CHAPTER I

FORMULATION OF THE ECONOMIC PROBLEM

1. The Mathematical Method in Economics

1.1. Introductory Remarks

1.1.1. The purpose of this book is to present a discussion of some fundamental questions of economic theory which require a treatment different from that which they have found thus far in the literature. The analysis is concerned with some basic problems arising from a study of economic behavior which have been the center of attention of economists for a long time. They have their origin in the attempts to find an exact description of the endeavor of the individual to obtain a maximum of utility, or, in the case of the entrepreneur, a maximum of profit. It is well known what considerable—and in fact unsurmounted—difficulties this task involves given even a limited number of typical situations, as, for example, in the case of the exchange of goods, direct or indirect, between two or more persons, of bilateral monopoly, of duopoly, of oligopoly, and of free competition. It will be made clear that the structure of these problems, familiar to every student of economics, is in many respects quite different from the way in which they are conceived at the present time. It will appear, furthermore, that their exact positing and subsequent solution can only be achieved with the aid of mathematical methods which diverge considerably from the techniques applied by older or by contemporary mathematical economists.

1.1.2. Our considerations will lead to the application of the mathematical theory of "games of strategy" developed by one of us in several successive stages in 1928 and 1940–1941.[1] After the presentation of this theory, its application to economic problems in the sense indicated above will be undertaken. It will appear that it provides a new approach to a number of economic questions as yet unsettled.

We shall first have to find in which way this theory of games can be brought into relationship with economic theory, and what their common elements are. This can be done best by stating briefly the nature of some fundamental economic problems so that the common elements will be seen clearly. It will then become apparent that there is not only nothing artificial in establishing this relationship but that on the contrary this

[1] The first phases of this work were published: *J. von Neumann*, "Zur Theorie der Gesellschaftsspiele," Math. Annalen, vol. 100 (1928), pp. 295–320. The subsequent completion of the theory, as well as the more detailed elaboration of the considerations of loc. cit. above, are published here for the first time.

theory of games of strategy is the proper instrument with which to develop a theory of economic behavior.

One would misunderstand the intent of our discussions by interpreting them as merely pointing out an analogy between these two spheres. We hope to establish satisfactorily, after developing a few plausible schematizations, that the typical problems of economic behavior become strictly identical with the mathematical notions of suitable games of strategy.

1.2. Difficulties of the Application of the Mathematical Method

1.2.1. It may be opportune to begin with some remarks concerning the nature of economic theory and to discuss briefly the question of the role which mathematics may take in its development.

First let us be aware that there exists at present no universal system of economic theory and that, if one should ever be developed, it will very probably not be during our lifetime. The reason for this is simply that economics is far too difficult a science to permit its construction rapidly, especially in view of the very limited knowledge and imperfect description of the facts with which economists are dealing. Only those who fail to appreciate this condition are likely to attempt the construction of universal systems. Even in sciences which are far more advanced than economics, like physics, there is no universal system available at present.

To continue the simile with physics: It happens occasionally that a particular physical theory appears to provide the basis for a universal system, but in all instances up to the present time this appearance has not lasted more than a decade at best. The everyday work of the research physicist is certainly not involved with such high aims, but rather is concerned with special problems which are "mature." There would probably be no progress at all in physics if a serious attempt were made to enforce that super-standard. The physicist works on individual problems, some of great practical significance, others of less. Unifications of fields which were formerly divided and far apart may alternate with this type of work. However, such fortunate occurrences are rare and happen only after each field has been thoroughly explored. Considering the fact that economics is much more difficult, much less understood, and undoubtedly in a much earlier stage of its evolution as a science than physics, one should clearly not expect more than a development of the above type in economics either.

Second we have to notice that the differences in scientific questions make it necessary to employ varying methods which may afterwards have to be discarded if better ones offer themselves. This has a double implication: In some branches of economics the most fruitful work may be that of careful, patient description; indeed this may be by far the largest domain for the present and for some time to come. In others it may be possible to develop already a theory in a strict manner, and for that purpose the use of mathematics may be required.

Mathematics has actually been used in economic theory, perhaps even in an exaggerated manner. In any case its use has not been highly successful. This is contrary to what one observes in other sciences: There mathematics has been applied with great success, and most sciences could hardly get along without it. Yet the explanation for this phenomenon is fairly simple.

1.2.2. It is not that there exists any fundamental reason why mathematics should not be used in economics. The arguments often heard that because of the human element, of the psychological factors etc., or because there is—allegedly—no measurement of important factors, mathematics will find no application, can all be dismissed as utterly mistaken. Almost all these objections have been made, or might have been made, many centuries ago in fields where mathematics is now the chief instrument of analysis. This "might have been" is meant in the following sense: Let us try to imagine ourselves in the period which preceded the mathematical or almost mathematical phase of the development in physics, that is the 16th century, or in chemistry and biology, that is the 18th century. Taking for granted the skeptical attitude of those who object to mathematical economics in principle, the outlook in the physical and biological sciences at these early periods can hardly have been better than that in economics—*mutatis mutandis*—at present.

As to the lack of measurement of the most important factors, the example of the theory of heat is most instructive; before the development of the mathematical theory the possibilities of quantitative measurements were less favorable there than they are now in economics. The precise measurements of the quantity and quality of heat (energy and temperature) were the outcome and not the antecedents of the mathematical theory. This ought to be contrasted with the fact that the quantitative and exact notions of prices, money and the rate of interest were already developed centuries ago.

A further group of objections against quantitative measurements in economics, centers around the lack of indefinite divisibility of economic quantities. This is supposedly incompatible with the use of the infinitesimal calculus and hence (!) of mathematics. It is hard to see how such objections can be maintained in view of the atomic theories in physics and chemistry, the theory of quanta in electrodynamics, etc., and the notorious and continued success of mathematical analysis within these disciplines.

At this point it is appropriate to mention another familiar argument of economic literature which may be revived as an objection against the mathematical procedure.

1.2.3. In order to elucidate the conceptions which we are applying to economics, we have given and may give again some illustrations from physics. There are many social scientists who object to the drawing of such parallels on various grounds, among which is generally found the assertion that economic theory cannot be modeled after physics since it is a science of social, of human phenomena, has to take psychology into account, etc. Such statements are at least premature. It is without doubt reasonable to discover what has led to progress in other sciences, and to investigate whether the application of the same principles may not lead to progress in economics also. Should the need for the application of different principles arise, it could be revealed only in the course of the actual development of economic theory. This would itself constitute a major revolution. But since most assuredly we have not yet reached such a state—and it is by no means certain that there ever will be need for entirely different scientific principles—it would be very unwise to consider anything else than the pursuit of our problems in the manner which has resulted in the establishment of physical science.

1.2.4. The reason why mathematics has not been more successful in economics must, consequently, be found elsewhere. The lack of real success is largely due to a combination of unfavorable circumstances, some of which can be removed gradually. To begin with, the economic problems were not formulated clearly and are often stated in such vague terms as to make mathematical treatment *a priori* appear hopeless because it is quite uncertain what the problems really are. There is no point in using exact methods where there is no clarity in the concepts and issues to which they are to be applied. Consequently the initial task is to clarify the knowledge of the matter by further careful descriptive work. But even in those parts of economics where the descriptive problem has been handled more satisfactorily, mathematical tools have seldom been used appropriately. They were either inadequately handled, as in the attempts to determine a general economic equilibrium by the mere counting of numbers of equations and unknowns, or they led to mere translations from a literary form of expression into symbols, without any subsequent mathematical analysis.

Next, the empirical background of economic science is definitely inadequate. Our knowledge of the relevant facts of economics is incomparably smaller than that commanded in physics at the time when the mathematization of that subject was achieved. Indeed, the decisive break which came in physics in the seventeenth century, specifically in the field of mechanics, was possible only because of previous developments in astronomy. It was backed by several millennia of systematic, scientific, astronomical observation, culminating in an observer of unparalleled caliber, Tycho de Brahe. Nothing of this sort has occurred in economic science. It would have been absurd in physics to expect Kepler and Newton without Tycho,—and there is no reason to hope for an easier development in economics.

These obvious comments should not be construed, of course, as a disparagement of statistical-economic research which holds the real promise of progress in the proper direction.

It is due to the combination of the above mentioned circumstances that mathematical economics has not achieved very much. The underlying vagueness and ignorance has not been dispelled by the inadequate and inappropriate use of a powerful instrument that is very difficult to handle.

In the light of these remarks we may describe our own position as follows: The aim of this book lies not in the direction of empirical research. The advancement of that side of economic science, on anything like the scale which was recognized above as necessary, is clearly a task of vast proportions. It may be hoped that as a result of the improvements of scientific technique and of experience gained in other fields, the development of descriptive economics will not take as much time as the comparison with astronomy would suggest. But in any case the task seems to transcend the limits of any individually planned program.

We shall attempt to utilize only some commonplace experience concerning human behavior which lends itself to mathematical treatment and which is of economic importance.

We believe that the possibility of a mathematical treatment of these phenomena refutes the "fundamental" objections referred to in 1.2.2.

It will be seen, however, that this process of mathematization is not at all obvious. Indeed, the objections mentioned above may have their roots partly in the rather obvious difficulties of any direct mathematical approach. We shall find it necessary to draw upon techniques of mathematics which have not been used heretofore in mathematical economics, and it is quite possible that further study may result in the future in the creation of new mathematical disciplines.

To conclude, we may also observe that part of the feeling of dissatisfaction with the mathematical treatment of economic theory derives largely from the fact that frequently one is offered not proofs but mere assertions which are really no better than the same assertions given in literary form. Very frequently the proofs are lacking because a mathematical treatment has been attempted of fields which are so vast and so complicated that for a long time to come—until much more empirical knowledge is acquired—there is hardly any reason at all to expect progress *more mathematico*. The fact that these fields have been attacked in this way—as for example the theory of economic fluctuations, the time structure of production, etc.—indicates how much the attendant difficulties are being underestimated. They are enormous and we are now in no way equipped for them.

1.2.5. We have referred to the nature and the possibilities of those changes in mathematical technique—in fact, in mathematics itself—which a successful application of mathematics to a new subject may produce. It is important to visualize these in their proper perspective.

It must not be forgotten that these changes may be very considerable. The decisive phase of the application of mathematics to physics—Newton's creation of a rational discipline of mechanics—brought about, and can hardly be separated from, the discovery of the infinitesimal calculus. (There are several other examples, but none stronger than this.)

The importance of the social phenomena, the wealth and multiplicity of their manifestations, and the complexity of their structure, are at least equal to those in physics. It is therefore to be expected—or feared—that mathematical discoveries of a stature comparable to that of calculus will be needed in order to produce decisive success in this field. (Incidentally, it is in this spirit that our present efforts must be discounted.) *A fortiori* it is unlikely that a mere repetition of the tricks which served us so well in physics will do for the social phenomena too. The probability is very slim indeed, since it will be shown that we encounter in our discussions some mathematical problems which are quite different from those which occur in physical science.

These observations should be remembered in connection with the current overemphasis on the use of calculus, differential equations, etc., as the main tools of mathematical economics.

1.3. Necessary Limitations of the Objectives

1.3.1. We have to return, therefore, to the position indicated earlier: It is necessary to begin with those problems which are described clearly, even if they should not be as important from any other point of view. It should be added, moreover, that a treatment of these manageable problems may lead to results which are already fairly well known, but the exact proofs may nevertheless be lacking. Before they have been given the respective theory simply does not exist as a scientific theory. The movements of the planets were known long before their courses had been calculated and explained by Newton's theory, and the same applies in many smaller and less dramatic instances. And similarly in economic theory, certain results—say the indeterminateness of bilateral monopoly—may be known already. Yet it is of interest to derive them again from an exact theory. The same could and should be said concerning practically all established economic theorems.

1.3.2. It might be added finally that we do not propose to raise the question of the practical significance of the problems treated. This falls in line with what was said above about the selection of fields for theory. The situation is not different here from that in other sciences. There too the most important questions from a practical point of view may have been completely out of reach during long and fruitful periods of their development. This is certainly still the case in economics, where it is of utmost importance to know how to stabilize employment, how to increase the national income, or how to distribute it adequately. Nobody can really answer these questions, and we need not concern ourselves with the pretension that there can be scientific answers at present.

The great progress in every science came when, in the study of problems which were modest as compared with ultimate aims, methods were developed which could be extended further and further. The free fall is a very trivial physical phenomenon, but it was the study of this exceedingly simple fact and its comparison with the astronomical material, which brought forth mechanics.

It seems to us that the same standard of modesty should be applied in economics. It is futile to try to explain—and "systematically" at that—everything economic. The sound procedure is to obtain first utmost precision and mastery in a limited field, and then to proceed to another, somewhat wider one, and so on. This would also do away with the unhealthy practice of applying so-called theories to economic or social reform where they are in no way useful.

We believe that it is necessary to know as much as possible about the behavior of the individual and about the simplest forms of exchange. This standpoint was actually adopted with remarkable success by the founders of the marginal utility school, but nevertheless it is not generally accepted. Economists frequently point to much larger, more "burning" questions, and brush everything aside which prevents them from making statements about these. The experience of more advanced sciences, for example physics, indicates that this impatience merely delays progress, including that of the treatment of the "burning" questions. There is no reason to assume the existence of shortcuts.

1.4. Concluding Remarks

1.4. It is essential to realize that economists can expect no easier fate than that which befell scientists in other disciplines. It seems reasonable to expect that they will have to take up first problems contained in the very simplest facts of economic life and try to establish theories which explain them and which really conform to rigorous scientific standards. We can have enough confidence that from then on the science of economics will grow further, gradually comprising matters of more vital importance than those with which one has to begin.[1]

The field covered in this book is very limited, and we approach it in this sense of modesty. We do not worry at all if the results of our study conform with views gained recently or held for a long time, for what is important is the gradual development of a theory, based on a careful analysis of the ordinary everyday interpretation of economic facts. This preliminary stage is necessarily *heuristic*, i.e. the phase of transition from unmathematical plausibility considerations to the formal procedure of mathematics. The theory finally obtained must be mathematically rigorous and conceptually general. Its first applications are necessarily to elementary problems where the result has never been in doubt and no theory is actually required. At this early stage the application serves to corroborate the theory. The next stage develops when the theory is applied

[1] The beginning is actually of a certain significance, because the forms of exchange between a few individuals are the same as those observed on some of the most important markets of modern industry, or in the case of barter exchange between states in international trade.

to somewhat more complicated situations in which it may already lead to a certain extent beyond the obvious and the familiar. Here theory and application corroborate each other mutually. Beyond this lies the field of real success: genuine prediction by theory. It is well known that all mathematized sciences have gone through these successive phases of evolution.

2. Qualitative Discussion of the Problem of Rational Behavior

2.1. The Problem of Rational Behavior

2.1.1. The subject matter of economic theory is the very complicated mechanism of prices and production, and of the gaining and spending of incomes. In the course of the development of economics it has been found, and it is now well-nigh universally agreed, that an approach to this vast problem is gained by the analysis of the behavior of the individuals which constitute the economic community. This analysis has been pushed fairly far in many respects, and while there still exists much disagreement the significance of the approach cannot be doubted, no matter how great its difficulties may be. The obstacles are indeed considerable, even if the investigation should at first be limited to conditions of economics statics, as they well must be. One of the chief difficulties lies in properly describing the assumptions which have to be made about the motives of the individual. This problem has been stated traditionally by assuming that the consumer desires to obtain a maximum of utility or satisfaction and the entrepreneur a maximum of profits.

The conceptual and practical difficulties of the notion of utility, and particularly of the attempts to describe it as a number, are well known and their treatment is not among the primary objectives of this work. We shall nevertheless be forced to discuss them in some instances, in particular in 3.3. and 3.5. Let it be said at once that the standpoint of the present book on this very important and very interesting question will be mainly opportunistic. We wish to concentrate on one problem—which is not that of the measurement of utilities and of preferences—and we shall therefore attempt to simplify all other characteristics as far as reasonably possible. We shall therefore assume that the aim of all participants in the economic system, consumers as well as entrepreneurs, is money, or equivalently a single monetary commodity. This is supposed to be unrestrictedly divisible and substitutable, freely transferable and identical, even in the quantitative sense, with whatever "satisfaction" or "utility" is desired by each participant. (For the quantitative character of utility, cf. 3.3. quoted above.)

It is sometimes claimed in economic literature that discussions of the notions of utility and preference are altogether unnecessary, since these are purely verbal definitions with no empirically observable consequences, i.e., entirely tautological. It does not seem to us that these notions are qualitatively inferior to certain well established and indispensable notions in physics, like force, mass, charge, etc. That is, while they are in their immediate form merely definitions, they become subject to empirical control through the theories which are built upon them—and in no other way. Thus the notion of utility is raised above the status of a tautology by such economic theories as make use of it and the results of which can be compared with experience or at least with common sense.

2.1.2. The individual who attempts to obtain these respective maxima is also said to act "rationally." But it may safely be stated that there exists, at present, no satisfactory treatment of the question of rational behavior. There may, for example, exist several ways by which to reach the optimum position; they may depend upon the knowledge and understanding which the individual has and upon the paths of action open to him. A study of all these questions in qualitative terms will not exhaust them, because they imply, as must be evident, quantitative relationships. It would, therefore, be necessary to formulate them in quantitative terms so that all the elements of the qualitative description are taken into consideration. This is an exceedingly difficult task, and we can safely say that it has not been accomplished in the extensive literature about the topic. The chief reason for this lies, no doubt, in the failure to develop and apply suitable mathematical methods to the problem; this would have revealed that the maximum problem which is supposed to correspond to the notion of rationality is not at all formulated in an unambiguous way. Indeed, a more exhaustive analysis (to be given in 4.3.-4.5.) reveals that the significant relationships are much more complicated than the popular and the "philosophical" use of the word "rational" indicates.

A valuable qualitative preliminary description of the behavior of the individual is offered by the Austrian School, particularly in analyzing the economy of the isolated "Robinson Crusoe." We may have occasion to note also some considerations of Böhm-Bawerk concerning the exchange between two or more persons. The more recent exposition of the theory of the individual's choices in the form of indifference curve analysis builds up on the very same facts or alleged facts but uses a method which is often held to be superior in many ways. Concerning this we refer to the discussions in 2.1.1. and 3.3.

We hope, however, to obtain a real understanding of the problem of exchange by studying it from an altogether different angle; this is, from the perspective of a "game of strategy." Our approach will become clear presently, especially after some ideas which have been advanced, say by Böhm-Bawerk—whose views may be considered only as a prototype of this theory—are given correct quantitative formulation.

2.2. "Robinson Crusoe" Economy and Social Exchange Economy

2.2.1. Let us look more closely at the type of economy which is represented by the "Robinson Crusoe" model, that is an economy of an isolated single person or otherwise organized under a single will. This economy is confronted with certain quantities of commodities and a number of wants which they may satisfy. The problem is to obtain a maximum satisfaction. This is—considering in particular our above assumption of the numerical character of utility—indeed an ordinary maximum problem, its difficulty depending apparently on the number of variables and on the nature of the function to be maximized; but this is more of a practical difficulty than a theoretical one.[1] If one abstracts from continuous production and from the fact that consumption too stretches over time (and often uses durable consumers' goods), one obtains the simplest possible model. It was thought possible to use it as the very basis for economic theory, but this attempt—notably a feature of the Austrian version—was often contested. The chief objection against using this very simplified model of an isolated individual for the theory of a social exchange economy is that it does not represent an individual exposed to the manifold social influences. Hence, it is said to analyze an individual who might behave quite differently if his choices were made in a social world where he would be exposed to factors of imitation, advertising, custom, and so on. These factors certainly make a great difference, but it is to be questioned whether they change the formal properties of the process of maximizing. Indeed the latter has never been implied, and since we are concerned with this problem alone, we can leave the above social considerations out of account.

Some other differences between "Crusoe" and a participant in a social exchange economy will not concern us either. Such is the non-existence of money as a means of exchange in the first case where there is only a standard of calculation, for which purpose any commodity can serve. This difficulty indeed has been ploughed under by our assuming in 2.1.2. a quantitative and even monetary notion of utility. We emphasize again: Our interest lies in the fact that even after all these drastic simplifications Crusoe is confronted with a formal problem quite different from the one a participant in a social economy faces.

2.2.2. Crusoe is given certain physical data (wants and commodities) and his task is to combine and apply them in such a fashion as to obtain a maximum resulting satisfaction. There can be no doubt that he controls exclusively all the variables upon which this result depends—say the allotting of resources, the determination of the uses of the same commodity for different wants, etc.[2]

Thus Crusoe faces an ordinary maximum problem, the difficulties of which are of a purely technical—and not conceptual—nature, as pointed out.

2.2.3. Consider now a participant in a social exchange economy. His problem has, of course, many elements in common with a maximum prob-

[1] It is not important for the following to determine whether its theory is complete in all its aspects.

[2] Sometimes uncontrollable factors also intervene, e.g. the weather in agriculture. These however are purely statistical phenomena. Consequently they can be eliminated by the known procedures of the calculus of probabilities: i.e., by determining the probabilities of the various alternatives and by introduction of the notion of "mathematical expectation." Cf. however the influence on the notion of utility, discussed in 3.3.

lem. But it also contains some, very essential, elements of an entirely different nature. He too tries to obtain an optimum result. But in order to achieve this, he must enter into relations of exchange with others. If two or more persons exchange goods with each other, then the result for each one will depend in general not merely upon his own actions but on those of the others as well. Thus each participant attempts to maximize a function (his above-mentioned "result") of which he does not control all variables. This is certainly no maximum problem, but a peculiar and disconcerting mixture of several conflicting maximum problems. Every participant is guided by another principle and neither determines all variables which affect his interest.

This kind of problem is nowhere dealt with in classical mathematics. We emphasize at the risk of being pedantic that this is no conditional maximum problem, no problem of the calculus of variations, of functional analysis, etc. It arises in full clarity, even in the most "elementary" situations, e.g., when all variables can assume only a finite number of values.

A particularly striking expression of the popular misunderstanding about this pseudo-maximum problem is the famous statement according to which the purpose of social effort is the "greatest possible good for the greatest possible number." A guiding principle cannot be formulated by the requirement of maximizing two (or more) functions at once.

Such a principle, taken literally, is self-contradictory. (In general one function will have no maximum where the other function has one.) It is no better than saying, e.g., that a firm should obtain maximum prices at maximum turnover, or a maximum revenue at minimum outlay. If some order of importance of these principles or some weighted average is meant, this should be stated. However, in the situation of the participants in a social economy nothing of that sort is intended, but all maxima are desired at once—by various participants.

One would be mistaken to believe that it can be obviated, like the difficulty in the Crusoe case mentioned in footnote 2 on p. 10, by a mere recourse to the devices of the theory of probability. Every participant can determine the variables which describe his own actions but not those of the others. Nevertheless those "alien" variables cannot, from his point of view, be described by statistical assumptions. This is because the others are guided, just as he himself, by rational principles—whatever that may mean —and no *modus procedendi* can be correct which does not attempt to understand those principles and the interactions of the conflicting interests of all participants.

Sometimes some of these interests run more or less parallel—then we are nearer to a simple maximum problem. But they can just as well be opposed. The general theory must cover all these possibilities, all intermediary stages, and all their combinations.

2.2.4. The difference between Crusoe's perspective and that of a participant in a social economy can also be illustrated in this way: Apart from those variables which his will controls, Crusoe is given a number of data which are "dead"; they are the unalterable physical background of the situation. (Even when they are apparently variable, cf. footnote 2 on p. 10, they are really governed by fixed statistical laws.) Not a single datum with which he has to deal reflects another person's will or intention of an economic kind—based on motives of the same nature as his own. A participant in a social exchange economy, on the other hand, faces data of this last type as well: they are the product of other participants' actions and volitions (like prices). His actions will be influenced by his expectation of these, and they in turn reflect the other participants' expectation of his actions.

Thus the study of the Crusoe economy and the use of the methods applicable to it, is of much more limited value to economic theory than has been assumed heretofore even by the most radical critics. The grounds for this limitation lie not in the field of those social relationships which we have mentioned before—although we do not question their significance—but rather they arise from the conceptual differences between the original (Crusoe's) maximum problem and the more complex problem sketched above.

We hope that the reader will be convinced by the above that we face here and now a really conceptual—and not merely technical—difficulty. And it is this problem which the theory of "games of strategy" is mainly devised to meet.

2.3. The Number of Variables and the Number of Participants

2.3.1. The formal set-up which we used in the preceding paragraphs to indicate the events in a social exchange economy made use of a number of "variables" which described the actions of the participants in this economy. Thus every participant is allotted a set of variables, "his" variables, which together completely describe his actions, i.e. express precisely the manifestations of his will. We call these sets the partial sets of variables. The partial sets of all participants constitute together the set of all variables, to be called the total set. So the total number of variables is determined first by the number of participants, i.e. of partial sets, and second by the number of variables in every partial set.

From a purely mathematical point of view there would be nothing objectionable in treating all the variables of any one partial set as a single variable, "the" variable of the participant corresponding to this partial set. Indeed, this is a procedure which we are going to use frequently in our mathematical discussions; it makes absolutely no difference conceptually, and it simplifies notations considerably.

For the moment, however, we propose to distinguish from each other the variables within each partial set. The economic models to which one is naturally led suggest that procedure; thus it is desirable to describe for every participant the quantity of every particular good he wishes to acquire by a separate variable, etc.

2.3.2. Now we must emphasize that any increase of the number of variables inside a participant's partial set may complicate our problem technically, but only technically. Thus in a Crusoe economy—where there exists only one participant and only one partial set which then coincides with the total set—this may make the necessary determination of a maximum technically more difficult, but it will not alter the "pure maximum" character of the problem. If, on the other hand, the number of participants—i.e., of the partial sets of variables—is increased, something of a very different nature happens. To use a terminology which will turn out to be significant, that of games, this amounts to an increase in the number of players in the game. However, to take the simplest cases, a three-person game is very fundamentally different from a two-person game, a four-person game from a three-person game, etc. The combinatorial complications of the problem—which is, as we saw, no maximum problem at all—increase tremendously with every increase in the number of players, —as our subsequent discussions will amply show.

We have gone into this matter in such detail particularly because in most models of economics a peculiar mixture of these two phenomena occurs. Whenever the number of players, i.e. of participants in a social economy, increases, the complexity of the economic system usually increases too; e.g. the number of commodities and services exchanged, processes of production used, etc. Thus the number of variables in every participant's partial set is likely to increase. But the number of participants, i.e. of partial sets, has increased too. Thus both of the sources which we discussed contribute *pari passu* to the total increase in the number of variables. It is essential to visualize each source in its proper role.

2.4. The Case of Many Participants: Free Competition

2.4.1. In elaborating the contrast between a Crusoe economy and a social exchange economy in 2.2.2.-2.2.4., we emphasized those features of the latter which become more prominent when the number of participants —while greater than 1—is of moderate size. The fact that every participant is influenced by the anticipated reactions of the others to his own measures, and that this is true for each of the participants, is most strikingly the crux of the matter (as far as the sellers are concerned) in the classical problems of duopoly, oligopoly, etc. When the number of participants becomes really great, some hope emerges that the influence of every particular participant will become negligible, and that the above difficulties may recede and a more conventional theory become possible. These are, of course, the classical conditions of "free competition." Indeed, this was the starting point of much of what is best in economic theory. Compared with this case of great numbers—free competition—the cases of small numbers on the side of the sellers—monopoly, duopoly, oligopoly—were even considered to be exceptions and abnormities. (Even in these cases the number of participants is still very large in view of the competition among the buyers. The cases involving really small numbers are those of bilateral monopoly, of exchange between a monopoly and an oligopoly, or two oligopolies, etc.)

2.4.2. In all fairness to the traditional point of view this much ought to be said: It is a well known phenomenon in many branches of the exact and physical sciences that very great numbers are often easier to handle than those of medium size. An almost exact theory of a gas, containing about 10^{25} freely moving particles, is incomparably easier than that of the solar system, made up of 9 major bodies; and still more than that of a multiple star of three or four objects of about the same size. This is, of course, due to the excellent possibility of applying the laws of statistics and probabilities in the first case.

This analogy, however, is far from perfect for our problem. The theory of mechanics for 2, 3, 4, · · · bodies is well known, and in its general theoretical (as distinguished from its special and computational) form is the foundation of the statistical theory for great numbers. For the social exchange economy—i.e. for the equivalent "games of strategy"—the theory of 2, 3, 4, · · · participants was heretofore lacking. It is this need that our previous discussions were designed to establish and that our subsequent investigations will endeavor to satisfy. In other words, only after the theory for moderate numbers of participants has been satisfactorily developed will it be possible to decide whether extremely great numbers of participants simplify the situation. Let us say it again: We share the hope—chiefly because of the above-mentioned analogy in other fields!—that such simplifications will indeed occur. The current assertions concerning free competition appear to be very valuable surmises and inspiring anticipations of results. But they are not results and it is scientifically unsound to treat them as such as long as the conditions which we mentioned above are not satisfied.

There exists in the literature a considerable amount of theoretical discussion purporting to show that the zones of indeterminateness (of rates of exchange)—which undoubtedly exist when the number of participants is small—narrow and disappear as the number increases. This then would provide a continuous transition into the ideal case of free competition—for a very great number of participants—where all solutions would be sharply and uniquely determined. While it is to be hoped that this indeed turns out to be the case in sufficient generality, one cannot concede that anything like this contention has been established conclusively thus far. There is no getting away from it: The problem must be formulated, solved and understood for small numbers of participants before anything can be proved about the changes of its character in any limiting case of large numbers, such as free competition.

2.4.3. A really fundamental reopening of this subject is the more desirable because it is neither certain nor probable that a mere increase in the number of participants will always lead *in fine* to the conditions of free competition. The classical definitions of free competition all involve further postulates besides the greatness of that number. E.g., it is clear that if certain great groups of participants will—for any reason whatsoever—act together, then the great number of participants may not become effective; the decisive exchanges may take place directly between large "coalitions,"[1] few in number, and not between individuals, many in number, acting independently. Our subsequent discussion of "games of strategy" will show that the role and size of "coalitions" is decisive throughout the entire subject. Consequently the above difficulty—though not new—still remains the crucial problem. Any satisfactory theory of the "limiting transition" from small numbers of participants to large numbers will have to explain under what circumstances such big coalitions will or will not be formed—i.e. when the large numbers of participants will become effective and lead to a more or less free competition. Which of these alternatives is likely to arise will depend on the physical data of the situation. Answering this question is, we think, the real challenge to any theory of free competition.

2.5. The "Lausanne" Theory

2.5. This section should not be concluded without a reference to the equilibrium theory of the Lausanne School and also of various other systems which take into consideration "individual planning" and interlocking individual plans. All these systems pay attention to the interdependence of the participants in a social economy. This, however, is invariably done under far-reaching restrictions. Sometimes free competition is assumed, after the introduction of which the participants face fixed conditions and act like a number of Robinson Crusoes—solely bent on maximizing their individual satisfactions, which under these conditions are again independent. In other cases other restricting devices are used, all of which amount to excluding the free play of "coalitions" formed by any or all types of participants. There are frequently definite, but sometimes hidden, assumptions concerning the ways in which their partly parallel and partly opposite interests will influence the participants, and cause them to cooperate or not, as the case may be. We hope we have shown that such a procedure amounts to a *petitio principii*—at least on the plane on which we should like to put the discussion. It avoids the real difficulty and deals with a verbal problem, which is not the empirically given one. Of course we do not wish to question the significance of these investigations—but they do not answer our queries.

3. The Notion of Utility

3.1. Preferences and Utilities

3.1.1. We have stated already in 2.1.1. in what way we wish to describe the fundamental concept of individual preferences by the use of a rather

[1] Such as trade unions, consumers' cooperatives, industrial cartels, and conceivably some organizations more in the political sphere.

far-reaching notion of utility. Many economists will feel that we are assuming far too much (cf. the enumeration of the properties we postulated in 2.1.1.), and that our standpoint is a retrogression from the more cautious modern technique of "indifference curves."

Before attempting any specific discussion let us state as a general excuse that our procedure at worst is only the application of a classical preliminary device of scientific analysis: To divide the difficulties, i.e. to concentrate on one (the subject proper of the investigation in hand), and to reduce all others as far as reasonably possible, by simplifying and schematizing assumptions. We should also add that this high handed treatment of preferences and utilities is employed in the main body of our discussion, but we shall incidentally investigate to a certain extent the changes which an avoidance of the assumptions in question would cause in our theory (cf. 66., 67.).

We feel, however, that one part of our assumptions at least—that of treating utilities as numerically measurable quantities—is not quite as radical as is often assumed in the literature. We shall attempt to prove this particular point in the paragraphs which follow. It is hoped that the reader will forgive us for discussing only incidentally in a condensed form a subject of so great a conceptual importance as that of utility. It seems however that even a few remarks may be helpful, because the question of the measurability of utilities is similar in character to corresponding questions in the physical sciences.

3.1.2. Historically, utility was first conceived as quantitatively measurable, i.e. as a number. Valid objections can be and have been made against this view in its original, naive form. It is clear that every measurement—or rather every claim of measurability—must ultimately be based on some immediate sensation, which possibly cannot and certainly need not be analyzed any further.[1] In the case of utility the immediate sensation of preference—of one object or aggregate of objects as against another—provides this basis. But this permits us only to say when for one person one utility is greater than another. It is not in itself a basis for numerical comparison of utilities for one person nor of any comparison between different persons. Since there is no intuitively significant way to add two utilities for the same person, the assumption that utilities are of non-numerical character even seems plausible. The modern method of indifference curve analysis is a mathematical procedure to describe this situation.

3.2. Principles of Measurement: Preliminaries

3.2.1. All this is strongly reminiscent of the conditions existant at the beginning of the theory of heat: that too was based on the intuitively clear concept of one body feeling warmer than another, yet there was no immediate way to express significantly by how much, or how many times, or in what sense.

[1] Such as the sensations of light, heat, muscular effort, etc., in the corresponding branches of physics.

This comparison with heat also shows how little one can forecast *a priori* what the ultimate shape of such a theory will be. The above crude indications do not disclose at all what, as we now know, subsequently happened. It turned out that heat permits quantitative description not by one number but by two: the quantity of heat and temperature. The former is rather directly numerical because it turned out to be additive and also in an unexpected way connected with mechanical energy which was numerical anyhow. The latter is also numerical, but in a much more subtle way; it is not additive in any immediate sense, but a rigid numerical scale for it emerged from the study of the concordant behavior of ideal gases, and the role of absolute temperature in connection with the entropy theorem.

3.2.2. The historical development of the theory of heat indicates that one must be extremely careful in making negative assertions about any concept with the claim to finality. Even if utilities look very unnumerical today, the history of the experience in the theory of heat may repeat itself, and nobody can foretell with what ramifications and variations.[1] And it should certainly not discourage theoretical explanations of the formal possibilities of a numerical utility.

3.3. Probability and Numerical Utilities

3.3.1. We can go even one step beyond the above double negations—which were only cautions against premature assertions of the impossibility of a numerical utility. It can be shown that under the conditions on which the indifference curve analysis is based very little extra effort is needed to reach a numerical utility.

It has been pointed out repeatedly that a numerical utility is dependent upon the possibility of comparing differences in utilities. This may seem—and indeed is—a more far-reaching assumption than that of a mere ability to state preferences. But it will seem that the alternatives to which economic preferences must be applied are such as to obliterate this distinction.

3.3.2. Let us for the moment accept the picture of an individual whose system of preferences is all-embracing and complete, i.e. who, for any two objects or rather for any two imagined events, possesses a clear intuition of preference.

More precisely we expect him, for any two alternative events which are put before him as possibilities, to be able to tell which of the two he prefers.

It is a very natural extension of this picture to permit such an individual to compare not only events, but even combinations of events with stated probabilities.[2]

By a combination of two events we mean this: Let the two events be denoted by B and C and use, for the sake of simplicity, the probability

[1] A good example of the wide variety of formal possibilities is given by the entirely different development of the theory of light, colors, and wave lengths. All these notions too became numerical, but in an entirely different way.

[2] Indeed this is necessary if he is engaged in economic activities which are explicitly dependent on probability. Cf. the example of agriculture in footnote 2 on p. 10.

50%-50%. Then the "combination" is the prospect of seeing B occur with a probability of 50% and (if B does not occur) C with the (remaining) probability of 50%. We stress that the two alternatives are mutually exclusive, so that no possibility of complementarity and the like exists. Also, that an absolute certainty of the occurrence of either B or C exists.

To restate our position. We expect the individual under consideration to possess a clear intuition whether he prefers the event A to the 50-50 combination of B or C, or conversely. It is clear that if he prefers A to B and also to C, then he will prefer it to the above combination as well; similarly, if he prefers B as well as C to A, then he will prefer the combination too. But if he should prefer A to, say B, but at the same time C to A, then any assertion about his preference of A against the combination contains fundamentally new information. Specifically: If he now prefers A to the 50-50 combination of B and C, this provides a plausible base for the numerical estimate that his preference of A over B is in excess of his preference of C over A.[1,2]

If this standpoint is accepted, then there is a criterion with which to compare the preference of C over A with the preference of A over B. It is well known that thereby utilities—or rather differences of utilities—become numerically measurable.

That the possibility of comparison between A, B, and C only to this extent is already sufficient for a numerical measurement of "distances" was first observed in economics by Pareto. Exactly the same argument has been made, however, by Euclid for the position of points on a line—in fact it is the very basis of his classical derivation of numerical distances.

The introduction of numerical measures can be achieved even more directly if use is made of all possible probabilities. Indeed: Consider three events, C, A, B, for which the order of the individual's preferences is the one stated. Let α be a real number between 0 and 1, such that A is exactly equally desirable with the combined event consisting of a chance of probability $1 - \alpha$ for B and the remaining chance of probability α for C. Then we suggest the use of α as a numerical estimate for the ratio of the preference of A over B to that of C over B.[3] An exact and exhaustive

[1] To give a simple example: Assume that an individual prefers the consumption of a glass of tea to that of a cup of coffee, and the cup of coffee to a glass of milk. If we now want to know whether the last preference—i.e., difference in utilities—exceeds the former, it suffices to place him in a situation where he must decide this: Does he prefer a cup of coffee to a glass the content of which will be determined by a 50%-50% chance device as tea or milk.

[2] Observe that we have only postulated an individual intuition which permits decision as to which of two "events" is preferable. But we have not directly postulated any intuitive estimate of the relative sizes of two preferences—i.e. in the subsequent terminology, of two differences of utilities.

This is important, since the former information ought to be obtainable in a reproducible way by mere "questioning."

[3] This offers a good opportunity for another illustrative example. The above technique permits a direct determination of the ratio q of the utility of possessing 1 unit of a certain good to the utility of possessing 2 units of the same good. The individual must

elaboration of these ideas requires the use of the axiomatic method. A simple treatment on this basis is indeed possible. We shall discuss it in 3.5-3.7.

3.3.3. To avoid misunderstandings let us state that the "events" which were used above as the substratum of preferences are conceived as future events so as to make all logically possible alternatives equally admissible. However, it would be an unnecessary complication, as far as our present objectives are concerned, to get entangled with the problems of the preferences between events in different periods of the future.[1] It seems, however, that such difficulties can be obviated by locating all "events" in which we are interested at one and the same, standardized, moment, preferably in the immediate future.

The above considerations are so vitally dependent upon the numerical concept of probability that a few words concerning the latter may be appropriate.

Probability has often been visualized as a subjective concept more or less in the nature of an estimation. Since we propose to use it in constructing an individual, numerical estimation of utility, the above view of probability would not serve our purpose. The simplest procedure is, therefore, to insist upon the alternative, perfectly well founded interpretation of probability as frequency in long runs. This gives directly the necessary numerical foothold.[2]

3.3.4. This procedure for a numerical measurement of the utilities of the individual depends, of course, upon the hypothesis of completeness in the system of individual preferences.[3] It is conceivable—and may even in a way be more realistic—to allow for cases where the individual is neither able to state which of two alternatives he prefers nor that they are equally desirable. In this case the treatment by indifference curves becomes impracticable too.[4]

How real this possibility is, both for individuals and for organizations, seems to be an extremely interesting question, but it is a question of fact. It certainly deserves further study. We shall reconsider it briefly in 3.7.2.

At any rate we hope we have shown that the treatment by indifference curves implies either too much or too little: if the preferences of the indi-

---

be given the choice of obtaining 1 unit with certainty or of playing the chance to get two units with the probability α, or nothing with the probability $1 - \alpha$. If he prefers the former, then $\alpha < q$; if he prefers the latter, then $\alpha > q$; if he cannot state a preference either way, then $\alpha = q$.

[1] It is well known that this presents very interesting, but as yet extremely obscure, connections with the theory of saving and interest, etc.

[2] If one objects to the frequency interpretation of probability then the two concepts (probability and preference) can be axiomatized together. This too leads to a satisfactory numerical concept of utility which will be discussed on another occasion.

[3] We have not obtained any basis for a comparison, quantitatively or qualitatively, of the utilities of different individuals.

[4] These problems belong systematically in the mathematical theory of ordered sets. The above question in particular amounts to asking whether events, with respect to preference, form a completely or a partially ordered set. Cf. 65.3.

vidual are not all comparable, then the indifference curves do not exist.[1] If the individual's preferences are all comparable, then we can even obtain a (uniquely defined) numerical utility which renders the indifference curves superfluous.

All this becomes, of course, pointless for the entrepreneur who can calculate in terms of (monetary) costs and profits.

3.3.5. The objection could be raised that it is not necessary to go into all these intricate details concerning the measurability of utility, since evidently the common individual, whose behavior one wants to describe, does not measure his utilities exactly but rather conducts his economic activities in a sphere of considerable haziness. The same is true, of course, for much of his conduct regarding light, heat, muscular effort, etc. But in order to build a science of physics these phenomena had to be measured. And subsequently the individual has come to use the results of such measurements—directly or indirectly—even in his everyday life. The same may obtain in economics at a future date. Once a fuller understanding of economic behavior has been achieved with the aid of a theory which makes use of this instrument, the life of the individual might be materially affected. It is, therefore, not an unnecessary digression to study these problems.

3.4. Principles of Measurement: Detailed Discussion

3.4.1. The reader may feel, on the basis of the foregoing, that we obtained a numerical scale of utility only by begging the principle, i.e. by really postulating the existence of such a scale. We have argued in 3.3.2. that if an individual prefers A to the 50-50 combination of B and C (while preferring C to A and A to B), this provides a plausible basis for the numerical estimate that this preference of A over B exceeds that of C over A. Are we not postulating here—or taking it for granted—that one preference may exceed another, i.e. that such statements convey a meaning? Such a view would be a complete misunderstanding of our procedure.

3.4.2. We are not postulating—or assuming—anything of the kind. We have assumed only one thing—and for this there is good empirical evidence—namely that imagined events can be combined with probabilities. And therefore the same must be assumed for the utilities attached to them,—whatever they may be. Or to put it in more mathematical language:

There frequently appear in science quantities which are *a priori* not mathematical, but attached to certain aspects of the physical world. Occasionally these quantities can be grouped together in domains within which certain natural, physically defined operations are possible. Thus the physically defined quantity of "mass" permits the operation of addition. The physico-geometrically defined quantity of "distance"[2] permits the same

[1] Points on the same indifference curve must be identified and are therefore no instances of incomparability.

[2] Let us, for the sake of the argument, view geometry as a physical discipline,—a sufficiently tenable viewpoint. By "geometry" we mean—equally for the sake of the argument—Euclidean geometry.

operation. On the other hand, the physico-geometrically defined quantity of "position" does not permit this operation,[1] but it permits the operation of forming the "center of gravity" of two positions.[2] Again other physico-geometrical concepts, usually styled "vectorial"—like velocity and acceleration—permit the operation of "addition."

3.4.3. In all these cases where such a "natural" operation is given a name which is reminiscent of a mathematical operation—like the instances of "addition" above—one must carefully avoid misunderstandings. This nomenclature is not intended as a claim that the two operations with the same name are identical,—this is manifestly not the case; it only expresses the opinion that they possess similar traits, and the hope that some correspondence between them will ultimately be established. This of course—when feasible at all—is done by finding a mathematical model for the physical domain in question, within which those quantities are defined by numbers, so that in the model the mathematical operation describes the synonymous "natural" operation.

To return to our examples: "energy" and "mass" became numbers in the pertinent mathematical models, "natural" addition becoming ordinary addition. "Position" as well as the vectorial quantities became triplets[3] of numbers, called coordinates or components respectively. The "natural" concept of "center of gravity" of two positions $\{x_1, x_2, x_3\}$ and $\{x'_1, x'_2, x'_3\}$,[4] with the "masses" α, $1 - \alpha$ (cf. footnote 2 above), becomes

$$\{\alpha x_1 + (1 - \alpha)x'_1,\ \alpha x_2 + (1 - \alpha)x'_2,\ \alpha x_3 + (1 - \alpha)x'_3\}.^5$$

The "natural" operation of "addition" of vectors $\{x_1, x_2, x_3\}$ and $\{x'_1, x'_2, x'_3\}$ becomes $\{x_1 + x'_1, x_2 + x'_2, x_3 + x'_3\}$.[6]

What was said above about "natural" and mathematical operations applies equally to natural and mathematical relations. The various concepts of "greater" which occur in physics—greater energy, force, heat, velocity, etc.—are good examples.

These "natural" relations are the best base upon which to construct mathematical models and to correlate the physical domain with them.[7,8]

[1] We are thinking of a "homogeneous" Euclidean space, in which no origin or frame of reference is preferred above any other.

[2] With respect to two given masses α, β occupying those positions. It may be convenient to normalize so that the total mass is the unit, i.e. $\beta = 1 - \alpha$.

[3] We are thinking of three-dimensional Euclidean space.

[4] We are now describing them by their three numerical coordinates.

[5] This is usually denoted by $\alpha\{x_1, x_2, x_3\} + (1 - \alpha)\{x'_1, x'_2, x'_3\}$. Cf. (16:A:c) in 16.2.1.

[6] This is usually denoted by $\{x_1, x_2, x_3\} + \{x'_1, x'_2, x'_3\}$. Cf. the beginning of 16.2.1.

[7] Not the only one. Temperature is a good counter-example. The "natural" relation of "greater" would not have sufficed to establish the present day mathematical model,—i.e. the absolute temperature scale. The devices actually used were different. Cf. 3.2.1.

[8] We do not want to give the misleading impression of attempting here a complete picture of the formation of mathematical models, i.e. of physical theories. It should be remembered that this is a very varied process with many unexpected phases. An important one is, e.g., the disentanglement of concepts: i.e. splitting up something which at

3.4.4. Here a further remark must be made. Assume that a satisfactory mathematical model for a physical domain in the above sense has been found, and that the physical quantities under consideration have been correlated with numbers. In this case it is not true necessarily that the description (of the mathematical model) provides for a *unique* way of correlating the physical quantities to numbers; i.e., it may specify an entire family of such correlations—the mathematical name is mappings—any one of which can be used for the purposes of the theory. Passage from one of these correlations to another amounts to a *transformation* of the numerical data describing the physical quantities. We then say that in this theory the physical quantities in question are described by numbers *up to* that system of transformations. The mathematical name of such transformation systems is *groups*.[1]

Examples of such situations are numerous. Thus the geometrical concept of distance is a number, up to multiplication by (positive) constant factors.[2] The situation concerning the physical quantity of mass is the same. The physical concept of energy is a number up to any linear transformation,—i.e. addition of any constant and multiplication by any (positive) constant.[3] The concept of position is defined up to an inhomogeneous orthogonal linear transformation.[4,5] The vectorial concepts are defined up to homogeneous transformations of the same kind.[5,6]

3.4.5. It is even conceivable that a physical quantity is a number up to any monotone transformation. This is the case for quantities for which only a "natural" relation "greater" exists—and nothing else. E.g. this was the case for temperature as long as only the concept of "warmer" was known;[7] it applies to the Mohs' scale of hardness of minerals; it applies to

---

superficial inspection seems to be one physical entity into several mathematical notions. Thus the "disentanglement" of force and energy, of quantity of heat and temperature, were decisive in their respective fields.

It is quite unforeseeable how many such differentiations still lie ahead in economic theory.

[1] We shall encounter groups in another context in 28.1.1, where references to the literature are also found.

[2] I.e. there is nothing in Euclidean geometry to fix a unit of distance.

[3] I.e. there is nothing in mechanics to fix a zero or a unit of energy. Cf. with footnote 2 above. Distance has a natural zero,—the distance of any point from itself.

[4] I.e. $\{x_1, x_2, x_3\}$ are to be replaced by $\{x_1^*, x_2^*, x_3^*\}$ where

$$\begin{aligned} x_1^* &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + b_1, \\ x_2^* &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + b_2, \\ x_3^* &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + b_3, \end{aligned}$$

the a_{ij}, b_i being constants, and the matrix (a_{ij}) what is known as orthogonal.

[5] I.e. there is nothing in geometry to fix either origin or the frame of reference when positions are concerned; and nothing to fix the frame of reference when vectors are concerned.

[6] I.e. the $b_i = 0$ in footnote 4 above. Sometimes a wider concept of matrices is permissible,—all those with determinants $\neq 0$. We need not discuss these matters here.

[7] But no quantitatively reproducible method of thermometry.

the notion of utility when this is based on the conventional idea of preference. In these cases one may be tempted to take the view that the quantity in question is not numerical at all, considering how arbitrary the description by numbers is. It seems to be preferable, however, to refrain from such qualitative statements and to state instead objectively up to what system of transformations the numerical description is determined. The case when the system consists of all monotone transformations is, of course, a rather extreme one; various graduations at the other end of the scale are the transformation systems mentioned above: inhomogeneous or homogeneous orthogonal linear transformations in space, linear transformations of one numerical variable, multiplication of that variable by a constant.[1] *In fine*, the case even occurs where the numerical description is absolutely rigorous, i.e. where no transformations at all need be tolerated.[2]

3.4.6. Given a physical quantity, the system of transformations up to which it is described by numbers may vary in time, i.e. with the stage of development of the subject. Thus temperature was originally a number only up to any monotone transformation.[3] With the development of thermometry—particularly of the concordant ideal gas thermometry—the transformations were restricted to the linear ones, i.e. only the absolute zero and the absolute unit were missing. Subsequent developments of thermodynamics even fixed the absolute zero so that the transformation system in thermodynamics consists only of the multiplication by constants. Examples could be multiplied but there seems to be no need to go into this subject further.

For utility the situation seems to be of a similar nature. One may take the attitude that the only "natural" datum in this domain is the relation "greater," i.e. the concept of preference. In this case utilities are numerical up to a monotone transformation. This is, indeed, the generally accepted standpoint in economic literature, best expressed in the technique of indifference curves.

To narrow the system of transformations it would be necessary to discover further "natural" operations or relations in the domain of utility. Thus it was pointed out by Pareto[4] that an equality relation for utility differences would suffice; in our terminology it would reduce the transformation system to the linear transformations.[5] However, since it does not

[1] One could also imagine intermediate cases of greater transformation systems than these but not containing all monotone transformations. Various forms of the theory of relativity give rather technical examples of this.

[2] In the usual language this would hold for physical quantities where an absolute zero as well as an absolute unit can be defined. This is, e.g., the case for the absolute value (not the vector!) of velocity in such physical theories as those in which light velocity plays a normative role: Maxwellian electrodynamics, special relativity.

[3] As long as only the concept of "warmer"—i.e. a "natural" relation "greater"—was known. We discussed this *in extenso* previously.

[4] *V. Pareto*, Manuel d'Economie Politique, Paris, 1907, p. 264.

[5] This is exactly what Euclid did for position on a line. The utility concept of "preference" corresponds to the relation of "lying to the right of" there, and the (desired) relation of the equality of utility differences to the geometrical congruence of intervals.

seem that this relation is really a "natural" one—i.e. one which can be interpreted by reproducible observations—the suggestion does not achieve the purpose.

3.5. Conceptual Structure of the Axiomatic Treatment of Numerical Utilities

3.5.1. The failure of one particular device need not exclude the possibility of achieving the same end by another device. Our contention is that the domain of utility contains a "natural" operation which narrows the system of transformations to precisely the same extent as the other device would have done. This is the combination of two utilities with two given alternative probabilities α, $1 - \alpha$, $(0 < \alpha < 1)$ as described in 3.3.2. The process is so similar to the formation of centers of gravity mentioned in 3.4.3. that it may be advantageous to use the same terminology. Thus we have for utilities u, v the "natural" *relation* $u > v$ (read: u is preferable to v), and the "natural" *operation* $\alpha u + (1 - \alpha)v$, $(0 < \alpha < 1)$, (read: center of gravity of u, v with the respective weights α, $1 - \alpha$; or: combination of u, v with the alternative probabilities α, $1 - \alpha$). If the existence—and reproducible observability —of these concepts is conceded, then our way is clear: We must find a correspondence between utilities and numbers which carries the relation $u > v$ and the operation $\alpha u + (1 - \alpha)v$ for utilities into the synonymous concepts for numbers.

Denote the correspondence by

$$u \to \rho = \mathrm{v}(u),$$

u being the utility and $\mathrm{v}(u)$ the number which the correspondence attaches to it. Our requirements are then:

(3:1:a) $$u > v \quad \text{implies} \quad \mathrm{v}(u) > \mathrm{v}(v),$$

(3:1:b) $$\mathrm{v}(\alpha u + (1 - \alpha)v) = \alpha\mathrm{v}(u) + (1 - \alpha)\mathrm{v}(v).$$ [1]

If two such correspondences

(3:2:a) $$u \to \rho = \mathrm{v}(u),$$

(3:2:b) $$u \to \rho' = \mathrm{v}'(u),$$

should exist, then they set up a correspondence between numbers

(3:3) $$\rho \leftrightarrows \rho',$$

for which we may also write

(3:4) $$\rho' = \phi(\rho).$$

Since (3:2:a), (3:2:b) fulfill (3:1:a), (3:1:b), the correspondence (3:3), i.e. the function $\phi(\rho)$ in (3:4) must leave the relation $\rho > \sigma$ [2] and the operation

[1] Observe that in in each case the left-hand side has the "natural" concepts for utilities, and the right-hand side the conventional ones for numbers.

[2] Now these are applied to numbers ρ, σ!

$\alpha\rho + (1 - \alpha)\sigma$ unaffected (cf footnote 1 on p. 24). I.e.

(3:5:a) $\rho > \sigma$ implies $\phi(\rho) > \phi(\sigma)$,

(3:5:b) $\phi(\alpha\rho + (1 - \alpha)\sigma) = \alpha\phi(\rho) + (1 - \alpha)\phi(\sigma)$.

Hence $\phi(\rho)$ must be a linear function, i.e.

$$\rho' = \phi(\rho) \equiv \omega_0\rho + \omega_1, \tag{3:6}$$

where ω_0, ω_1 are fixed numbers (constants) with $\omega_0 > 0$.

So we see: If such a numerical valuation of utilities[1] exists at all, then it is determined up to a linear transformation.[2,3] I.e. then utility is a number up to a linear transformation.

In order that a numerical valuation in the above sense should exist it is necessary to postulate certain properties of the relation $u > v$ and the operation $\alpha u + (1 - \alpha)v$ for utilities. The selection of these postulates or axioms and their subsequent analysis leads to problems of a certain mathematical interest. In what follows we give a general outline of the situation for the orientation of the reader; a complete discussion is found in the Appendix.

3.5.2. A choice of axioms is not a purely objective task. It is usually expected to achieve some definite aim—some specific theorem or theorems are to be derivable from the axioms—and to this extent the problem is exact and objective. But beyond this there are always other important desiderata of a less exact nature: The axioms should not be too numerous, their system is to be as simple and transparent as possible, and each axiom should have an immediate intuitive meaning by which its appropriateness may be judged directly.[4] In a situation like ours this last requirement is particularly vital, in spite of its vagueness: we want to make an intuitive concept amenable to mathematical treatment and to see as clearly as possible what hypotheses this requires.

The objective part of our problem is clear: the postulates must imply the existence of a correspondence (3:2:a) with the properties (3:1:a), (3:1:b) as described in 3.5.1. The further heuristic, and even esthetic desiderata, indicated above, do not determine a unique way of finding this axiomatic treatment. In what follows we shall formulate a set of axioms which seems to be essentially satisfactory.

[1] I.e. a correspondence (3:2:a) which fulfills (3:1:a), (3:1:b).

[2] I.e. one of the form (3:6).

[3] Remember the physical examples of the same situation given in 3.4.4. (Our present discussion is somewhat more detailed.) We do not undertake to fix an absolute zero and an absolute unit of utility.

[4] The first and the last principle may represent—at least to a certain extent—opposite influences: If we reduce the number of axioms by merging them as far as technically possible, we may lose the possibility of distinguishing the various intuitive backgrounds. Thus we could have expressed the group (3:B) in 3.6.1. by a smaller number of axioms, but this would have obscured the subsequent analysis of 3.6.2.

To strike a proper balance is a matter of practical—and to some extent even esthetic—judgment.

3.6. The Axioms and Their Interpretation

3.6.1. Our axioms are these:

We consider a system U of entities[1] $u, v, w, \cdots$. In U a *relation* is given, $u > v$, and for any number α, $(0 < \alpha < 1)$, an *operation*

$$\alpha u + (1 - \alpha)v = w.$$

These concepts satisfy the following axioms:

(3:A) *$u > v$ is a complete ordering of U.*[2]

This means: Write $u < v$ when $v > u$. Then:

(3:A:a) For any two u, v one and only one of the three following relations holds:

$$u = v, \qquad u > v, \qquad u < v.$$

(3:A:b) $u > v$, $v > w$ imply $u > w$.[3]

(3:B) *Ordering and combining.*[4]

(3:B:a) $u < v$ implies that $u < \alpha u + (1 - \alpha)v$.

(3:B:b) $u > v$ implies that $u > \alpha u + (1 - \alpha)v$.

(3:B:c) $u < w < v$ implies the existence of an α with

$$\alpha u + (1 - \alpha)v < w.$$

(3:B:d) $u > w > v$ implies the existence of an α with

$$\alpha u + (1 - \alpha)v > w.$$

(3:C) *Algebra of combining.*

(3:C:a) $\alpha u + (1 - \alpha)v = (1 - \alpha)v + \alpha u.$

(3:C:b) $\alpha(\beta u + (1 - \beta)v) + (1 - \alpha)v = \gamma u + (1 - \gamma)v$ where $\gamma = \alpha\beta$.

One can show that these axioms imply the existence of a correspondence (3:2:a) with the properties (3:1:a), (3:1:b) as described in 3.5.1. Hence the conclusions of 3.5.1. hold good: The system U—i.e. in our present interpretation, the system of (abstract) utilities—is one of numbers up to a linear transformation.

The construction of (3:2:a) (with (3:1:a), (3:1:b) by means of the axioms (3:A)-(3:C)) is a purely mathematical task which is somewhat lengthy, although it runs along conventional lines and presents no par-

[1] This is, of course, meant to be the system of (abstract) utilities, to be characterized by our axioms. Concerning the general nature of the axiomatic method, cf. the remarks and references in the last part of 10.1.1.

[2] For a more systematic mathematical discussion of this notion, cf. 65.3.1. The equivalent concept of the completeness of the system of preferences was previously considered at the beginning of 3.3.2. and of 3.4.6.

[3] These conditions (3:A:a), (3:A:b) correspond to (65:A:a), (65:A:b) in 65.3.1.

[4] Remember that the α, β, γ occurring here are always > 0, < 1.

ticular difficulties. (Cf. Appendix.)

It seems equally unnecessary to carry out the usual logistic discussion of these axioms[1] on this occasion.

We shall however say a few more words about the intuitive meaning—i.e. the justification—of each one of our axioms (3:A)-(3:C).

3.6.2. The analysis of our postulates follows:

(3:A:a*) This is the statement of the completeness of the system of individual preferences. It is customary to assume this when discussing utilities or preferences, e.g. in the "indifference curve analysis method." These questions were already considered in 3.3.4. and 3.4.6.

(3:A:b*) This is the "transitivity" of preference, a plausible and generally accepted property.

(3:B:a*) We state here: If v is preferable to u, then even a chance $1 - \alpha$ of v—alternatively to u—is preferable. This is legitimate since any kind of complementarity (or the opposite) has been excluded, cf. the beginning of 3.3.2.

(3:B:b*) This is the dual of (3:B:a*), with "less preferable" in place of "preferable."

(3:B:c*) We state here: If w is preferable to u, and an even more preferable v is also given, then the combination of u with a chance $1 - \alpha$ of v will not affect w's preferability to it if this chance is small enough. I.e.: However desirable v may be in itself, one can make its influence as weak as desired by giving it a sufficiently small chance. This is a plausible "continuity" assumption.

(3:B:d*) This is the dual of (3:B:c*), with "less preferable" in place of "preferable."

(3:C:a*) This is the statement that it is irrelevant in which order the constituents u, v of a combination are named. It is legitimate, particularly since the constituents are alternative events, cf. (3:B:a*) above.

(3:C:b*) This is the statement that it is irrelevant whether a combination of two constituents is obtained in two successive steps,—first the probabilities α, $1 - \alpha$, then the probabilities β, $1 - \beta$; or in one operation,—the probabilities γ, $1 - \gamma$ where $\gamma = \alpha\beta$.[2] The same things can be said for this as for (3:C:a*) above. It may be, however, that this postulate has a deeper significance, to which one allusion is made in 3.7.1. below.

[1] A similar situation is dealt with more exhaustively in 10.; those axioms describe a subject which is more vital for our main objective. The logistic discussion is indicated there in 10.2. Some of the general remarks of 10.3. apply to the present case also.

[2] This is of course the correct arithmetic of accounting for two successive admixtures of v with u.

3.7. General Remarks Concerning the Axioms

3.7.1. At this point it may be well to stop and to reconsider the situation. Have we not shown too much? We can derive from the postulates (3:A)-(3:C) the numerical character of utility in the sense of (3:2:a) and (3:1:a), (3:1:b) in 3.5.1.; and (3:1:b) states that the numerical values of utility combine (with probabilities) like mathematical expectations! And yet the concept of mathematical expectation has been often questioned, and its legitimateness is certainly dependent upon some hypothesis concerning the nature of an "expectation."[1] Have we not then begged the question? Do not our postulates introduce, in some oblique way, the hypotheses which bring in the mathematical expectation?

More specifically: May there not exist in an individual a (positive or negative) utility of the mere act of "taking a chance," of gambling, which the use of the mathematical expectation obliterates?

How did our axioms (3:A)-(3:C) get around this possibility?

As far as we can see, our postulates (3:A)-(3:C) do not attempt to avoid it. Even that one which gets closest to excluding a "utility of gambling" (3:C:b) (cf. its discussion in 3.6.2.), seems to be plausible and legitimate,—unless a much more refined system of psychology is used than the one now available for the purposes of economics. The fact that a numerical utility—with a formula amounting to the use of mathematical expectations—can be built upon (3:A)-(3:C), seems to indicate this: We have practically defined numerical utility as being that thing for which the calculus of mathematical expectations is legitimate.[2] Since (3:A)-(3:C) secure that the necessary construction can be carried out, concepts like a "specific utility of gambling" cannot be formulated free of contradiction on this level.[3]

3.7.2. As we have stated, the last time in 3.6.1., our axioms are based on the relation $u > v$ and on the operation $\alpha u + (1 - \alpha)v$ for utilities. It seems noteworthy that the latter may be regarded as more immediately given than the former: One can hardly doubt that anybody who could imagine two alternative situations with the respective utilities u, v could not also conceive the prospect of having both with the given respective probabilities α, $1 - \alpha$. On the other hand one may question the postulate of axiom (3:A:a) for $u > v$, i.e. the completeness of this ordering.

Let us consider this point for a moment. We have conceded that one may doubt whether a person can always decide which of two alternatives—

[1] Cf. *Karl Menger:* Das Unsicherheitsmoment in der Wertlehre, Zeitschrift für Nationalökonomie, vol. 5, (1934) pp. 459ff. and *Gerhard Tintner:* A contribution to the non-static Theory of Choice, Quarterly Journal of Economics, vol. LVI, (1942) pp. 274ff.

[2] Thus Daniel Bernoulli's well known suggestion to "solve" the "St. Petersburg Paradox" by the use of the so-called "moral expectation" (instead of the mathematical expectation) means defining the utility numerically as the logarithm of one's monetary possessions.

[3] This may seem to be a paradoxical assertion. But anybody who has seriously tried to axiomatize that elusive concept, will probably concur with it.

with the utilities u, v—he prefers.[1] But, whatever the merits of this doubt are, this possibility—i.e. the completeness of the system of (individual) preferences—must be assumed even for the purposes of the "indifference curve method" (cf. our remarks on (3:A:a) in 3.6.2.). But if this property of $u > v$ [2] is assumed, then our use of the much less questionable $\alpha u + (1 - \alpha)v$ [3] yields the numerical utilities too![4]

If the general comparability assumption is not made,[5] a mathematical theory—based on $\alpha u + (1 - \alpha)v$ together with what remains of $u > v$—is still possible.[6] It leads to what may be described as a many-dimensional vector concept of utility. This is a more complicated and less satisfactory set-up, but we do not propose to treat it systematically at this time.

3.7.3. This brief exposition does not claim to exhaust the subject, but we hope to have conveyed the essential points. To avoid misunderstandings, the following further remarks may be useful.

(1) We re-emphasize that we are considering only utilities experienced by one person. These considerations do not imply anything concerning the comparisons of the utilities belonging to different individuals.

(2) It cannot be denied that the analysis of the methods which make use of mathematical expectation (cf. footnote 1 on p. 28 for the literature) is far from concluded at present. Our remarks in 3.7.1. lie in this direction, but much more should be said in this respect. There are many interesting questions involved, which however lie beyond the scope of this work. For our purposes it suffices to observe that the validity of the simple and plausible axioms (3:A)-(3:C) in 3.6.1. for the relation $u > v$ and the operation $\alpha u + (1 - \alpha)v$ makes the utilities numbers up to a linear transformation in the sense discussed in these sections.

3.8. The Role of the Concept of Marginal Utility

3.8.1. The preceding analysis made it clear that we feel free to make use of a numerical conception of utility. On the other hand, subsequent

[1] Or that he can assert that they are precisely equally desirable.

[2] I.e. the completeness postulate (3:A:a).

[3] I.e. the postulates (3:B), (3:C) together with the obvious postulate (3:A:b).

[4] At this point the reader may recall the familiar argument according to which the unnumerical ("indifference curve") treatment of utilities is preferable to any numerical one, because it is simpler and based on fewer hypotheses. This objection might be legitimate if the numerical treatment were based on Pareto's equality relation for utility differences (cf. the end of 3.4.6.). This relation is, indeed, a stronger and more complicated hypothesis, added to the original ones concerning the general comparability of utilities (completeness of preferences).

However, we used the operation $\alpha u + (1 - \alpha)v$ instead, and we hope that the reader will agree with us that it represents an even safer assumption than that of the completeness of preferences.

We think therefore that our procedure, as distinguished from Pareto's, is not open to the objections based on the necessity of artificial assumptions and a loss of simplicity.

[5] This amounts to weakening (3:A:a) to an (3:A:a') by replacing in it "one and only one" by "at most one." The conditions (3:A:a'), (3:A:b) then correspond to (65:B:a), (65:B:b).

[6] In this case some modifications in the groups of postulates (3:B), (3:C) are also necessary.

discussions will show that we cannot avoid the assumption that all subjects of the economy under consideration are completely informed about the physical characteristics of the situation in which they operate and are able to perform all statistical, mathematical, etc., operations which this knowledge makes possible. The nature and importance of this assumption has been given extensive attention in the literature and the subject is probably very far from being exhausted. We propose not to enter upon it. The question is too vast and too difficult and we believe that it is best to "divide difficulties." I.e. we wish to avoid this complication which, while interesting in its own right, should be considered separately from our present problem.

Actually we think that our investigations—although they assume "complete information" without any further discussion—do make a contribution to the study of this subject. It will be seen that many economic and social phenomena which are usually ascribed to the individual's state of "incomplete information" make their appearance in our theory and can be satisfactorily interpreted with its help. Since our theory assumes "complete information," we conclude from this that those phenomena have nothing to do with the individual's "incomplete information." Some particularly striking examples of this will be found in the concepts of "discrimination" in 33.1., of "incomplete exploitation" in 38.3., and of the "transfer" or "tribute" in 46.11., 46.12.

On the basis of the above we would even venture to question the importance usually ascribed to incomplete information in its conventional sense[1] in economic and social theory. It will appear that some phenomena which would *prima facie* have to be attributed to this factor, have nothing to do with it.[2]

3.8.2. Let us now consider an isolated individual with definite physical characteristics and with definite quantities of goods at his disposal. In view of what was said above, he is in a position to determine the maximum utility which can be obtained in this situation. Since the maximum is a well-defined quantity, the same is true for the increase which occurs when a unit of any definite good is added to the stock of all goods in the possession of the individual. This is, of course, the classical notion of the marginal utility of a unit of the commodity in question.[3]

These quantities are clearly of decisive importance in the "Robinson Crusoe" economy. The above marginal utility obviously corresponds to

[1] We shall see that the rules of the games considered may explicitly prescribe that certain participants should not possess certain pieces of information. Cf. 6.3., 6.4. (Games in which this does not happen are referred to in 14.8. and in (15:B) of 15.3.2., and are called games with "perfect information.") We shall recognize and utilize this kind of "incomplete information" (according to the above, rather to be called "imperfect information"). But we reject all other types, vaguely defined by the use of concepts like complication, intelligence, etc.

[2] Our theory attributes these phenomena to the possibility of multiple "stable standards of behavior" cf. 4.6. and the end of 4.7.

[3] More precisely: the so-called "indirectly dependent expected utility."

the maximum effort which he will be willing to make—if he behaves according to the customary criteria of rationality—in order to obtain a further unit of that commodity.

It is not clear at all, however, what significance it has in determining the behavior of a participant in a social exchange economy. We saw that the principles of rational behavior in this case still await formulation, and that they are certainly not expressed by a maximum requirement of the Crusoe type. Thus it must be uncertain whether marginal utility has any meaning at all in this case.[1]

Positive statements on this subject will be possible only after we have succeeded in developing a theory of rational behavior in a social exchange economy,—that is, as was stated before, with the help of the theory of "games of strategy." It will be seen that marginal utility does, indeed, play an important role in this case too, but in a more subtle way than is usually assumed.

4. Structure of the Theory: Solutions and Standards of Behavior

4.1. The Simplest Concept of a Solution for One Participant

4.1.1. We have now reached the point where it becomes possible to give a positive description of our proposed procedure. This means primarily an outline and an account of the main technical concepts and devices.

As we stated before, we wish to find the mathematically complete principles which define "rational behavior" for the participants in a social economy, and to derive from them the general characteristics of that behavior. And while the principles ought to be perfectly general—i.e., valid in all situations—we may be satisfied if we can find solutions, for the moment, only in some characteristic special cases.

First of all we must obtain a clear notion of what can be accepted as a solution of this problem; i.e., what the amount of information is which a solution must convey, and what we should expect regarding its formal structure. A precise analysis becomes possible only after these matters have been clarified.

4.1.2. The immediate concept of a solution is plausibly a set of rules for each participant which tell him how to behave in every situation which may conceivably arise. One may object at this point that this view is unnecessarily inclusive. Since we want to theorize about "rational behavior," there seems to be no need to give the individual advice as to his behavior in situations other than those which arise in a rational community. This would justify assuming rational behavior on the part of the others as well,—in whatever way we are going to characterize that. Such a procedure would probably lead to a unique sequence of situations to which alone our theory need refer.

[1] All this is understood within the domain of our several simplifying assumptions. If they are relaxed, then various further difficulties ensue.

This objection seems to be invalid for two reasons:

First, the "rules of the game,"—i.e. the physical laws which give the factual background of the economic activities under consideration may be explicitly statistical. The actions of the participants of the economy may determine the outcome only in conjunction with events which depend on chance (with known probabilities), cf. footnote 2 on p. 10 and 6.2.1. If this is taken into consideration, then the rules of behavior even in a perfectly rational community must provide for a great variety of situations—some of which will be very far from optimum.[1]

Second, and this is even more fundamental, the rules of rational behavior must provide definitely for the possibility of irrational conduct on the part of others. In other words: Imagine that we have discovered a set of rules for all participants—to be termed as "optimal" or "rational"—each of which is indeed optimal provided that the other participants conform. Then the question remains as to what will happen if some of the participants do not conform. If that should turn out to be advantageous for them—and, quite particularly, disadvantageous to the conformists—then the above "solution" would seem very questionable. We are in no position to give a positive discussion of these things as yet—but we want to make it clear that under such conditions the "solution," or at least its motivation, must be considered as imperfect and incomplete. In whatever way we formulate the guiding principles and the objective justification of "rational behavior," provisos will have to be made for every possible conduct of "the others." Only in this way can a satisfactory and exhaustive theory be developed. But if the superiority of "rational behavior" over any other kind is to be established, then its description must include rules of conduct for all conceivable situations—including those where "the others" behaved irrationally, in the sense of the standards which the theory will set for them.

4.1.3. At this stage the reader will observe a great similarity with the everyday concept of games. We think that this similarity is very essential; indeed, that it is more than that. For economic and social problems the games fulfill—or should fulfill—the same function which various geometrico-mathematical models have successfully performed in the physical sciences. Such models are theoretical constructs with a precise, exhaustive and not too complicated definition; and they must be similar to reality in those respects which are essential in the investigation at hand. To recapitulate in detail: The definition must be precise and exhaustive in order to make a mathematical treatment possible. The construct must not be unduly complicated, so that the mathematical treatment can be brought beyond the mere formalism to the point where it yields complete numerical results. Similarity to reality is needed to make the operation significant. And this similarity must usually be restricted to a few traits

[1] That a unique optimal behavior is at all conceivable in spite of the multiplicity of the possibilities determined by chance, is of course due to the use of the notion of "mathematical expectation." Cf. loc. cit. above.

deemed "essential" *pro tempore*—since otherwise the above requirements would conflict with each other.[1]

It is clear that if a model of economic activities is constructed according to these principles, the description of a game results. This is particularly striking in the formal description of markets which are after all the core of the economic system—but this statement is true in all cases and without qualifications.

4.1.4. We described in 4.1.2. what we expect a solution—i.e. a characterization of "rational behavior"—to consist of. This amounted to a complete set of rules of behavior in all conceivable situations. This holds equivalently for a social economy and for games. The entire result in the above sense is thus a combinatorial enumeration of enormous complexity. But we have accepted a simplified concept of utility according to which all the individual strives for is fully described by one numerical datum (cf. 2.1.1. and 3.3.). Thus the complicated combinatorial catalogue—which we expect from a solution—permits a very brief and significant summarization: the statement of how much[2,3] the participant under consideration can get if he behaves "rationally." This "can get" is, of course, presumed to be a minimum; he may get more if the others make mistakes (behave irrationally).

It ought to be understood that all this discussion is advanced, as it should be, preliminary to the building of a satisfactory theory along the lines indicated. We formulate desiderata which will serve as a gauge of success in our subsequent considerations; but it is in accordance with the usual heuristic procedure to reason about these desiderata—even before we are able to satisfy them. Indeed, this preliminary reasoning is an essential part of the process of finding a satisfactory theory.[4]

4.2. Extension to All Participants

4.2.1. We have considered so far only what the solution ought to be for one participant. Let us now visualize all participants simultaneously. I.e., let us consider a social economy, or equivalently a game of a fixed number of (say n) participants. The complete information which a solution should convey is, as we discussed it, of a combinatorial nature. It was indicated furthermore how a single quantitative statement contains the decisive part of this information, by stating how much each participant

[1] E.g., Newton's description of the solar system by a small number of "masspoints." These points attract each other and move like the stars; this is the similarity in the essentials, while the enormous wealth of the other physical features of the planets has been left out of account.

[2] Utility; for an entrepreneur,—profit; for a player,—gain or loss.

[3] We mean, of course, the "mathematical expectation," if there is an explicit element of chance. Cf. the first remark in 4.1.2. and also the discussion of 3.7.1.

[4] Those who are familiar with the development of physics will know how important such heuristic considerations can be. Neither general relativity nor quantum mechanics could have been found without a "pre-theoretical" discussion of the desiderata concerning the theory-to-be.

obtains by behaving rationally. Consider these amounts which the several participants "obtain." If the solution did nothing more in the quantitative sense than specify these amounts,[1] then it would coincide with the well known concept of imputation: it would just state how the total proceeds are to be distributed among the participants.[2]

We emphasize that the problem of imputation must be solved both when the total proceeds are in fact identically zero and when they are variable. This problem, in its general form, has neither been properly formulated nor solved in economic literature.

4.2.2. We can see no reason why one should not be satisfied with a solution of this nature, providing it can be found: i.e. a single imputation which meets reasonable requirements for optimum (rational) behavior. (Of course we have not yet formulated these requirements. For an exhaustive discussion, cf. loc. cit. below.) The structure of the society under consideration would then be extremely simple: There would exist an absolute state of equilibrium in which the quantitative share of every participant would be precisely determined.

It will be seen however that such a solution, possessing all necessary properties, does not exist in general. The notion of a solution will have to be broadened considerably, and it will be seen that this is closely connected with certain inherent features of social organization that are well known from a "common sense" point of view but thus far have not been viewed in proper perspective. (Cf. 4.6. and 4.8.1.)

4.2.3. Our mathematical analysis of the problem will show that there exists, indeed, a not inconsiderable family of games where a solution can be defined and found in the above sense: i.e. as one single imputation. In such cases every participant obtains at least the amount thus imputed to him by just behaving appropriately, rationally. Indeed, he gets exactly this amount if the other participants too behave rationally; if they do not, he may get even more.

These are the games of two participants where the sum of all payments is zero. While these games are not exactly typical for major economic processes, they contain some universally important traits of all games and the results derived from them are the basis of the general theory of games. We shall discuss them at length in Chapter III.

4.3. The Solution as a Set of Imputations

4.3.1. If either of the two above restrictions is dropped, the situation is altered materially.

[1] And of course, in the combinatorial sense, as outlined above, the procedure how to obtain them.

[2] In games—as usually understood—the total proceeds are always zero; i.e. one participant can gain only what the others lose. Thus there is a pure problem of distribution—i.e. imputation—and absolutely none of increasing the total utility, the "social product." In all economic questions the latter problem arises as well, but the question of imputation remains. Subsequently we shall broaden the concept of a game by dropping the requirement of the total proceeds being zero (cf. Ch. XI).

The simplest game where the second requirement is overstepped is a two-person game where the sum of all payments is variable. This corresponds to a social economy with two participants and allows both for their interdependence and for variability of total utility with their behavior.[1] As a matter of fact this is exactly the case of a bilateral monopoly (cf. 61.2.-61.6.). The well known "zone of uncertainty" which is found in current efforts to solve the problem of imputation indicates that a broader concept of solution must be sought. This case will be discussed loc. cit. above. For the moment we want to use it only as an indicator of the difficulty and pass to the other case which is more suitable as a basis for a first positive step.

4.3.2. The simplest game where the first requirement is disregarded is a three-person game where the sum of all payments is zero. In contrast to the above two-person game, this does not correspond to any fundamental economic problem but it represents nevertheless a basic possibility in human relations. The essential feature is that any two players who combine and cooperate against a third can thereby secure an advantage. The problem is how this advantage should be distributed among the two partners in this combination. Any such scheme of imputation will have to take into account that any two partners can combine; i.e. while any one combination is in the process of formation, each partner must consider the fact that his prospective ally could break away and join the third participant.

Of course the rules of the game will prescribe how the proceeds of a coalition should be divided between the partners. But the detailed discussion to be given in 22.1. shows that this will not be, in general, the final verdict. Imagine a game (of three or more persons) in which two participants can form a very advantageous coalition but where the rules of the game provide that the greatest part of the gain goes to the first participant. Assume furthermore that the second participant of this coalition can also enter a coalition with the third one, which is less effective *in toto* but promises him a greater individual gain than the former. In this situation it is obviously reasonable for the first participant to transfer a part of the gains which he could get from the first coalition to the second participant in order to save this coalition. In other words: One must expect that under certain conditions one participant of a coalition will be willing to pay a compensation to his partner. Thus the apportionment within a coalition depends not only upon the rules of the game but also upon the above principles, under the influence of the alternative coalitions.[2]

Common sense suggests that one cannot expect any theoretical statement as to which alliance will be formed[3] but only information concerning

[1] It will be remembered that we make use of a transferable utility, cf. 2.1.1.

[2] This does not mean that the rules of the game are violated, since such compensatory payments, if made at all, are made freely in pursuance of a rational consideration.

[3] Obviously three combinations of two partners each are possible. In the example to be given in 21., any preference within the solution for a particular alliance will be a

how the partners in a possible combination must divide the spoils in order to avoid the contingency that any one of them deserts to form a combination with the third player. All this will be discussed in detail and quantitatively in Ch. V.

It suffices to state here only the result which the above qualitative considerations make plausible and which will be established more rigorously loc. cit. A reasonable concept of a solution consists in this case of a system of three imputations. These correspond to the above-mentioned three combinations or alliances and express the division of spoils between respective allies.

4.3.3. The last result will turn out to be the prototype of the general situation. We shall see that a consistent theory will result from looking for solutions which are not single imputations, but rather systems of imputations.

It is clear that in the above three-person game no single imputation from the solution is in itself anything like a solution. Any particular alliance describes only one particular consideration which enters the minds of the participants when they plan their behavior. Even if a particular alliance is ultimately formed, the division of the proceeds between the allies will be decisively influenced by the other alliances which each one might alternatively have entered. Thus only the three alliances and their imputations together form a rational whole which determines all of its details and possesses a stability of its own. It is, indeed, this whole which is the really significant entity, more so than its constituent imputations. Even if one of these is actually applied, i.e. if one particular alliance is actually formed, the others are present in a "virtual" existence: Although they have not materialized, they have contributed essentially to shaping and determining the actual reality.

In conceiving of the general problem, a social economy or equivalently a game of n participants, we shall—with an optimism which can be justified only by subsequent success—expect the same thing: A solution should be a system of imputations[1] possessing in its entirety some kind of balance and stability the nature of which we shall try to determine. We emphasize that this stability—whatever it may turn out to be—will be a property of the system as a whole and not of the single imputations of which it is composed. These brief considerations regarding the three-person game have illustrated this point.

4.3.4. The exact criteria which characterize a system of imputations as a solution of our problem are, of course, of a mathematical nature. For a precise and exhaustive discussion we must therefore refer the reader to the subsequent mathematical development of the theory. The exact definition

limine excluded by symmetry. I.e. the game will be symmetric with respect to all three participants. Cf. however 33.1.1.

[1] They may again include compensations between partners in a coalition, as described in 4.3.2.

itself is stated in 30.1.1. We shall nevertheless undertake to give a preliminary, qualitative outline. We hope this will contribute to the understanding of the ideas on which the quantitative discussion is based. Besides, the place of our considerations in the general framework of social theory will become clearer.

4.4. The Intransitive Notion of "Superiority" or "Domination"

4.4.1. Let us return to a more primitive concept of the solution which we know already must be abandoned. We mean the idea of a solution as a single imputation. If this sort of solution existed it would have to be an imputation which in some plausible sense was superior to all other imputations. This notion of superiority as between imputations ought to be formulated in a way which takes account of the physical and social structure of the milieu. That is, one should define that an imputation x is superior to an imputation y whenever this happens: Assume that society, i.e. the totality of all participants, has to consider the question whether or not to "accept" a static settlement of all questions of distribution by the imputation y. Assume furthermore that at this moment the alternative settlement by the imputation x is also considered. Then this alternative x will suffice to exclude acceptance of y. By this we mean that a sufficient number of participants prefer in their own interest x to y, and are convinced or can be convinced of the possibility of obtaining the advantages of x. In this comparison of x to y the participants should not be influenced by the consideration of any third alternatives (imputations). I.e. we conceive the relationship of superiority as an elementary one, correlating the two imputations x and y only. The further comparison of three or more—ultimately of all—imputations is the subject of the theory which must now follow, as a superstructure erected upon the elementary concept of superiority.

Whether the possibility of obtaining certain advantages by relinquishing y for x, as discussed in the above definition, can be made convincing to the interested parties will depend upon the physical facts of the situation—in the terminology of games, on the rules of the game.

We prefer to use, instead of "superior" with its manifold associations, a word more in the nature of a *terminus technicus*. When the above described relationship between two imputations x and y exists,[1] then we shall say that x *dominates* y.

If one restates a little more carefully what should be expected from a solution consisting of a single imputation, this formulation obtains: Such an imputation should dominate all others and be dominated by none.

4.4.2. The notion of domination as formulated—or rather indicated—above is clearly in the nature of an ordering, similar to the question of

[1] That is, when it holds in the mathematically precise form, which will be given in 30.1.1.

preference, or of size in any quantitative theory. The notion of a single imputation solution[1] corresponds to that of the first element with respect to that ordering.[2]

The search for such a first element would be a plausible one if the ordering in question, i.e. our notion of domination, possessed the important property of transitivity; that is, if it were true that whenever x dominates y and y dominates z, then also x dominates z. In this case one might proceed as follows: Starting with an arbitrary x, look for a y which dominates x; if such a y exists, choose one and look for a z which dominates y; if such a z exists, choose one and look for a u which dominates z, etc. In most practical problems there is a fair chance that this process either terminates after a finite number of steps with a w which is undominated by anything else, or that the sequence $x, y, z, u, \cdots$, goes on *ad infinitum*, but that these $x, y, z, u, \cdots$ tend to a limiting position w undominated by anything else. And, due to the transitivity referred to above, the final w will in either case dominate all previously obtained $x, y, z, u, \cdots$.

We shall not go into more elaborate details which could and should be given in an exhaustive discussion. It will probably be clear to the reader that the progress through the sequence $x, y, z, u, \cdots$ corresponds to successive "improvements" culminating in the "optimum," i.e. the "first" element w which dominates all others and is not dominated.

All this becomes very different when transitivity does not prevail. In that case any attempt to reach an "optimum" by successive improvements may be futile. It can happen that x is dominated by y, y by z, and z in turn by x.[3]

4.4.3. Now the notion of domination on which we rely is, indeed, not transitive. In our tentative description of this concept we indicated that x dominates y when there exists a group of participants each one of whom prefers his individual situation in x to that in y, and who are convinced that they are able as a group—i.e. as an alliance—to enforce their preferences. We shall discuss these matters in detail in 30.2. This group of participants shall be called the "effective set" for the domination of x over y. Now when x dominates y and y dominates z, the effective sets for these two dominations may be entirely disjunct and therefore no conclusions can be drawn concerning the relationship between z and x. It can even happen that z *dominates* x with the help of a third effective set, possibly disjunct from both previous ones.

[1] We continue to use it as an illustration although we have shown already that it is a forlorn hope. The reason for this is that, by showing what is involved if certain complications did not arise, we can put these complications into better perspective. Our real interest at this stage lies of course in these complications, which are quite fundamental.

[2] The mathematical theory of ordering is very simple and leads probably to a deeper understanding of these conditions than any purely verbal discussion. The necessary mathematical considerations will be found in 65.3.

[3] In the case of transitivity this is impossible because—if a proof be wanted—x never dominates itself. Indeed, if e.g. y dominates x, z dominates y, and x dominates z, then we can infer by transitivity that x dominates x.

This lack of transitivity, especially in the above formalistic presentation, may appear to be an annoying complication and it may even seem desirable to make an effort to rid the theory of it. Yet the reader who takes another look at the last paragraph will notice that it really contains only a circumlocution of a most typical phenomenon in all social organizations. The domination relationships between various imputations $x, y, z, \cdots$—i.e. between various states of society—correspond to the various ways in which these can unstabilize—i.e. upset—each other. That various groups of participants acting as effective sets in various relations of this kind may bring about "cyclical" dominations—e.g., y over x, z over y, and x over z—is indeed one of the most characteristic difficulties which a theory of these phenomena must face.

4.5. The Precise Definition of a Solution

4.5.1. Thus our task is to replace the notion of the optimum—i.e. of the first element—by something which can take over its functions in a static equilibrium. This becomes necessary because the original concept has become untenable. We first observed its breakdown in the specific instance of a certain three-person game in 4.3.2.-4.3.3. But now we have acquired a deeper insight into the cause of its failure: it is the nature of our concept of domination, and specifically its intransitivity.

This type of relationship is not at all peculiar to our problem. Other instances of it are well known in many fields and it is to be regretted that they have never received a generic mathematical treatment. We mean all those concepts which are in the general nature of a comparison of preference or "superiority," or of order, but lack transitivity: e.g., the strength of chess players in a tournament, the "paper form" in sports and races, etc.[1]

4.5.2. The discussion of the three-person game in 4.3.2.-4.3.3. indicated that the solution will be, in general, a set of imputations instead of a single imputation. That is, the concept of the "first element" will have to be replaced by that of a set of elements (imputations) with suitable properties. In the exhaustive discussion of this game in 32. (cf. also the interpretation in 33.1.1. which calls attention to some deviations) the system of three imputations, which was introduced as the solution of the three-person game in 4.3.2.-4.3.3., will be derived in an exact way with the help of the postulates of 30.1.1. These postulates will be very similar to those which characterize a first element. They are, of course, requirements for a set of elements (imputations), but if that set should turn out to consist of a single element only, then our postulates go over into the characterization of the first element (in the total system of all imputations).

We do not give a detailed motivation for those postulates as yet, but we shall formulate them now hoping that the reader will find them to be some-

[1] Some of these problems have been treated mathematically by the introduction of chance and probability. Without denying that this approach has a certain justification, we doubt whether it is conducive to a complete understanding even in those cases. It would be altogether inadequate for our considerations of social organization.

what plausible. Some reasons of a qualitative nature, or rather one possible interpretation, will be given in the paragraphs immediately following.

4.5.3. The postulates are as follows: A set S of elements (imputations) is a solution when it possesses these two properties:

(4:A:a) No y contained in S is dominated by an x contained in S.

(4:A:b) Every y not contained in S is dominated by some x contained in S.

(4:A:a) and (4:A:b) can be stated as a single condition:

(4:A:c) The elements of S are precisely those elements which are undominated by elements of S.[1]

The reader who is interested in this type of exercise may now verify our previous assertion that for a set S which consists of a single element x the above conditions express precisely that x is the first element.

4.5.4. Part of the malaise which the preceding postulates may cause at first sight is probably due to their circular character. This is particularly obvious in the form (4:A:c), where the elements of S are characterized by a relationship which is again dependent upon S. It is important not to misunderstand the meaning of this circumstance.

Since our definitions (4:A:a) and (4:A:b), or (4:A:c), are circular—i.e. implicit—for S, it is not at all clear that there really exists an S which fulfills them, nor whether—if there exists one—the S is unique. Indeed these questions, at this stage still unanswered, are the main subject of the subsequent theory. What is clear, however, is that these definitions tell unambiguously whether any particular S is or is not a solution. If one insists on associating with the concept of a definition the attributes of existence and uniqueness of the object defined, then one must say: We have not given a definition of S, but a definition of a property of S—we have not defined the solution but characterized all possible solutions. Whether the totality of all solutions, thus circumscribed, contains no S, exactly one S, or several S's, is subject for further inquiry.[2]

4.6. Interpretation of Our Definition in Terms of "Standards of Behavior"

4.6.1. The single imputation is an often used and well understood concept of economic theory, while the sets of imputations to which we have been led are rather unfamiliar ones. It is therefore desirable to correlate them with something which has a well established place in our thinking concerning social phenomena.

[1] Thus (4:A:c) is an exact equivalent of (4:A:a) and (4:A:b) together. It may impress the mathematically untrained reader as somewhat involved, although it is really a straightforward expression of rather simple ideas.

[2] It should be unnecessary to say that the circularity, or rather implicitness, of (4:A:a) and (4:A:b), or (4:A:c), does not at all mean that they are tautological. They express, of course, a very serious restriction of S.

Indeed, it appears that the sets of imputations S which we are considering correspond to the "standard of behavior" connected with a social organization. Let us examine this assertion more closely.

Let the physical basis of a social economy be given,—or, to take a broader view of the matter, of a society.[1] According to all tradition and experience human beings have a characteristic way of adjusting themselves to such a background. This consists of not setting up one rigid system of apportionment, i.e. of imputation, but rather a variety of alternatives, which will probably all express some general principles but nevertheless differ among themselves in many particular respects.[2] This system of imputations describes the "established order of society" or "accepted standard of behavior."

Obviously no random grouping of imputations will do as such a "standard of behavior": it will have to satisfy certain conditions which characterize it as a possible order of things. This concept of possibility must clearly provide for conditions of stability. The reader will observe, no doubt, that our procedure in the previous paragraphs is very much in this spirit: The sets S of imputations $x, y, z, \cdots$ correspond to what we now call "standard of behavior," and the conditions (4:A:a) and (4:A:b), or (4:A:c), which characterize the solution S express, indeed, a stability in the above sense.

4.6.2. The disjunction into (4:A:a) and (4:A:b) is particularly appropriate in this instance. Recall that domination of y by x means that the imputation x, if taken into consideration, excludes acceptance of the imputation y (this without forecasting what imputation will ultimately be accepted, cf. 4.4.1. and 4.4.2.). Thus (4:A:a) expresses the fact that the standard of behavior is free from inner contradictions: No imputation y belonging to S—i.e. conforming with the "accepted standard of behavior" —can be upset—i.e. dominated—by another imputation x of the same kind. On the other hand (4:A:b) expresses that the "standard of behavior" can be used to discredit any non-conforming procedure: Every imputation y not belonging to S can be upset—i.e. dominated—by an imputation x belonging to S.

Observe that we have not postulated in 4.5.3. that a y belonging to S should never be dominated by any x.[3] Of course, if this should happen, then x would have to be outside of S, due to (4:A:a). In the terminology of social organizations: An imputation y which conforms with the "accepted

[1] In the case of a game this means simply—as we have mentioned before—that the rules of the game are given. But for the present simile the comparison with a social economy is more useful. We suggest therefore that the reader forget temporarily the analogy with games and think entirely in terms of social organization.

[2] There may be extreme, or to use a mathematical term, "degenerate" special cases where the setup is of such exceptional simplicity that a rigid single apportionment can be put into operation. But it seems legitimate to disregard them as non-typical.

[3] It can be shown, cf. (31:M) in 31.2.3., that such a postulate cannot be fulfilled in general; i.e. that in all really interesting cases it is impossible to find an S which satisfies it together with our other requirements.

standard of behavior" may be upset by another imputation x, but in this case it is certain that x does not conform.[1] It follows from our other requirements that then x is upset in turn by a third imputation z which again conforms. Since y and z both conform, z cannot upset y—a further illustration of the intransitivity of "domination."

Thus our solutions S correspond to such "standards of behavior' as have an inner stability: once they are generally accepted they overrule everything else and no part of them can be overruled within the limits of the accepted standards. This is clearly how things are in actual social organizations, and it emphasizes the perfect appropriateness of the circular character of our conditions in 4.5.3.

4.6.3. We have previously mentioned, but purposely neglected to discuss, an important objection: That neither the existence nor the uniqueness of a solution S in the sense of the conditions (4:A:a) and (4:A:b), or (4:A:c), in 4.5.3. is evident or established.

There can be, of course, no concessions as regards existence. If it should turn out that our requirements concerning a solution S are, in any special case, unfulfillable,—this would certainly necessitate a fundamental change in the theory. Thus a general proof of the existence of solutions S for all particular cases[2] is most desirable. It will appear from our subsequent investigations that this proof has not yet been carried out in full generality but that in all cases considered so far solutions were found.

As regards uniqueness the situation is altogether different. The often mentioned "circular" character of our requirements makes it rather probable that the solutions are not in general unique. Indeed we shall in most cases observe a multiplicity of solutions.[3] Considering what we have said about interpreting solutions as stable "standards of behavior" this has a simple and not unreasonable meaning, namely that given the same physical background different "established orders of society" or "accepted standards of behavior" can be built, all possessing those characteristics of inner stability which we have discussed. Since this concept of stability is admittedly of an "inner" nature—i.e. operative only under the hypothesis of general acceptance of the standard in question—these different standards may perfectly well be in contradiction with each other.

4.6.4. Our approach should be compared with the widely held view that a social theory is possible only on the basis of some preconceived principles of social purpose. These principles would include quantitative statements concerning both the aims to be achieved *in toto* and the apportionments between individuals. Once they are accepted, a simple maximum problem results.

[1] We use the word "conform" (to the "standard of behavior") temporarily as a synonym for *being contained in S*, and the word "upset" as a synonym for *dominate*.

[2] In the terminology of games: for all numbers of participants and for all possible rules of the game.

[3] An interesting exception is 65.8.

Let us note that no such statement of principles is ever satisfactory *per se*, and the arguments adduced in its favor are usually either those of inner stability or of less clearly defined kinds of desirability, mainly concerning distribution.

Little can be said about the latter type of motivation. Our problem is not to determine what ought to happen in pursuance of any set of—necessarily arbitrary—*a priori* principles, but to investigate where the equilibrium of forces lies.

As far as the first motivation is concerned, it has been our aim to give just those arguments precise and satisfactory form, concerning both global aims and individual apportionments. This made it necessary to take up the entire question of inner stability as a problem in its own right. A theory which is consistent at this point cannot fail to give a precise account of the entire interplay of economic interests, influence and power.

4.7. Games and Social Organizations

4.7. It may now be opportune to revive the analogy with games, which we purposely suppressed in the previous paragraphs (cf. footnote 1 on p. 41). The parallelism between the solutions S in the sense of 4.5.3. on one hand and of stable "standards of behavior" on the other can be used for corroboration of assertions concerning these concepts in both directions. At least we hope that this suggestion will have some appeal to the reader. We think that the procedure of the mathematical theory of games of strategy gains definitely in plausibility by the correspondence which exists between its concepts and those of social organizations. On the other hand, almost every statement which we—or for that matter anyone else—ever made concerning social organizations, runs afoul of some existing opinion. And, by the very nature of things, most opinions thus far could hardly have been proved or disproved within the field of social theory. It is therefore a great help that all our assertions can be borne out by specific examples from the theory of games of strategy.

Such is indeed one of the standard techniques of using models in the physical sciences. This two-way procedure brings out a significant function of models, not emphasized in their discussion in 4.1.3.

To give an illustration: The question whether several stable "orders of society" or "standards of behavior" based on the same physical background are possible or not, is highly controversial. There is little hope that it will be settled by the usual methods because of the enormous complexity of this problem among other reasons. But we shall give specific examples of games of three or four persons, where one game possesses several solutions in the sense of 4.5.3. And some of these examples will be seen to be models for certain simple economic problems. (Cf. 62.)

4.8. Concluding Remarks

4.8.1. In conclusion it remains to make a few remarks of a more formal nature.

We begin with this observation: Our considerations started with single imputations—which were originally quantitative extracts from more detailed combinatorial sets of rules. From these we had to proceed to sets S of imputations, which under certain conditions appeared as solutions. Since the solutions do not seem to be necessarily unique, the complete answer to any specific problem consists not in finding a solution, but in determining the set of all solutions. Thus the entity for which we look in any particular problem is really a set of sets of imputations. This may seem to be unnaturally complicated in itself; besides there appears no guarantee that this process will not have to be carried further, conceivably because of later difficulties. Concerning these doubts it suffices to say: First, the mathematical structure of the theory of games of strategy provides a formal justification of our procedure. Second, the previously discussed connections with "standards of behavior" (corresponding to sets of imputations) and of the multiplicity of "standards of behavior" on the same physical background (corresponding to sets of sets of imputations) make just this amount of complicatedness desirable.

One may criticize our interpretation of sets of imputations as "standards of behavior." Previously in 4.1.2. and 4.1.4. we introduced a more elementary concept, which may strike the reader as a direct formulation of a "standard of behavior": this was the preliminary combinatorial concept of a solution as a set of rules for each participant, telling him how to behave in every possible situation of the game. (From these rules the single imputations were then extracted as a quantitative summary, cf. above.) Such a simple view of the "standard of behavior" could be maintained, however, only in games in which coalitions and the compensations between coalition partners (cf. 4.3.2.) play no role, since the above rules do not provide for these possibilities. Games exist in which coalitions and compensations can be disregarded: e.g. the two-person game of zero-sum mentioned in 4.2.3., and more generally the "inessential" games to be discussed in 27.3. and in (31:P) of 31.2.3. But the general, typical game—in particular all significant problems of a social exchange economy—cannot be treated without these devices. Thus the same arguments which forced us to consider sets of imputations instead of single imputations necessitate the abandonment of that narrow concept of "standard of behavior." Actually we shall call these sets of rules the "strategies" of the game.

4.8.2. The next subject to be mentioned concerns the static or dynamic nature of the theory. We repeat most emphatically that our theory is thoroughly static. A dynamic theory would unquestionably be more complete and therefore preferable. But there is ample evidence from other branches of science that it is futile to try to build one as long as the static side is not thoroughly understood. On the other hand, the reader may object to some definitely dynamic arguments which were made in the course of our discussions. This applies particularly to all considerations concerning the interplay of various imputations under the influence of "domination," cf. 4.6.2. We think that this is perfectly legitimate. A static theory deals with equilibria.[1] The essential characteristic of an equilibrium is that it has no tendency to change, i.e. that it is not conducive to dynamic developments. An analysis of this feature is, of course, inconceivable without the use of certain rudimentary dynamic concepts. The important point is that they are rudimentary. In other words: For the real dynamics which investigates the precise motions, usually far away from equilibria, a much deeper knowledge of these dynamic phenomena is required.[2,3]

4.8.3. Finally let us note a point at which the theory of social phenomena will presumably take a very definite turn away from the existing patterns of mathematical physics. This is, of course, only a surmise on a subject where much uncertainty and obscurity prevail.

Our static theory specifies equilibria—i.e. solutions in the sense of 4.5.3.—which are sets of imputations. A dynamic theory—when one is found—will probably describe the changes in terms of simpler concepts: of a single imputation—valid at the moment under consideration—or something similar. This indicates that the formal structure of this part of the theory—the relationship between statics and dynamics—may be generically different from that of the classical physical theories.[4]

All these considerations illustrate once more what a complexity of theoretical forms must be expected in social theory. Our static analysis alone necessitated the creation of a conceptual and formal mechanism which is very different from anything used, for instance, in mathematical physics. Thus the conventional view of a solution as a uniquely defined number or aggregate of numbers was seen to be too narrow for our purposes, in spite of its success in other fields. The emphasis on mathematical methods seems to be shifted more towards combinatorics and set theory—and away from the algorithm of differential equations which dominate mathematical physics.

[1] The dynamic theory deals also with inequilibria—even if they are sometimes called "dynamic equilibria."

[2] The above discussion of statics *versus* dynamics is, of course, not at all a construction *ad hoc*. The reader who is familiar with mechanics for instance will recognize in it a reformulation of well known features of the classical mechanical theory of statics and dynamics. What we do claim at this time is that this is a general characteristic of scientific procedure involving forces and changes in structures.

[3] The dynamic concepts which enter into the discussion of static equilibria are parallel to the "virtual displacements" in classical mechanics. The reader may also remember at this point the remarks about "virtual existence" in 4.3.3.

[4] Particularly from classical mechanics. The analogies of the type used in footnote 2 above, cease at this point.