Denote by $K = \{i \in N, x \succ_i a\}$. Propping x to the top in the preferences of agents $i \in K$, we get $x \in C(R_K', R_{\overline{K}}^*) \subset L(a, R_K)$, but this is impossible. Conversely, let the conditions a), b) be fulfilled. Then no $x \neq a$ belongs to $C(R_N^*)$ and since the blocking is stable CE1 is true. Show CE2. Let $x \notin L(a, R_K)$. Then xR_ja for all $j \in K$. So by a) $K \cap K(x) = \emptyset$, and by b) $K(x)BL(x, R_{K(x)}^*)$, therefore $x \notin C(R_K', R_{\overline{K}}^*)$ for any R_K'.■

(5.4.11) Corollary. *If $a \in CE(R_N)$ then there exists a core equilibrium (R_N^{**}, a) such that $R_N^{**} \preceq_a R_N$.*

Proof. Let (R_N^*, a) be a core equilibrium for the preference profile R_N and let $(K(x), x \neq a)$ be a family of coalitions satisfying a) and b) from Proposition (5.4.10). We form the new profile R_N^{**}, for each i, $R_i^{**} = (X_i, R_i^*|\overline{K}_i)$, where $X_i = L(a, R_i^*) \backslash L(a, R_i)$. According to Proposition (5.4.10), we should prove the condition b) for R_N^{**}. Let $x \neq a$ and $i \in K(x)$. Prove that $L(x, R_i^{**}) \subset L(x, R_i^*)$. Since $i \in K(x)$ then aR_ix and $x \in L(a, R_i) \subset \overline{X}_i$. So $L(x, R_i^{**}) \subset L(x, R_i^*)$ for any $i \in K(x)$, and $L(x, R_{K(x)}^{**}) \subset L(x, R_{K(x)}^*)$. Therefore $K(x)BL(x, R_{K(x)}^{**})$. ■

(5.4.12) Corollary. $CE(B,R_N) \subset RE(B, R_N)$.■

We establish below the converse inclusion.

As becomes clear now, the existence of regular equilibria issue is closely connected with the notions of elimination scheme ((4.4.5)) and of supporting scheme ((4.1.5)), but more precisely with some their combination. Let some blocking B be fixed.

(5.4.13) Definition. A *laminar supporting scheme* of an alternative a at a profile R_N is an elimination scheme $\mathbf{K}$ of the set $A \backslash \{a\}$ such that aR_ix for every $i \in K(x)$.

That is for any alternative $x \neq a$, there is a coalition $K(x)$ of "opponents" to x (axiom C1) summoned to "prevent" the outcome to be x (axiom C2). Given a blocking B and a profile R_N, denote by $Nu(B, R_N)$ the set of alternatives a, for which there exist laminar supporting schemes. This set generalizes Holzman's nucleus concept (1987), in which the blocking coalitions in the elimination scheme were assumed to be disjoint.

(5.4.14) Proposition. *Let B be an almost additive blocking ((4.4.2)), R_N be a profile and $a \in A$. The following assertions are equivalent:*

1) there exists an a-singular profile R_N^* such that $R_N^* \preceq_a R_N$;

2) there exists a laminar supporting scheme of alternative a at the profile R_N.

Proof. 1) $\Rightarrow$ 2). Applying the Corollary (4.4.8) to the profile R_N^* and $X = A \backslash \{a\}$, we know that there exists an elimination scheme $\mathbf{K}\colon A \backslash \{a\} \to 2^N$ such that aR_i^*x for $i \in K(x)$. Since $R_N^* \preceq_a R_N$ then aR_ix for $i \in K(x)$.

2) $\Rightarrow$ 1). Let $\mathbf{K}\colon A \backslash \{a\} \to 2^N$ be a laminar supporting scheme of a at profile R_N. We explicitly construct a regular profile R_N^*. To this end we introduce the sets

$$L_i = \{x \in A \backslash \{a\},\ i \in K(x)\}.$$

After that as R_i^* we take any linear order such that:

i) the set L_i stand at the bottom of R_i^* ;

ii) the order $R_i^*|L_i$ be such that the relation xR_i^*y implies $K(x) \supset K(y)$. Obviously such R_i^* exist. Then by i) aR_i^*x obtains for all $i \in K(x)$. By ii) we have for $i \in K(x)$

$$L(x, R_i^*) \subset \{y \in L_i,\ K(x) \subset K(y)\}.$$

One can see from here and C2 that the coalition $K(x)$ blocks the set $L(x, R_{K(x)}^*) \subset \{y \in L_i,\ K(x) \subset K(y)\}$.■

Now we can establish the following

(5.4.15) Theorem. *For any almost-additive blocking B the following equalities hold*

$$RE(B, \cdot) = CE(B, \cdot) = Nu(B, \cdot).$$

Proof. $RE(B, \cdot) = Nu(B, \cdot)$ by Proposition (5.4.14). $RE(B, \cdot) \supset CE(B, \cdot)$ by Corollary (5.4.12). Let $a \in Nu(B, R_N)$ and take the profile constructed above in Proposition (5.4.14) as R_N^*. Then the pair (R_N^*, a) is a core equilibrium for R_N, that is $a \in CE(B, R_N)$. ■

5.5 Laminable Blockings

(5.5.1) We saw that we can associate three correspondences RE, CE and Nu to every blocking. They are equal to each other when blockings are almost-additive. However they might be empty-valued for some profiles. To ensure that these correspondences be non-empty-valued, we need to restrict the class of blockings. This section is devoted to the description of one of these classes and its properties. In addition we develop a procedure, based on laminar elimination schemes, which explicitly yields regular equilibria.

Let us start by exploring the procedure in which regular equilibria and elimination schemes are constructed step by step. For example the proof of Theorem (4.4.7) used such a procedure. Three ideas underlie the notion of elimination schemes. The first is that one should seek a coalition $K(x)$ of "opponents" to x among those participants which rank x low. The second is that a coalition $K(x)$ which "rejects" x should be of minimal size. The third is that one can prop alternative x up in the rankings of other agents in order to enable agents to reject a next alternative.

(5.5.2) Elimination procedures. Procedures emerge from these ideas. We need the notion of a minimal elimination scheme. An elimination scheme $\mathbf{K} : X \to 2^N$ is called *minimal,* if for any elimination scheme $\mathbf{K}' : X \to 2^N$ such that $K'(x) \subset K(x)$, the equality $\mathbf{K}' = \mathbf{K}$ is fulfilled.

The *elimination procedure $\Pi 2$* presents a sequence of steps. At every step r both an alternative x_r and a coalition $K(x_r)$ form under the following conditions:

a) the elimination scheme $\mathbf{K}$ formed at this step is minimal for the set $\{x_1, ..., x_r\}$;

b) $x_r = \min(R_i | A \setminus \{x_1, ..., x_{r-1}\})$ for any $i \in K(x_r)$.

We say that the procedure *carries to the end* if as a result one succeeds in eliminating all alternatives from A but one. This alternative is called the *outcome* of the procedure $\Pi 2$ and is denoted by a. The procedure by construction yields a laminar supporting scheme of a at profile R_N provided it carries to the end.

We illustrate how this procedure works with an example.

(5.5.3) Example. $N = \{1,2,3,4\}$, $A = \{x,y,z,u\}$. The blocking is additive and is given by the weights $\mu(1) = ... = \mu(4) = 3$; $\beta(x) = \beta(y) = \beta(u) = 4$, $\beta(z) = 1$. Let the profile R_N be as follows:

z	u	z	z
x	z	u	y
u	y	y	x
y	x	x	u
1	2	3	4

.

At the first step, the coalition $K(x) = \{2,3\}$ eliminates x since $\mu(K(x)) = 6 \geq \beta(x) = 4$.

At the second step, the coalition $K(y) = \{1,2,3\}$ eliminates y since $\mu(K(y)) = 9 \geq \beta(x,y) = 8$.

At the third step, the coalition $K(z) = \{2\}$ eliminates z since $\mu(K(y)) = 9 \geq \beta(x,y,z) = 9$.

Therefore, the outcome of the procedure is u. Note that $C(B, R_N) = \{z,u\}$. The corresponding regular equilibrium profile has the form:

$*$	u	z	$*$
$*$	y	u	$*$
u	x	y	$*$
y	z	x	u
1	2	3	4

.

Thus if the procedure $\Pi 2$ carries to the end, for a given profile R_N, there exists both a laminar elimination scheme and a regular equilibrium R_N^*. And what if it does not? Then, as we proceed to show, there is no laminar elimination scheme and regular equilibria for some (maybe, differing slightly from R_N) preference profile. Let us introduce a preliminary definition.

(5.5.4) Definition. A blocking B is called *laminable* if for any profile $R_N \in \mathbf{L}^N$ there exists a laminar supporting scheme of some alternative.

A laminable blocking is obviously stable, since any supported alternative belongs to the core (see Lemma (4.1.6)).

(5.5.5) Theorem. *The following conditions on a blocking B are equivalent:*

1) *B is laminable, that is the correspondence $Nu(B,\cdot)$ is non-empty-valued;*

1′) *B is maximal and the correspondence $CE(B,\cdot)$ is non-empty-valued;*

1″) *B is maximal and the correspondence $RE(B,\cdot)$ is non-empty-valued;*

2) *B is maximal and any core mechanism $\pi(\cdot) \in C(B,\cdot)$ possesses regular equilibria for any preference profile R_N,*

2′) *B is maximal and some core mechanism $\pi(\cdot) \in C(B,\cdot)$ possesses regular equilibria at any preference profile R_N,*

3) *procedure $\Pi 2$ carries to the end at any profile R_N.*

Proof. If B is laminable, then B is maximal and stable by Lemma (4.1.7). Moreover it is almost-additive according to the Theorem (4.5.11). Therefore 1) implies 1′) and 1″) by Theorem (5.4.15). Conversely, if CE or RE is non-empty-valued, then B is stable, and adding maximality, almost-additive. Thus by Theorem (5.4.15), B is laminar. Therefore 1), 1′) and 1″) are equivalent.

It is clear that 1′) $\Rightarrow$ 2). Obviously 2) $\Rightarrow$ 2′). If 2′) holds, then for any profile R_N, there exists a a-singular profile $R_N^* \preceq_a R_N$, that is 1″) is satisfied.

Clearly, 3) $\Rightarrow$ 1).

It remains to establish that 1) $\Rightarrow$ 3). Let R_N be a profile, we check now that procedure $\Pi 2$ carries to the end. Assume we have succeeded in making several steps and constructed a minimal eliminating scheme $\mathbf{K} : X \to 2^N$ where the set $\overline{X} = \{a_1, ..., a_p\}$ contains more than one element, $p \geq 2$. We show that one can make a next step and eliminate one more alternative a_j. Let us form, for any alternative $a \in \overline{X}$, the coalition

$$\Pi(a) = \{i \in N, \ \min R_i|\overline{X} = a\}.$$

Next, consider the following auxiliary profile R'_N. The set $X_i = \{x \in X, \ i \in K(x)\}$, ordered as in the proof of Proposition (5.4.14), is propped at the bottom of R'_i. Then above X_i, we have the set $\overline{X}$. More exactly for $i \in \Pi(a_j)$ the order R'_i has the form $R'_i = \{*, \, a_{j-p}, ..., a_{j-1}, a_j, \, X_i)$. Briefly the profile R'_N has the form:

$$\begin{array}{|c|c|c|c|}
* & * & \cdots & * \\
a_2 & a_3 & \cdots & a_1 \\
\vdots & \vdots & & \vdots \\
a_p & a_1 & \cdots & a_{p-1} \\
a_1 & a_2 & \cdots & a_p \\
\vdots & \vdots & & \vdots \\
\hline
\Pi(a_1) & \Pi(a_2) & \cdots & \Pi(a_p)
\end{array}.$$

Some coalitions $\Pi(a_j)$ might be empty. Since the blocking B is laminable, there exists a laminar supporting scheme $\mathbf{K}'$ of some alternative a, for a profile R'_N. This alternative a does not belong to X. In effect, any alternative $x \in X$ is rejected (at profile R'_N) by coalition $K(x)$ (see the proof of Proposition (5.4.14)). So we consider a restriction $\mathbf{K}'|X$. A coalition $K(x)$, $x \in X$, by definition consists of "opponents" of x with respect to a at the profile R'_N; so $K'(x) \subset K(x)$. Hence by minimality of the scheme $\mathbf{K}$, we conclude that $\mathbf{K} = \mathbf{K}'|X$. Let now $a = a_j$. Then, for all agents $i \notin \Pi(a_j)$, alternative a_{j-1} is located above a_j, $a_{j-1}R'_i a_j$. Moreover since alternative $a = a_j$ is Pareto optimal, then $\Pi(a_{j-1}) \neq \emptyset$. From here we conclude that $K'(a_{j-1}) \subset \Pi(a_{j-1})$. Additionally $\mathbf{K}'|(X \cup \{a_{j-1}\})$ is an elimination scheme of the set $X \cup \{a_{j-1}\}$

and in fact, extends the scheme **K**. However this extension might turn out to be not minimal. To find a minimal scheme, we have to choose $K(a_{j-1}) \subset K'(a_{j-1})$ somehow minimal. Thus we have at hand a minimal elimination scheme $X \cup \{a_{j-1}\}$ which extends **K**. The fact that it is consistent with the profile R_N is readily noted from the followed inclusion $K(a_{j-1}) \subset \Pi(a_{j-1})$.

Thus we proceed one step further with this procedure. ■

The procedure $\Pi 2$ gives a constructive and efficient means both to check a blocking's laminability and to exhibit regular equilibria. We establish the laminability of several classes of blockings by means of this procedure.

Let us give some examples of laminable blockings.

(5.5.6) Peleg Blockings. We recall that these are additive blockings (see (4.2.6)) for which all participants have weight 1. Laminability of such blockings is simple to establish. Let R_N be a profile and, for $x \in A$, let

$$\Pi(x) = \{i \in N,\ \min R_i = x\}.$$

By supermaximality, $\Pi(x)B\{x\}$ for some x. Let $K \subset \Pi(x)$ be a minimal subcoalition blocking x; this means that $|K| = \beta(x)$. If now we both exclude the coalition K and the alternative x, the resulting blocking is a Peleg blocking as well. Thus the line of reasoning (and the construction of an elimination scheme) can be pursued till $|A| \geq 2$.

Note that the resulting elimination scheme satisfies the following property: $K(x) \cap K(y) = \emptyset$ if $x \neq y$, which is stronger than the laminarity axiom C1 ((4.4.5)).

(5.5.7) Holzman condition. The crux of the argument given above is that the coalition K uses all its power to block x. Formally this means that the weight of K is equal to the weight of x. Can we devise this in a more general case, that is when the blocking is not necessarily additive? Incidentally, Holzman (1986) answered this question. He introduced a condition on blockings, which we call the *Holzman condition*. Assume that K is a minimal coalition blocking x (i.e. $KB\{x\}$ and $K' \subset K$, $K'B\{x\} \Rightarrow K = K'$); then for any partitions $N = K \sqcup K' \sqcup K''$ and $A = \{x\} \sqcup X' \sqcup X''$ either $K'BX'$ or $K''BX''$.

(5.5.8) Proposition. *If a blocking B is maximal and satisfies the Holzman condition, then B is laminable.*

We give a sketch of the proof. Proceeding as above in the Peleg case, we start by finding a minimal coalition K which blocks alternative x. Then we exclude both K and x, bringing the set of agents to $N' = N \backslash K$, and the set of alternatives to $A' = A \backslash \{x\}$. A blocking B' is defined in the obvious way: $K'B'X' \Leftrightarrow K'BX'$ (where $K' \subset N' \subset N,\ X' \subset A' \subset A$). By the Holzman condition for B, B' is maximal. We affirm that B' satisfies the Holzman condition as well. Let $N = K \sqcup K' \sqcup K'' \sqcup K'''$ and $A = \{x, x'\} \sqcup X'' \sqcup X'''$ be partitions where K "minimally" blocks x, K' "minimally" blocks x'. Applying the Holzman condition to the block[2] (K, x), we conclude that either $K'''BX'''$ (and then everything is fine) or $K' \sqcup K''$ blocks $\{x''\} \sqcup X''$. In that case, $K \sqcup K'''$ does not block $\{x\} \sqcup X'''$ and by the Holzman condition for the block (K', x') we conclude that $K''BX''$.

Thus the proposition's assumptions are satisfied for the blocking B', and the proof is completed by induction.■

(5.5.9) Moulin Blockings. They are additive blockings, in which the weight of each alternative is equal to 1. Every Moulin blocking is laminable and the proof is quite simple. Let i be an agent whose weight is positive and let $x = \min R_i$. We exclude x and substract 1 from the weight of agent i. The resulting blocking is a Moulin blocking as well and we pursue the same line of reasoning. The resulting elimination scheme is such that all coalitions are singletons.

We can give a kind of "qualitative" generalization of Moulin blockings. To this end, one needs to introduce the following condition: let KBX, $K'BX'$ and $(K \sqcup K')B(X \sqcup X' \sqcup \{y\})$; then either $KB(X \cup \{y\})$ or $K'B(X' \sqcup \{y\})$. In other words, this is a kind of "additivity" condition, in the following strong sense, coalitions cannot increase their power by uniting. Danilov and Sotskov (1988) introduced a class of decomposable blockings, which contains both Peleg-Holzman blockings and Moulin's and their "qualitative" generalizations.

(5.5.10) Maskin Blockings. We encountered these blockings in Example (4.1.11) and in Example (5.1.10). They are also laminable. In order to assert it, we explicitly exhibit a laminar supporting scheme of an arbitrary stable outcome $b \in C(B, R_N)$. We distinguish two cases.

In the first case: $b = a$, where a is the alternative which is blocked only by the full coalition. Then, for $x \neq a$, we can take $K(x)$ to be formed of any single agent i, for which bR_ix. This agent i exists due to the efficiency of outcome b.

In the second case: $b \neq a$. Then $K(a) = N$, and for $x \neq a, b$ the coalition $K(x)$ is chosen as in the first case.

In the second case, the resulting elimination scheme differs from those considered in (5.5.6) and (5.5.9). Further and for the sake of brevity, we call this kind of schemes links.

(5.5.11) Links. More exactly, *a link* is a minimal elimination scheme $\mathbf{K} : X \to 2^N$ satisfying the following property: there is some distinguished alternative $x_0 \in X$ such that $|K(x_0)| > 1$ and the other coalitions $K(x) \subset K(x_0)$ consist of single agents.

[2] For brevity, we call "block" any pair $(K, X) \in B$ consisting of a coalition K and a subset of alternatives X.

In order to generalize the Peleg-Holzman and Maskin blockings, we consider blockings whose elimination schemes consist of disjoint links. These blockings are defined through the following *generalized Holzman condition*: let $N = K \sqcup K' \sqcup K''$ and $A = X \sqcup X' \sqcup X''$ be partitions and the block (K, X) be a link. Then either $K'BX'$ or $K''BX''$.

(5.5.12) Proposition. *Let a blocking B be maximal and satisfy the generalized Holzman condition. Then B is laminable and any associated minimal elimination scheme consists of disjoint links.* ■

In Proposition (5.7.1) we shall show that for neutral blockings the generalized Holzman condition is also necessary for laminability.

(5.5.13) Example of a More Complex Blocking. There are laminable blockings, whose elimination schemes have a more complicated structure, for they cannot be decomposed into disjoint links. Let there be four agents and four alternatives, $A = \{x, y, z, u\}$ and the following additive blocking with weights:

$\mu(1) = 1,\ \mu(2) = \mu(3) = \mu(4) = 7,$
$\beta(x) = \beta(y) = 10,\ \ \beta(z) = 2,\ \ \beta(u) = 1.$

This blocking is laminable. The following scheme

$$X = \{x, y\},\ K(x) = \{2, 3\},\ K(y) = \{2, 3, 4\},$$

is a minimal elimination scheme, but it is not a link.

5.6 A Necessary and Sufficient Condition of Laminability

(5.6.1) We defined the notion of a laminable blocking in (5.5.4) in terms of agents' preferences. Now we characterize the property of laminability purely in terms of blockings. To this end, we have to strengthen the supermaximality property on the one hand and weaken the generalized Holzman condition on the other hand.

Let us start by introducing the following notion.

(5.6.2) Definition. A pair $(K, X) \in B$ is called an *irreducible block* if there exist a minimal elimination scheme $K(\cdot) : X \to 2^N$ and $x_0 \in X$ such that $K = K(x_0) \supset K(x)$ for any $x \in X$. The coalition $K(x_0)$ is called the *base* of the irreducible block $K(\cdot)$.

For instance, a link is an irreducible block.

Obviously any minimal elimination scheme can be written as a disjoint union of irreducible blocks.

We consider now the following *L-condition* on a blocking B. Let (K, X) be an irreducible block, $|K| > 1$, and $N = K \sqcup K' \sqcup K''$, $A = X \sqcup X' \sqcup X''$ be partitions. Then either $K'BX'$, or $K''BX''$, or $KB(X \cup \{x'\})$ for any $x' \in X'$, or $KB(X \cup \{x''\})$ for any $x'' \in X''$.

(5.6.3) Theorem. *A blocking B is laminable if and only if it is supermaximal and satisfies the L-condition.*

Proof. *Necessity.* Let B be a laminable blocking. By the Theorem (5.5.5) (see 1) $\Rightarrow$ 2)) B is supermaximal. Suppose that the L-condition is violated. Then from supermaximality we have $K'B(X'\backslash\{x'\})$ and $K''B(X''\backslash\{x''\})$. But then one can construct a minimal elimination scheme of the set $X'\backslash\{x'\}$ with coalitions from K' (and analogously for the set $X''\backslash\{x''\}$). Together with the scheme $\mathbf{K}$ this gives a minimal elimination scheme $\widetilde{\mathbf{K}}$ of the set $A\backslash\{x', x''\}$. Now one can take the preference profile R_N as in the proof of Theorem (5.5.5) so that $\Pi(x') \supset K'$, $\Pi(x'') \supset K''$ and $\Pi(x'), \Pi(x'')$ intersect $K(x_0)$. Since the L-condition holds, the procedure $\Pi 2$ carries to the end and gives $\widetilde{K}(x') \subset \Pi(x')$ (or analogously for x''). If $\widetilde{K}(x') \subset K(x_0)$ we would get a contradiction with Lemma (4.4.6) that $K(x_0)$ blocks $X \cup \{x'\}$. Otherwise $\widetilde{K}(x') \subset K'$ and again by the same Lemma $K'BX'$, which contradicts the assumption.

Sufficiency. We show that if a blocking B is both supermaximal and satisfies the L-condition, then the procedure $\Pi 2$ carries to the end at any given profile R_N. Suppose we are at step k and we have a minimal elimination scheme $\mathbf{K}(\cdot) : X \to 2^N$, $|\overline{X}| > 1$. We show that we can go to step $k+1$ and can discard an additional alternative.

First we will get an auxiliary fact.

(5.6.4) Lemma. *Let $\mathbf{K} : X \to 2^N$ be both a minimal elimination scheme and an irreducible block with base K. Suppose K' is such that $K' \subset K$ and $K' \neq K$, then $K'\overline{B}\{x \in X,\ \mathbf{K}(x) \cap K' \neq \emptyset\}$.*

Assume the converse and take a minimal coalition $K' \neq K$, $K' \subset K$ which blocks the set $\{x \in X,\ \mathbf{K}(x) \cap K' \neq \emptyset\}$. Then coalition $\mathbf{K}(x) \cap K'$, where $x \in X$, blocks the set of alternatives y, such that $\mathbf{K}(y) \subset \mathbf{K}(x)$ and $\mathbf{K}(y) \cap K' \neq \emptyset$ (for $\mathbf{K}(x) \neq K$ this follows from both almost-additivity and minimality of coalition K'). So, one can construct a new elimination scheme $\widetilde{\mathbf{K}}$, imbedded in the scheme $\mathbf{K}$, by setting $\widetilde{\mathbf{K}}(x) = \mathbf{K}(x) \cap K'$ if $\mathbf{K}(x) \cap K' \neq \emptyset$ and $\widetilde{\mathbf{K}}(x) = \mathbf{K}(x)$ otherwise. Since (K, X) is an irreducible block, then the scheme $\widetilde{\mathbf{K}}(\cdot)$ is not equal to the scheme $\mathbf{K}(\cdot)$. This contradicts minimality of the scheme $\mathbf{K}(\cdot)$.■

Now, we come back to the proof of Theorem (5.6.3). Denote by

$$\Pi(x) = \{i \in N, \min(R_i|\overline{K} = x\}, x \in \overline{X}.$$

Let $(K_1, X_1), ..., (K_m, X_m)$ be a set of irreducible blocks of scheme $\mathbf{K}(\cdot)$. We successively consider three cases.

The first case. There exist both $x \in \overline{X}$ and an irreducible block (K_j, X_j), $j \in \{1, ..., m\}$, such that $\Pi(x) \cap K_j \neq \emptyset$ and $K_j B(X_j \cup \{x\})$. We assert that the scheme $\mathbf{K}$ can be extended to $X \cup \{x\}$ such that $\mathbf{K}(x) \subset \Pi(x)$. $\mathbf{K}(x)$ can be chosen to be equal to $\{i\}$ where $i \in \Pi(x) \cap K_j$. One only need check condition C2 (recall our definition of an elimination scheme in (4.4.5)), as obviously both condition C1 and minimality of the extended scheme are fulfilled. Let us check that the coalition $\mathbf{K}(x')$, where $x' \in X \cup \{x\}$, blocks the set $\{y \in X \cup \{x\},\ \mathbf{K}(y) \subset \mathbf{K}(x')\}$. This is true when $i \notin \mathbf{K}(x')$. Now, let $i \in \mathbf{K}(x')$. We check that $\mathbf{K}(x')$ blocks the set $\{x\} \cup \{y \in X,\ \mathbf{K}(y) \subset \mathbf{K}(x')\}$. Since $i \in K_j$, then due to C1, $\mathbf{K}(x') \subset K_j$. We know that K_j blocks $X \cup \{x\}$. Moreover by Lemma (5.6.4), $K_j \backslash \mathbf{K}(x')$ does not block the set $\{y \in X_j,\ \mathbf{K}(y) \cap (K_j \backslash \mathbf{K}(x')) \neq \emptyset\}$. So from almost-additivity, $K(x')$ blocks $\{x\} \cup \{y \in X_j,\ K(y) \subset K(x')\}$.

The second case. There exists an alternative $x \in \overline{X}$, such that the coalition $\Pi(x)$ blocks the set $\{x\} \cup [\bigcup_{K_j \subset \Pi(x)} X_j]$.

We assert, in this case, that the scheme $\mathbf{K}$ can be extended to $X \cup \{x\}$. To this end, we take a minimal coalition $\Pi \subset \Pi(x)$, which blocks the set $\{x\} \cup [\bigcup_{K_j \subset \Pi} X_j]$. If Π coincides with one of the coalitions K_j, then we are back to the first case. One can then set $\mathbf{K}(x) = \{i\}$ for any agent $i \in \Pi$. Otherwise, we pose $\mathbf{K}(x) = \Pi$. Clearly in both cases, we end up with a minimal elimination scheme of $X \cup \{x\}$.

The third case. Now suppose that neither the first nor the second case obtains, that is:

a) whatever $x \in \overline{X}$, the coalition $\Pi(x)$ does not block the set $C(x) = \{x\} \cup [\bigcup_{K_j \subset \Pi(x)} X_j]$;

b) if $K_j \cap \Pi(x) \neq \emptyset$ then K_j does not block $X_j \cup \{x\}$.

We show that this contradicts the L-condition. In order to do this we have to play around with our blocks. We consider only the blocks, which are not contained in some $\Pi(x)$, $x \in \overline{X}$, among the blocks (K_j, X_j). More precisely, we associate to every alternative $x \in \overline{X}$, both the set

$$C(x) = \{x\} \cup [\bigcup_{K_j \subset \Pi(x)} X_j]$$

and the coalition

$$\Pi_0(x) = \Pi(x) \backslash [\bigcup_{K_j \not\subset \Pi(x)} \cup K_j].$$

This done, we end up with several blocks $(K_1, X_1), ..., (K_n, X_n)$, whose K_j is contained in no $\Pi(x)$. For these the following holds,

1) $N = (\bigsqcup_{x \in \overline{X}} \Pi_0(x)) \sqcup K_1 \sqcup ... \sqcup K_n$;

2) $A = (\bigsqcup_{x \in \overline{X}} C(x)) \sqcup X_1 \sqcup ... \sqcup X_n$;

3) $\Pi_0(x) \overline{B} C(x)$ for any $x \in X$ (see a));

4) (K_j, X_j) are irreducible blocks, $1 \leq j \leq n$, $|K_j| > 1$;

5) $K_j \overline{B}(X_j \cup \{x\})$ if $K_j \cap \Pi(x) \neq \emptyset$ (see b)).

By construction, there is no $\Pi(x)$ which contains K_j. Therefore, for any given j $(1 \leq j \leq n)$, we can find two different alternatives x_j and x'_j from $\overline{X}$, such that K_j blocks neither sets $X_j \cup \{x_j\}$ or $X_j \cup \{x'_j\}$ (see property 5)). We show (below, see Corollary (5.6.6)) that there exist two alternatives x_0 and $x'_0 \in \overline{X}$ independent from j, which satisfy this previous condition. First we start by establishing the following auxiliary result. We denote

$$Y_j = \{x \in \overline{X},\ K_j \overline{B}(X_j \cup \{x\})\}$$

for every j.

(5.6.5) Lemma. *Either $Y_j \subset Y_{j'}$ or $Y_{j'} \subset Y_j$.*

Proof. Assume the converse, that is that Y_j does not belong to $Y_{j'}$ and $Y_{j'}$ does not belong to Y_j. In other words, there exist $x \in Y_j \setminus Y_{j'}$ and $x' \in Y_{j'} \setminus Y_j$. The inclusion $x \in Y_j \setminus Y_{j'}$, implies that

$$K_j \overline{B}(X_j \cup \{x\}), \quad K_{j'} B(X_{j'} \cup \{x\}),$$

while $x' \in Y_{j'} \setminus Y_j$ means that

$$K_{j'} \overline{B}(X_j \cup \{x'\}), \quad K_j B(X_j \cup \{x'\}).$$

Almost-additivity implies that

$$(K_j \cup K_{j'}) \overline{B}(X_j \cup X_{j'} \cup \{x, x'\}).$$

And superadditivity (i.e. B2) implies that

$$(K_j \cup K_{j'}) B(X_j \cup X_{j'} \cup \{x, x'\}).$$

We have a contradiction. ■

(5.6.6) Corollary. *There exist two alternatives x_0, $x'_0 \in \overline{X}$ such that for any $1 \leq j \leq n$ the coalition K_j blocks neither sets $X_j \cup \{x_0\}$ nor $X_j \cup \{x'_0\}$.*

Indeed, let Y_{j_0} be a minimal set among $Y_1, ..., Y_n$. By Lemma (5.6.5), $Y_{j_0} \subset Y_0$, for all j. Following the line of reasoning appearing after property 5) above, $|Y_{j_0}| \geq 2$.

Last part of the proof of Theorem. We are ready now to exhibit two coalitions K', K'', two sets X', X'', and two alternatives x', x'' for which the L-condition is not fulfilled. To this end, we take the alternative x_0 as in Corollary (5.6.6). From property 3), the pair $(\Pi_0(x_0), C(x_0))$ does not constitute a block. Let us now complete this pair $(\Pi_0(x_0), C(x_0))$ with elements (K_j, X_j). We stop as soon as adding any remaining block (K_k, X_k) makes the coalition $\Pi_0(x_0) \cup K_j \cup ...$ block $C(x_0) \cup X_j ...$. One can assume that

$K' = \Pi_0(x_0) \cup K_{s+1} \cup K_{s+2} \cup ... K_n$ does not block $X' = C(x_0) \cup X_{s+1} \cup ... \cup X_n$ but $K' \cup K_k$ blocks $X' \cup X_k$ for any k, $1 \leq k \leq s$.

We do the same for the pair $(\Pi_0(x'_0), C(x'_0))$ and the left over pairs $(K_1, X_1), ..., (K_s, X_s)$. We end up with a pair (K'', X''), which is not a block. Some blocks $(K_1, X_1), ..., (K_t, X_t)$ might possibly remain. And, in fact, *there remains just one block* (K_1, X_1).

Proof. Why should $t \geq 1$? If $t = 0$, we would have two partitions $N = K' \sqcup K''$ and $A = X' \sqcup X''$ and this would contradict the maximality of B.

Why should $t \leq 1$? If $t > 1$ we would have two blocks $(K' \cup K_1, X' \cup X_1)$, $(K'' \cup K_2, X'' \cup X_2)$ and may be still several blocks (K_3, X_3) and etc. This contradicts B2 and B3. Thus $t = 1$.

Now we have:

A) an irreducible block (K_1, X_1), $|K_1| > 1$ (see 4));

B) two pairs (K', X'), (K'', X'') which are not blocks and which associated to (K_1, X_1) yield a partition of (N, A);

C) two elements $x_0 \in X'$ and $x_{0'} \in X''$ such that

$$K_1 \overline{B}(X_1 \cup \{x_0\}), \;\; K_1 \overline{B}(X_1 \cup \{x'_0\}).$$

However this contradicts the L-condition and proves the sufficiency assertion. This ends the proof of the theorem. ■

There are some important and particular cases for which the L-condition simplifies and in fact, boils down to the generalized Holzman condition or Holzman condition. One of those cases is when agents are weak (see below), another one is when the blocking is neutral (see Section 5.7).

(5.6.7) Corollary. *Let a blocking B be such that all agents are weak* (see (2.4.3) . *Then the following assertions are equivalent:*

a) *B is laminable,*

b) *B is maximal and satisfies the generalized Holzman condition,*

c) *B is maximal and satisfies the Holzman condition.*

We can give an additional important result in the case of weak agents.

(5.6.8) Theorem (Holzman). *Let B be a maximal blocking with weak agents. A strongly consistent mechanism $\pi(\cdot) \in C(B, \cdot)$ exists if and only if the blocking B satisfies the Holzman condition.*

The reader will understand the proof of this result after reading that of a similar in spirit theorem about neutral blockings presented in the following Section.

5.7 Neutral Laminable Blockings

Laminability issues and existence of core SC-mechanisms issues are essentially simplified in the case of neutral blockings.

(5.7.1) Proposition. *A neutral blocking is laminable if and only if it is maximal and satisfies the generalized Holzman condition.*

Proof. By Proposition (5.5.12), it suffices to check the generalized Holzman condition. Let there be two given partitions $N = K \sqcup K' \sqcup K''$ and $A = X \sqcup X' \sqcup X''$ and let (K, X) be a link. Suppose the generalized Holzman condition does not hold, then by Theorem (5.6.3), coalition K blocks, say $X \cup \{x'\}$, where $x' \in X'$. Let $K = \mathbf{K}(x_0)$ and i be an arbitrary agent from K. Since links are minimal, the coalition $K \setminus \{i\}$ does not block the set $\{x \in X,\ \mathbf{K}(x) \neq \{i\}\}$. But then almost-additivity implies that agent i blocks the set $\{x \in X,\ \mathbf{K}(x) = \{i\}\} \cup \{x'\}$. Neutrality of the blocking means that the alternative x' can be replaced by x_0. Then $\{i\} B \{x \in X,\ \mathbf{K}(x) = \{i\}\} \cup \{x_0\}$. And this contradicts the fact that the link (K, X) is minimal. ■

(5.7.2) Thus we can decompose any minimal elimination scheme for a neutral laminable blocking into disjoint links. Proposition (5.7.1) helps describe fully neutral and anonymous blockings, which are laminable. Namely, let $n = |N|,\ m = |A|$. Then the numbers n and m should satisfy the following equality:

$$m = m_0 + kn,$$

where k is a non-negative integer and m_0 is a divisor of $n + 1$. For example,

if $n = 4$, then $m = 1, 5, 9, 13, 17, ...;$

if $n = 5$, then $m = 1, 2, 3, 6, 7, 8, 11, 12, 13, ...;$

if $n = 6$, then $m = 1, 7, 12, 19, 25, ...$

We can see from this list that laminable blockings as rare as prime numbers among integers.

We turn now to the issue of existence of core SC-mechanisms for neutral blockings. Then the theorem from (5.5.5) can be stated in a sharper form. Recall that an equilibrium profile R_N^* of a mechanism $\pi : \mathbf{L}^N \to A$ corresponding to a preference profile R_N is called regular if $R_N^* \preceq_a R_N$, where $a = \pi(R_N^*)$.

(5.7.3) Theorem. *Let B be a maximal neutral blocking. The two following assertions are equivalent:*

1) *there exists an SC-mechanism $\pi(\cdot) \in C(B, \cdot)$ which admits, for any profile of preferences, a regular equilibrium profile;*

2) *the blocking B satisfies the generalized Holzman condition.*

Proof. The implication 2) $\Longrightarrow$ 1). From Proposition (5.7.1) the blocking B is laminable. Then Theorem (5.5.5) yields 1).

We prove the implication 1) $\Longrightarrow$ 2). Let there be two partitions $N = K \sqcup K' \sqcup K''$ and $A = X \sqcup X' \sqcup X''$, in which the block (K, X) is a link. Let x_0 be a distinguished alternative in the link so that $\mathbf{K}(x_0) = K$ and let $K = S_1 \sqcup S_2$, $S_1 \neq \emptyset$, $S_2 \neq \emptyset$. We consider the profile R_N

$$R_N = \left| \begin{array}{c|c|c|c} X & * & * & X \\ & X'' & X' & \\ X'' & X' & X'' & X' \\ & \overline{x_0} & \overline{x_0} & \\ X' & \vdots Z_i & \vdots & X'' \\ \hline K' & S_1 & S_2 & K'' \end{array} \right| .$$

(The alternatives within the sets X, X', X'' are ranked arbitrarily in the member's rankings.) Let R_N^* be an equilibrium profile of the mechanism $\pi(\cdot)$, both corresponding to the preference profile R_N and "similar" to it, $\pi(R_N^*) = a$ and $R_N^* \preceq_a R_N$. Obviously, $a \notin X$. Since the profile R_N is symmetrical, one can without loss of generality assume that $a \in X'$.

There are two possible cases:

1. $L(x_0, R_K^*) = L(x_0, R_K)$.
2. $L(x_0, R_K^*) \neq L(x_0, R_K)$.

We show below that case 2 is impossible, and if case 1 obtains, then either $K'BX'$ or $K''BX''$. In order to investigate these cases, we start by establishing the following fact.

Denote by $Z_i = \{z \in X, \, K(z) = \{i\}\}$; Z_i is i-th "column" of the link (K, X). We assert that

$$\{i\}BL(z, R_i^*) \text{ for any } i \in K, \; z \in Z_i.$$

In fact, if $\{i\}\overline{B}L(z, R_i^*)$ for some $i \in K$ and $z \in Z_i$, then the remaining agents $N \backslash \{i\}$ can force the outcome to be z (and thus gain with respect to outcome a!). Therefor they use the following ranking $R' = (z \succ L(z, R_i^*) \backslash \{z\} \succ A \backslash L(z, R_i^*))$. Then the core of the resulting profile $R'_N = (R'_{N \backslash \{i\}}, R_i^*)$ consists of the single element z. So the outcome $\pi(R'_N) = z \succ_{N \backslash \{i\}} a$ and this contradicts the fact that profile R_N^* is an equilibrium profile. In particular, it follows from here that $L(x_0, R_i^*) \supset Z_i$ for any $i \in K$.

Now let us investigate case 1. The equality $L(x_0, R_K^*) = L(x_0, R_K)$ and the fact that $L(x_0, R_i^*) \supset Z_i$ for any $i \in K$ means that $L(x_0, R_i^*) \equiv L(x_0, R_i)$, $i \in K$. We show that then $K''BX''$. Suppose the converse. Then we check that the coalition $K' \sqcup S_1$ is better off when its members set alternative x_0 over X' in the profile $R_{K' \cup S_1}$:

$$R'_i = (* \succ x_0 \succ X'), \; i \in K'; \quad R'_i = (* \succ x_0 \succ X' \succ Z_i), \; i \in S_1.$$

Let us consider the profile $R'_N = (R'_{K' \sqcup S_1}, R^*_{S_2 \sqcup K''})$. We assert that

$$C(B, R'_N) \subset X''.$$

Proof the assertion. We check that all alternatives outside X'' are rejected by some coalition.

a) The set Z_i is rejected by the corresponding participant i.

b) The alternative x_0 is rejected (at the profile R'_N) by the coalition $K' \sqcup K$. Indeed, $(K' \sqcup K)B(X' \sqcup X)$ because the blocking B is maximal and K'' does not block X''.

c) Alternatives from X' are rejected by the coalition $K' \sqcup S_1$. Indeed, $S_2\overline{B}(\{x_0\} \cup (\bigsqcup_{i \in S_2} Z_i))$ because a link is minimal elimination scheme. Since we assume that K'' does not block X'' then $(K' \sqcup S_1)$ blocks $X' \sqcup (\bigsqcup_{i \in S_1} Z_i)$ by supermaximality of B.

This proves the assertion.

Since π is a core mechanism, the outcome $\pi(R'_N)$ belongs to X'' and therefore it is strictly better than a for the coalition $K' \sqcup S_1$. This contradicts the assumption that R^*_N is an equilibrium profile and ascertains that, in case 1, we have $K''BX''$.

We now examine case 2. According to what we have proved above, the ranking R^*_i of every agent $i \in K$ is such that the set Z_i and possibly some set Y_i, of external elements, are placed under x_0. Moreover $\{i\}BL(z, R^*_i)$ for any $z \in Z_i$. Since the profile R^*_N is "similar" to R_N, then $Y_i \cap X = \emptyset$. Moreover, $Y_i \cap L(z, R^*_i) = \emptyset$ because if not, by neutrality, we would have $\{i\}B(Z_i \cup \{x_0\})$. This contradicts the definition of a link. Thus the ranking R^*_i takes the form $R^*_i = (* \succ x_0 \succ Y_i \succ Z_i)$. Take then any agent $i \in K$ for which $Y_i \neq \emptyset$. The coalition $K \backslash \{i\}$ does not block the set $X \backslash Z_i$ (by minimality of a link). The coalition $K \backslash \{i\} \cup \{i\} = K$ does not block the set $X \sqcup Y_i$ (since B is neutral). Under these conditions, the coalition $K' \sqcup K''$ can force the outcome x_0, wherewith it gains with respect to outcome a. To this end, this coalition will settle for the ranking $R' = (X \succ Y_i \succ A \backslash (X \sqcup Y_i))$. At profile $R'_N = (R'_{K' \sqcup K''}, R^*_K)$, the core consists of the single point x_0, since $(K' \sqcup K'')B(A \backslash (X \sqcup Y_i))$ and $(K' \sqcup K'' \sqcup \{i\})B(A \backslash (X \backslash Z_i))$ (by maximality of B). Thus the outcome $\pi(R'_N) = x_0 \succ_{K' \sqcup K''} a$ which contradicts the fact that R^*_N is an equilibrium profile. The case 2 is therefore impossible. ■

(5.7.4) In the special case where the blocking is not neutral, but the agents are weak (see the previous section and Holzman's theorem), then in the link (K, X) the sets $Z_i = \emptyset$. In this case, the requirement of existence of an equilibrium profile "similar" to the true one and which does not allow "littering" lower contours $L(x_0, R^*_i)$ by elements of Z_j, $j \neq i$, becomes superfluous. As for the rest, a simplified but essentially similar line of reasoning proves Holzman's theorem (5.6.8).

5.A Implementation via Strong Equilibria

(5.A.1) We consider here the conditions under which a given SCC can be implemented through the strong equilibria of some mechanism. As usual, N denotes the set of participants, A denotes the set of alternatives, and $\pi : \prod_{i \in N} S_i \to A$ is a mechanism. A strategy profile $s^*_N = (s^*_i)$ is a strong equilibrium if for any coalition $K \subset N$ we have the inclusion

$$\pi(S_K, s^*_{N-K}) \subset L(\pi(s^*_N), R_K).$$

A *strong-equilibrium outcome correspondence* (SEOC) of a mechanism π is

$$Eq(\pi) = \{(R_N, a), \exists s^*_N \in S_N \text{ s.t. } \pi(S_K, s^*_{N-K}) \subset L(a, R_K)\ \forall K \subset N\}.$$

An SCC $F : \mathbf{L}^N \Longrightarrow A$ is *strongly implementable* if $F = Eq(\pi)$ for some mechanism π. The target of this section is to characterize strongly implementable SCCs.

The simplest condition for strong implementation is monotonicity (see Remark (2.3.3)). This condition is necessary but, of course, not sufficient. Below we give a few finer necessary conditions, which prove to be also sufficient.

(5.A.2) Necessary Conditions for Strong Implementation. In order to formulate our conditions, we provide a few more notations. Fix some mechanism $\pi: S_N \to A$, and let $F = Eq(\pi)$ be its SEOC.

Let (R_N, a) be a pair where R_N is a preference profile, and $a \in A$. Denote as $S^*_N(R_N, a)$ the set of all strong equilibria s^*_N in the game $G(\pi, R_N)$ such that $\pi(s^*_N) = a$. This set is non-empty if and only if $(R_N, a) \in F = Eq(\pi)$. The projection of $S^*_N(R_N, a) \subset S_N$ on S_K is denoted $S^*_K(R_N, a)$. In other words, s^*_K belongs to $S^*_K(R_N, a)$ if s^*_K can be extended to $s^*_N \in S^*_N(R_N, a)$. Of course, in this case

$$\pi(s^*_K, S_{N-K}) \subset L(a, R_{N-K}). \tag{5.1}$$

(5.A.3) Definition. Given any SCC F, we call the map $\xi_N: N \to F$ a *situation*. In other words, a situation is a family $((R^i_N, a^i))$ of elements of $F \subset \mathbf{L}^N \times A$ parametrized by participants $i \in N$. In a situation ξ_N a participant i sends the *message* $\xi_i = (R^i_N, a^i)$.

Let ξ_N be a situation for the correspondence $F = Eq(\pi)$. Assume we group the "like-minded" participants, that is those who send the same message. This forms a partition $N = K_1 \amalg ... \amalg K_m$; all the members of a coalition K_j send the message (R^j_N, a^j). Define $S^*(\xi_N)$ as the following product

$$S^*(\xi_N) = S^*_{K_1}(R^1_N., a^1) \times \ldots \times S^*_{K_m}(R^m_N, a^m) \subset S_N.$$

Finally, set $\varepsilon_\pi(\xi_N) = \pi(S^*(\xi_N)) \subset A$. Since each pair (R^j_N, a^j) belongs to $Eq(\pi)$ the set $S^*(\xi_N)$ is non-empty as well as $\varepsilon_\pi(\xi_N)$. Moreover, we see from (5.1) that $\varepsilon_\pi(\xi_N) \subset L(a^j, R^j_{N-K_j})$ for any $j = 1, ..., m$. If we set

$$\lambda(\xi_N) = L(a^1, R^1_{N-K_1}) \cap ... \cap L(a^m, R^m_{N-K_m}),$$

then $\varepsilon_\pi(\xi_N) \subset \lambda(\xi_N)$ for any situation ξ_N. Note that the definition of the correspondence $\lambda : F^N \Longrightarrow A$ does not call upon the mechanism π.

We are interested in the sets $\varepsilon_\pi(\xi_N)$ for the following reason. Let a be an element of $\varepsilon_\pi(\xi_N)$. Then a is an equilibrium outcome for any preference profile in which a is situated "high enough". We give a more precise assertion just below.

(5.A.4) Proposition. *Let R_N be a preference profile, ξ_N be a situation, and $a \in \varepsilon_\pi(\xi_N)$. Suppose that for any non-empty coalition K the inclusion $\varepsilon_\pi(*_K, \xi_{N-K}) \subset L(a, R_K)$ holds. Then $(R_N, a) \in Eq(\pi)$.*

Proof. By definition, $a \in \varepsilon_\pi(\xi_N)$ means that $a = \pi(s^*_N)$ where $s^*_N = (s^1_{K_1}, ..., s^m_{K_m})$, and $s^j_{K_j} \in S^*_{K_j}(R^j_N, a^j)$ for $j = 1, ..., m$. We show that s^*_N is a strong equilibrium in the game $G(\pi, R_N)$. Suppose that some coalition K deviates from this strategy and indeed proposes to use another strategy s_K. Extend s_K to $s_N \in S_N$. Obviously, the strategy profile s_N is a strong equilibrium at some preference profile Q_N (for example, when the alternative $\pi(s_N)$ is the best one for each participant). Therefore $\pi(s_K, s^*_{N-K}) \in \varepsilon_\pi(\vartheta_K, \xi_{N-K})$ where $\vartheta_i = (Q_N, \pi(s_N))$ for $i \in K$. Due to the assumption underlying Proposition (5.A.4), $\varepsilon_\pi(\vartheta_K, \xi_{N-K}) \in L(a, R_K)$. Hence $\pi(s_K, s^*_{N-K}) \in L(\pi(s^*_N), R_N)$ for any non-empty coalition K, and s^*_N is a strong equilibrium.■

(5.A.5) Let us sum up omitting any reference to the mechanism π. Suppose that SCC is strongly implementable. Then, for any situation $\xi_N \in F^N$, there exists a set $\varepsilon(\xi_N) \subset A$ (i.e. there exists a correspondence $\varepsilon: F^N \Longrightarrow A$) such that the three following properties are satisfied:

1) $\varepsilon(\xi_N)$ is non-empty for any situation ξ_N;

2) $\varepsilon(\xi_N) \subset \lambda(\xi_N)$ for any situation ξ_N;

3) if $a \in \varepsilon(\xi_N)$, and R_N is preference profile such that $\varepsilon(\xi'_K, \xi_{N-K}) \subset L(a, R_N)$ for every non-empty coalition K and every ξ'_K, then $(R_N, a) \in F$.

(5.A.6) Corollary. *Assume that SCC F be strongly implementable. Then for any partition $N = K_1 \amalg ... \amalg K_m$ and a family $(R^1_N, a^1), ..., (R^m_N, a^m)$ of pairs from F, the intersection $L(a^1, R^1_{N-K_1}) \cap ... \cap L(a^m, R^m_{N-K_m})$ is non-empty.*

This is a formal consequence of the properties 1) and 2). ■

Monotonicity is a formal consequence of the properties 1)-3) as well. Indeed, let $(R_N, a) \in F$, and R'_N be another preference profile such that $L(a, R'_i) \supset L(a, R_i)$, for any $i \in N$. Form the message $\xi = (R_N, a)$ and consider a situation $\xi_N = (\xi, ..., \xi)$. Then $\lambda(\xi_N) = L(a, R_N) = \{a\}$; because of 1) and 2), $\varepsilon(\xi_N) = \{a\}$. Further, $\lambda(\xi'_K, \xi_{N-K}) \subset L(a, R_K)$, hence (see the property 2)) $\varepsilon(\xi'_K, \xi_{N-K}) \subset \lambda(\xi'_K, \xi_{N-K}) \subset L(a, R_K) \subset L(a, R'_K)$ for any coalition K and ξ'_K. From 3) we conclude that $(R'_N, a) \in F$. This proves the monotonicity of correspondence F.■

In some sense, the necessary properties $1) - 3)$ are a far-reaching generalization of monotonicity. We assert that these properties are also sufficient for strong implementation. More exactly, let F be a SCC such that there exists a correspondence $\varepsilon : F^N \Longrightarrow A$, with the properties 1) − 3). Then F is strongly implemented by some mechanism π_ε, which we construct below.

(5.A.7) Construction of an Implementation Mechanism. Suppose F is an SCC, and suppose that $\varepsilon : F^N \Longrightarrow A$ be a correspondence fulfilling property 1) (that is $\varepsilon(\xi_N)$ is non-empty for every situation ξ_N). The mechanism π_ε is set-up as a composite mechanism (see (1.5.10)). The basic strategies of this mechanism consist in pairs $(R_N, a) \in F$. Suppose that participants send messages ξ_i; then we have a situation $\xi_N = (\xi_i, i \in N)$. Given this situation, a final outcome of the mechanism π_ε will be defined using a roulette with values in the set $\varepsilon(\xi_N)$. Note that the set $\varepsilon(\xi_N)$ is not empty by property 1).

The SCC implemented by the mechanism π_ε is described in the following proposition.

(5.A.8) Proposition. *A pair (R_N, a) belongs to the correspondence $Eq(\pi_\varepsilon)$ if and only if there exists a situation $\xi_N \in F$ such that $a \in \varepsilon(\xi_N)$ and inclusions $\varepsilon(\xi'_K, \xi_{N-K}) \subset L(a, R_K)$ hold for any non-empty coalition $K \subset N$ and every ξ'_K.*

Proof. In one direction, the assertion is almost obvious: if a is strong equilibrium outcome in the game $G(\pi, R_N)$, then it suffices to take ξ_N to be the "basic part" of the equilibrium strategy. Since the coalition K can force any element from $\varepsilon(\xi'_K, \xi_{N-K})$ to be an outcome, we have the inclusion $\varepsilon(\xi'_K, \xi_{N-K}) \subset L(a, R_K)$. The converse implication is simpler: suppose we have a situation ξ_N as in Proposition (5.A.8), then it suffices to complete ξ_N adding a "roulette" strategy, which yields a as an outcome (recall that $a \in \varepsilon(\xi_N)$). Then we have a strong equilibrium since the following inclusions hold, $\varepsilon(\xi'_K, \xi_{N-K}) \subset L(a, R_K)$.■

(5.A.9) Corollary. *Suppose a correspondence F has the property* 2) *then $F \subset Eq(\pi_\varepsilon)$.*

Proof. Let $(R_N, a) \in F$. We show that $(R_N, a) \in Eq(\pi_\varepsilon)$ using Proposition (5.A.8). We choose as ξ_N, the situation $(\xi, ..., \xi)$, where $\xi = (R_N, a)$. Due to the property 2), $\varepsilon(\xi_N) \subset \lambda(\xi_N) = L(a, R_\emptyset) = \{a\}$ which jointly with the property 1) yields $\varepsilon(\xi_N) = \{a\}$.

Let now K be a non-empty coalition. We prove that $\varepsilon(\xi'_K, \xi_{N-K}) \subset L(a, R_K)$. Let $\xi'_N = \varepsilon(\xi'_K, \xi_{N-K})$ be a new situation; let its associated partition have the form $K_1 \amalg ...$, where $K \supset N - K_1$ and $(R^1_N, a^1) = (R_N, a)$. Therefore $\lambda(\xi'_N) \subset L(a, R_K)$, and because of the property 2) $\varepsilon(\xi'_N) \subset \lambda(\xi'_N) \subset L(a, R_K)$.■

(5.A.10) Corollary. *Suppose a correspondence F has the property 3). Then $F \supset Eq(\pi_\varepsilon)$.*

Proof. Let $(R_N, a) \in Eq(\pi_\varepsilon)$ and a situation ξ_N be as in Proposition (5.A.8). Then property 3) states that $(R_N, a) \in F$. ■

We sum all this up in the following assertion.

(5.A.11) Theorem. *An SCC F is strongly implementable if and only if there exists a correspondence $\varepsilon : F^N \Longrightarrow A$ satisfying the properties* 1)-3). ■

(5.A.12) Modification of the Criterion. The criterion of strong implementation, which we obtained above, is not satisfactory. In fact, it says nothing about whether there exists a correspondence ε at all. We try here to provide some more satisfactory answer to this issue.

Let $F : \mathbf{L}^N \Longrightarrow A$ be an SCC. Consider the set of all correspondences $\varepsilon : F^N \Longrightarrow A$ satisfying properties 2) and 3). This set is not empty: for example, the empty correspondence satisfies these conditions. Any union of two (or more) correspondences in this set satisfies these conditions as well. Therefore there exists a correspondence $\varepsilon_F : F^N \Longrightarrow A$ satisfying the properties 2) and 3) and which is maximal (by inclusion). Of course, in general, the condition 1) can be violated.

(5.A.13) Theorem. *An SCC F is strongly implementable if and only if the correspondence ε_F satisfies the condition* 1) (that is, for any situation $\xi_N \in F^N$, the set $\varepsilon_F(\xi_N)$ is nonempty).

Proof. If ε_F satisfies 1) then by Theorem (5.A.11), F is strongly implementable. Conversely, if F is strongly implementable then by Theorem (5.A.11), there exists a correspondence ε satisfying the conditions $1) - 3)$. But then $\varepsilon_F \supset \varepsilon$ also satisfies 1).■

(5.A.14) The criterion given in Theorem (5.A.13) improves on the previous one. However, we still do not know what the correspondence ε_F should look like or how it could be constructed. We provide below a constructive way of generating this correspondence. Namely, we proceed inductively. We form a sequence of correspondences $\varepsilon_0 \supset \varepsilon_1 \supset \ldots$, whose limit $\varepsilon_\infty = \cap_k \varepsilon_k$ is equal to ε_F.

We start this sequence by setting $\varepsilon_0 = \lambda$, where the correspondence ε_0 tautologically fulfills condition 2). However, in general, it does not satisfy condition 3). In order to come closer to fulfilling condition 3), we construct a new correspondence ε_1 as follows:

$$\varepsilon_1(\xi_N) = \{a \in \varepsilon_0(\xi_N) \text{ such that } a \in F(R_N) \text{ as soon as}$$
$$L(a, R_K) \supset \varepsilon_0(\xi'_K, \xi_{N-K}) \text{ for any non-empty coalition } K \text{ and every } \xi'_K\}.$$

Obviously $\varepsilon_1 \subset \varepsilon_0$, satisfies 2); but in general ε_1 does not satisfy 3). Thus we repeat the operation and form ε_2 and so on. Suppose we have constructed a correspondence ε_k, we form the correspondence ε_{k+1} where we take as situation ξ_N

$$\varepsilon_{k+1}(\xi_N) = \{a \in \varepsilon_k(\xi_N), a \in F(R_N) \text{ for a preference profile } R_N \text{ such that}$$
$$L(a,R_K) \supset \varepsilon_k(\xi'_K, \xi_{N-K}) \text{ for any non-empty coalition } K \text{ and every } \xi'_K\}.$$

So we are left with a decreasing sequence of correspondences, $\lambda = \varepsilon_0 \supset \varepsilon_1 \supset ... \supset \varepsilon_k \supset ...$, and define $\varepsilon_\infty = \cap_k \varepsilon_k$. Note that this sequence becomes stationary after some step (because both A and N are finite); therefore $\varepsilon_k = \varepsilon_{k+1} = ... = \varepsilon_\infty$ and ε_∞ satisfies the conditions 2) and 3).

(5.A.15) Proposition. *For any SCC F, the following equality holds:* $\varepsilon_F = \varepsilon_\infty$.

Proof. Since the correspondence ε_∞ satisfies 2) and 3), then $\varepsilon_\infty \subset \varepsilon_F$. The reverse inclusion obtains if we show that $\varepsilon_F \subset \varepsilon_k$ for any $k \geq 0$. It is obvious that $\varepsilon_F \subset \varepsilon_0 = \lambda$. Suppose that the assertion is true, for some k. Let us prove it for $k+1$.

Let $a \in \varepsilon_F(\xi_N)$; we need to show that $a \in \varepsilon_{k+1}(\xi_N)$. Due to the definition of ε_{k+1}, we need to prove that $a \in F(R_N)$ for any preference profile R_N such that $L(a, R_K) \supset \varepsilon_k(\xi'_K, \xi_{N-K})$, where K is a non-empty coalition and every ξ'_K. The inductive proposition implies that $\varepsilon_k \supset \varepsilon_F$; therefore if $L(a, R_K) \supset \varepsilon_k(\xi'_K, \xi_{N-K})$ then $L(a, R_K) \supset \varepsilon_F(*_K, \xi_{N-K})$. Since the condition 3) is true for ε_F, then $a \in F(R_N)$.■

Thus we are left with yet another form of strong implementation criterion.

(5.A.16) Theorem. *An SCC F is strongly implementable if and only if the correspondences* ε_k $(k = 0, 1, ...)$ *are non-empty-valued.* ■

So, we have the series of conditions. The most important condition is that $\varepsilon_0 = \lambda$ be non-empty-valued. In essence, this is the assertion of Corollary (5.A.6). If $\lambda(\xi_N) = \emptyset$ for some a situation ξ_N then the SCC F can not be strongly implemented. If, on the contrary, $\lambda(\xi_N) \neq \emptyset$ for any situation ξ_N, then we form the next correspondence ε_1, and so on. The process ends either when ε_k has an empty value (then F is not strongly implementable) or when $\varepsilon_{k+1} = \varepsilon_k$ (then F is strongly implementable).

For an illustration, we suggest that the reader return back to the Theorem (5.3.6). It asserts that if a blocking B is maximal then the core correspondence $C(B, \cdot)$ is strongly implementable. There we explicitly constructed an implementing mechanism (by the way, its construction is similar to that of π_ε and is based on non-empty-valuedness of λ). But one can choose another way, that is to prove that the correspondence λ satisfies condition 3) (the condition 1) follows from supermaximality of B).

Bibliographic Comments

The idea of strongly consistent mechanism was introduced by Peleg (1978). He also proposed the notion of an exactly strongly consistent social choice function, as an answer to the Gibbard-Satterthwaite result about the manipulability of non-dictatorial voting schemes. Peleg's idea consisted in constructing mechanisms for which "truth-telling", whilst not necessarily an equilibrium strategy, would lead to a "good" equilibrium outcome. For such mechanisms, truthful behavior became more attractive to agents. This idea led to the procedure of elimination "bad" alternatives given blocking coefficients of alternatives (see Sections 5.1 and 5.5). At the same time, Dutta and Pattanaik (1978) and Maskin (1979) constructed other examples of exact mechanisms. Oren (1981) studied the relationships between properties of exact mechanisms and blocking coefficients. He introduced for the first time the concepts of maximality and of almost-additivity of blocking coefficients and showed that they are necessary for the strong consistency of a mechanism (Proposition (5.3.2)) to hold. Additionally he established that using the Peleg procedure any element from the core is attainable (Section 4.2).

Our presentation of the mechanism with tokens is probably new, albeit Moulin (1983) describes a similar mechanism, in which the Peleg procedure is used for "splitted" (weight dependent) agents and alternatives.

Theorem (5.3.3) about implementation of the core of a maximal blocking is a generalization of the theorem by Moulin-Peleg (1982) on stable blockings. We use here a simpler mechanism (see Danilov and Sotskov (1988)).

The results in Sections 4 and 5 obtain as generalizations of the results of both Peleg (1978) and Holzman (1986). In our study of direct mechanisms, we put the emphasis on core mechanisms, selectors of core correspondences (and not on exact mechanisms). The main reason is that they are by definition sufficiently reasonable mechanisms, moreover given a blocking they are easily characterized (Section 4.6). Strong consistency of core mechanisms is closely connected with elimination schemes (Section 5.5). These schemes appeared in Peleg (1978), and also Holzman (1986), albeit in a somewhat rudimentary form. The generalization of such schemes opens a wide class of laminable blockings, whose core mechanisms are strongly consistent.

Holzman (1986) presents the first serious result about necessary conditions for strong consistency. The theorem he proves, states that a blocking with "weak" agents, which has a core SC-mechanism satisfies the Holzman condition (see Theorem (5.6.8)). The two remaining and yet more general results on necessary conditions, namely Theorem (5.5.5) and Theorem (5.6.3), are new.

The $\Pi 2$ procedure in Section 5 directly generalizes the Peleg procedure and is essentially a sum of our work on strongly consistent mechanisms.

Necessary and sufficient conditions for strong implementation (stated somewhat differently) were obtained by Dutta and Sen (1991a). Danilov and Sotskov (1991a) proposed another condition described in Appendix 5.A. Suh (1995, 1996) obtained similar results.

References

Abdou, J. (1987): Stable effectivity functions with an infinity of players and alternatives. Journal of Mathematical Economics 16, 291-295

Abdou, J., Keiding, H. (1991): Effectivity functions in social choice. Dordrecht, Kluwer Academic Publishers

Abreu, D., Sen, A. (1990): Subgame perfect Implementation: A Necessary and Almost Sufficient Condition. Journal of Economic Theory 50, 285-299

Abreu, D., Sen, A. (1991): Virtual implementation in Nash Equilibrium. Econometrica 59, 997-1021

Aizerman, M.A., Malishevski, A.V. (1981): Some aspects of general theory of choice of the best variants. Automatika i Telemachanika. n.2, 65-83 (in Russian)

Afanasiev, V.M., Lezina, Z.M. (1982): Auctions and iterative methods of social choice of variants. Automatika i Telemachanika. n.9, 9-24 (in Russian)

Arrow, K.J. (1951): Social choice and individual values. New York, Wiley

Barbera, S. (1983): Strategy-proofness and pivotal voters. A direct proof of the Gibbard-Satterthwaite theorem. International Economic Review 24, 413-417

Barbera, S., Peleg, B. (1990): Strategy-proof voting schemes with continuous preferences. Social Choice and Welfare 7, 31-38

Barbera, S. (1977): The Manipulation of Social Choice Mechanisms that Do Not Leave "Too Much" to Chance. Econometrica 45, 1573-1588

Barbera, S. (1978): Nice decision schemes. In: Gottinger, H.W., Leinfelner, W. (Eds.): Decision theory and social ethics: Issues in social choice. London, 101-117

Barbera, S., Bogomolnaia, A., van der Stel, H. (1998): Strategy-proof probabilistic rules for expected utility maximizers. Journal of Mathematical Economics 35, 89-103

Berga, D. (1998): Strategy-proofness and single-plateaued preferences. Math. Soc. Sciences 35, 105-120

Black, D.J. (1948): The decision of a committee using a special majority. Econometrica 16, 245-261

Bolker, E.D. (1969): A class of convex bodies. Trans. Amer. Math. Soc. 145, 323-345

Border, K.G., Jordan, J.S. (1983): Straightforward elections, unanimity and phantom voters. Review of Economic Studies 50, 153-170

Boros, E., Gurvich, V. (2000): Stable effectivity functions and perfect graphs. Mathematical Social Scinces 39, 175-194

Burkov, V.N., Enaleev, A.K. (1985): Optimality of the open control principle. Automatika i Telemachanika, n. 3, 73-80 (in Russian)

Clarke, E.H. (1971): Multipart pricing of public goods. Public Choice 11, 17-33

van Damme, E. (1987): Stability and Perfection of Nash equilibria. Springer, Berlin

Danilov, V.I. (1983): Models of group choice. Isvestia AS SSSR, ser. Technicheskaja kibernetika, n. 1, 143-164 (in Russian)

Danilov, V.I. (1988): Dominant-strategy mechanisms in euclidean environment. In: Makarov, V.L.,Polterovich, V.M. (Eds.): Mathematicheskie modeli ekonomicheskikh mechanismov. Moscow, CEMI AS SSSR (in Russian), 37-51

Danilov, V.I. (1992): Implementation via Nash equilibrium. Econometrica 60, 43-56

Danilov, V.I. (1994): The structure of non-manipulable social choice rules on a tree. Mathematical Social Sciences 27, 123-131

Danilov, V.I., Sotskov, A.I. (1987a): Stable blockings in social choice mechanisms. Ekonomika i Mathematicheskie Metody, v. 23, 888-898 (in Russian)

Danilov, V.I., Sotskov, A.I. (1987b): Balanced states and theorems on a core. Optimizatzija, v. 41(58) 36-49 (in Russian)

Danilov, V.I., Sotskov, A.I. (1988): Coalitionally stable social choice mechanisms. Ekonomika i Mathematicheskie Metody, v. 24, 113-124 (in Russian)

Danilov, V.I., Sotskov, A.I. (1991a): Implementation of social choice correspondences by coalition equilibria. Ekonomika i Mathematicheskie Metody, v. 27, 720-732 (in Russian)

Danilov, V.I., Sotskov, A.I. (1991b): Mechanisms of Social Choice. Nauka, Moscow (in Russian)

Danilov, V.I., Sotskov, A.I. (1993): On strongly consistent social choice functions. Journal of Mathematical Economics 22, 327-346

Demange, G. (1987): Nonmanipulable cores. Econometrica 55, 1057-1074

Demange, G. (1982): Single peaked orders on a tree. Mathematical Social Sciences 3(4) 389-396

Dutta, B. (1984): Effectivity functions and acceptable game forms. Econometrica 52, 1151-1166

Dutta, B., Pattanaik, P.K. (1978): On nicely consistent voting systems. Econometrica 46, 163-171

Dutta, B., Sen, A. (1991a): Implementation under strong equilibrium: a complete characterization. Journal of Mathematical Economics 20, 49-68

Dutta, B., Sen, A. (1991b): A necessary and sufficient condition for two-person Nash implementation. Review of Economic Studies 58, 121-128

Dutta, B., Sen, A. (1993): Implementing Generalized Condorcet Social Choice Functions via Backward Induction. Social Choice and Welfare 10, 149-160

Dutta, B., Sen, A., Vohra, R. (1995): Nash implementation through elementary mechanism in economic environments. Economic Design 1, 173-203

Dzjadyk, V.K. (1977): Introduction in a theory of the best approximation of function by polynomials. Nauka, Moscow (in Russian)

Farquharson, R. (1969): Theory of voting. Yale Univ. Press, New Haven

Fishburn, P.C. (1973): The theory of social choice. Univ. Press, Princeton

Freixas, X. (1984): A cardinal approach to straightforward probabilistic mechanisms. Journal of Mathematical Economics 34, 227-251

Gibbard, A. (1973): Manipulation of voting: A general result. Econometrica 41, 587-601

Gibbard, A. (1977): Manipulation of schemes that mix voting with chance. Econometrica 45, 665-681

Gibbard, A. (1978): Straightforwardness of game forms with lotteries as outcomes. Econometrica 46, 595-614

Golberg, A.I., Gurvich, V.A. (1986): Collective choice based on veto principle. VINITI, Moscow (in Russian)

Golberg, A.I., Gurvich, V.A., Zhizhin, M.N. (1986): Stability of collective choice. VINITI, Moscow (in Russian)

Green, J., Laffont, J.J. (1979): Incentives in public decision making. North-Holland, Amsterdam

Greenberg, J. (1985): Cores of Convex Games without Side Payments. Math. Oper. Research 10, 523-525

Groves, T. (1973): Incentives in teams. Econometrica 41, 617-631

Groves, T. (1979): Efficient collective choice with compensation. Review of Economic Studies 46, 227-242

Groves, T., Ledyard, J. (1977): Optimal allocation of public goods: A solution to the "free rider" problem. Econometrica 45, 783-809

Gurvich, V.A. (1975): Solvability of position games in pure strategies. Journal vychislitelnoi matematiki i matematicheskoi fiziki 15, 357-371 (in Russian)

Gurvich, V.A., Menshikov, I.S. (1989): Institutes for agreement. Moscow, Znanie (Matematika, Kibernetika, n. 6) (in Russian)

Holmstrom, B. (1979): Groves' scheme on restricted domains. Econometrica 47, 1137-1144

Holzman, L. (1986): On strong representations of games by social choice functions. Journal of Mathematical Economics 15, 39-57

Hurwicz, L. (1972): On informationally decentralized systems. In: McGuire, C.B., Radner, R. (Eds.): Decision and organization. North-Holland, Amsterdam

Hurwicz, L. (1979): Outcome functions yielding Walrasian and Lindahl allocations at Nash equilibrium points. Review of Economic Studies 46, 217-225

Hurwicz, L., Schmeidler, D. (1978): Construction of outcome guarantying existence and Pareto optimality of Nash equilibria. Econometrica 46, 1447-1474

Hylland, A. (1980): Strategyproofness of voting procedures with lotteries as outcomes and infinite sets of strategies. Mimeo, University of Oslo

Jackson, M.O. (1992): Implementation in undominated strategies: a look at bounded mechanisms. Review of Economic Studies 59, 757-775

Kannai, Y., Peleg, B. (1984): A note on the extension of an order on a set to the power set. Journal of Economic Theory 32, 172-175

Keiding, H. (1985): Necessary and sufficient condition for stability of effectivity functions. International Journal of Game Theory 14, 93-101

Kim, T. (1993): A stable Nash mechanism implementing Lindahl allocations for quasi-linear environments. Journal of Mathematical Economics 22, 359-372

Kiruta, A.J., Rubinov, A.M., Janovskaja, E.B. (1980): Optimal choice of distributions is complicate social-economical problems. Nauka, Leningrad (in Russian)

Kreps, D.M. (1988): Notes on the Theory of Choice. Westview Press, Boulder

Kukushkin, N.S., Morozov, V.V. (1984): Theory of non-antagonistic games. MGU, Moscow (in Russian)

Lezina, Z.M. (1987): Collective choice procedures. Automatika i Telemechanika, n. 8, 3-35 (in Russian)

Luce, R.D., Raifa, H. (1957): Games and Decisions. Wiley, New York

Makarov, I.M., Vinogradskaja, T.M., Rubchinski, A.A., Sokolov, V.V. (1982): Choice theory and decision making. Nauka, Moscow (in Russian)

Makowski, L., Ostroy, J. M. (1992): Vickrey-Clarke-Groves mechanisms in continuum economics: Characterization and existence. Journal of Mathematical Economics 21, 1-36

Maskin, E. (1977): Nash equilibrium and welfare optimality. Mimeo, Institute of Technology, Massachusetts

Maskin, E. (1979): Implementation and strong Nash equilibrium. In: Laffont, J.J. (Ed.): Aggregation and Revelation of Preferences. North-Holland, Amsterdam, 433-439

Maskin, E., Moore, J. (1987): Implementation with renegotiation. Mimeo, Harvard University

Mirkin, B.G. (1974): Problem of group choice. Nauka, Moscow (in Russian). English translation: Group Choice. Winston & Sons, Washington, 1979

Monjardet, B. (1978): An axiomatic theory of tournament aggregation. Math. Oper. Research 3, 334-351

Moore, J., Repullo, R. (1988): Subgame Perfect Implementation. Econometrica 56, 1191-1210

Moore, J., Repullo, R. (1990): Nash Implementation:A Full Characterization. Econometrica 58, 1083-1100

Moulin, H. (1979): Dominance Solvable Voting Schemes. Econometrica 47, 1337-1351

Moulin, H. (1980): On strategy-proofness and single peakedness. Public Choice 35, 437-455

Moulin, H. (1981a): The proportional veto principle. Review of Economic Studies 48, 407-416

Moulin, H. (1981b): Theorie des jeux pour léconomie et la politique. Hermann, Paris

Moulin, H. (1982): Non-cooperative implementation: a survey of recent results. Mathematical Social Sciences 3, 233-257

Moulin, H. (1983): The strategy of social choice. North-Holland, Amsterdam

Moulin, H. (1986): Characterizations of the pivotal mechanism. Journal of Public Economics 31, 53-78

Moulin, H. (1994): Social Choice. In: Aumann, R.J., Hart, S. (Eds.): Handbook of Game Theory, vol. 2. Elsevier Science, 1091-1125.

Moulin, H. (1995): Cooperative microeconomics. Prentice Hall, London

Moulin, H., Peleg, B. (1982): Cores of effectivity functions and implementation theory. Journal of Mathematical Economics, 10, 115-145

Muller, E., Satterthwaite, M.A. (1977): The equivalence of strong positive association and strategy-proofness. Journal of Economic Theory 14, 412-418

Myerson, R.B. (1991): Game Theory (Analysis of Conflict). Harvard Univ. Press, Cambridge

Nakamura, K. (1979): The vetoers in a simple game with ordinal preferences. International Journal of Game Theory 8, 55-61

Nash, J.F. (1951): Non cooperative games. Annales of Mathematics 54, 286-295

Nehring, K. (2000): Monotonicity implies generalized strategy-proofness for correspondences. Social Choice and Welfare 17, 367-375

von Neumann, J., Morgenstern, O. (1947): Theory of Games and Economic Behavior. Univ. Press, Princeton

Oren, I. (1981): The structure of exactly strongly consistent social choice functions. Journal of Mathematical Economics 8, 207-220

Peleg, B. (1978): Consistent voting systems. Econometrica 46, 153-161

Peleg, B. (1984): Game Theoretic Analysis of Voting in Committees. Univ. Press, Cambridge

Polterovich, V.M. (1973): Economic equilibrium and optimum. Ekonomika i Mathematicheskie Metody, v. 9, 835-845 (in Russian)

Roberts, K. (1979): The characterization of implementable choice rules. In: Laffont, J.J. (Ed.): Aggregation and Revelation of Preferences. North-Holland, Amsterdam, 321-348

Roth, A. Sotomayor, M. (1990): Two-sided Matching. A Study in Game-theoretic Modeling and Analysis. Econometric Society Monographs, vol. 18, Cambridge University Press

Saijo, T. (1988): Strategy space reduction in Maskin's theorem sufficient conditions for Nash implementation. Econometrica 56, 693-700

Satterthwaite, M. (1975) Strategy-proofness and Arrow's conditions. Journal of Economic Theory 10, 187-217

Sen, A. (1970): Collective Choice and Social Welfare. Holden Day, San Francisco

Sen, A. (1986): Social choice theory. In Arrow, K.J., Intriligator, M.D. (Eds.): Handbook of Mathematical Economics, vol. 3. North-Holland, Amsterdam, 1073-1181

Shapley, L., Scarf, H. (1971): On cores and indivisibility. Journal of Mathematical Economics 1, 23-37

Sönmez, T. (1996): Implementation in generalized matching problems. Journal of Mathematical Economics 26, 429-435

Sönmez, T. (1999): Strategy-Proofness and Essentially Single-Valued Cores. Econometrica 67, 677-689

Sotskov, A.I. (1994): Vetoing in social choice with blockings. Mathematical Social Sciences 27, 203-216

Suh, S.-C. (1995): An algorithm for checking strong Nash implementability. Journal of Mathematical Economics 25, 109-122

Suh, S.-C. (1996): Implementation with coalition formation: A complete characterization. Journal of Mathematical Economics 26, 409-428

Tian, G. (1993): Implementing Lindahl allocations by a withholding mechanism. Journal of Mathematical Economics 22, 169-180

Vasin, A.A., Gurvich, V.A. (1980): Coalition situations of equilibria in meta-games. Vestnik MGU, Ser. 15, Vychislitelnaja matematika i kibernetika 38-44 (in Russian)

Vilkov, V.B. (1977): Convex games without of side payments. Vestnik LGU, n. 7, 21-24 (in Russian)

Walker, M. (1980): On the non-existence of a dominant strategy mechanism for making optimal public decisions. Econometrica 48, 1521-1540

Walker, M. (1981): A simple incentive compatible scheme for attaining Lindahl allocations. Econometrica 49, 65-72

Wilson, R.J. (1972): Introduction to Graph Theory. Oliver and Boyd, Edinburgh

Index