physics
MDPI

Review

Selected Topics of Social Physics: Equilibrium Systems †

Vyacheslav I. Yukalov [1,2]

[1] Bogolubov Laboratory of Theoretical Physics, Joint Institute for Nuclear Research, Dubna 141980, Russia; yukalov@theor.jinr.ru
[2] Instituto de Fisica de São Carlos, Universidade de São Paulo, CP 369, São Paulo 13560-970, Brazil

† The article represents the first part of the review based on the lectures that the author gave during several years at the Swiss Federal Institute of Technology in Zürich (ETH Zürich).

Abstract: The paper gives an introduction to the physics approach to social systems providing the main definitions and notions used in the modeling of these systems. The behavior of social systems is illustrated by several quite simple, typical models. The present part considers equilibrium systems. Nonequilibrium systems will be presented in the second part of the review. The style of the paper combines the features of a tutorial and a survey, which, from one side, makes it simpler to read for nonspecialists aiming to grasp the basics of social physics, and from the other side, describes several rather recent original models containing new ideas that could be of interest to experienced researchers in the field. The selection of the material is limited and motivated by the author's research interests.

Keywords: social systems; typical agents; information and statistics; social phase transitions; yes–no model; regulation cost; fluctuating groups; self-organized disorder; population coexistence; country disintegration; formation of collective decisions

check for updates

Citation: Yukalov, V.I. Selected Topics of Social Physics: Equilibrium Systems. *Physics* **2023**, *5*, 590–635. https://doi.org/10.3390/physics5020041

Received: 10 February 2023
Revised: 11 May 2023
Accepted: 16 May 2023
Published: 2 June 2023

1. Introduction

Nowadays, physics methods and models are often used in social science for characterizing the behavior of different social systems. It is useful to stress that physics provides methods and models for describing social systems, but it does not pretend to replace social science.

Under a social system, one implies a bound collective of a number of interacting agents. These can be various human societies, starting with families, professional groups, members of organizations, financial markets, country populations, etc., or these can be animal societies forming biological systems, e.g., wolf packs, fish shoals, bird flocks, horse herds, swarms of bees, ant colonies, and the like. Thus, the number, N, of agents in a society is more than one, and generally it is much larger than one: $N \gg 1$.

The collective of agents in a society is bound in the sense of having some features uniting the society members. For example, the members of a society can participate in a joint activity or share similar goals or beliefs. The common features uniting the members make the collective stable or metastable, that is, bound for a sufficiently long time that is much larger than the interaction time between the members. Agent interactions can be physical, economic, financial, and so on. Societies can be formed by force, as in the army and in prison, or they can be self-organized.

Social systems pertain to the class of complex systems. A system is complex if some of its properties are qualitatively different from the agglomerated properties of its parts. Complex systems can be structured, consisting of parts that are complex systems themselves. For instance, the human world is composed of countries that, actually, are also complex systems. In that sense, social systems are often hierarchical, containing several levels of complex systems. For example, the human world is made up of countries that are composed of social groups that form organizations that include individuals whose bodies are made of biological cells and whose brains making decisions are extremely complex systems.

To represent a complex system, the society needs to have the following properties. The number of agents composing the society must be large: $N \gg 1$; however, this is not sufficient. If the members of the society are independent, the overall system can be characterized by a set of independent agents, this does not make such a system complex. In a complex society, its features are not just a sum of the features of the separate members. Thus, bees can fly, but a bee colony is not just a swarm of flying bees. A group of people walking separately is not a social system. A society becomes complex due to its agents' interactions and mutual relations, which cause the nonadditivity of the society features. Strictly speaking, there is no generally accepted mathematical measure of social system complexity [1–5]. Complexity is rather a qualitative notion, assuming that a society is complex if it is stable and consists of many interacting agents, so the typical society features are not just an arithmetic average of the features of separate individuals.

For example, an ant colony is a complex system since the behavior of separate ants is regulated by their interactions and distribution of jobs; thus, ant queens lay eggs, worker ants form ant-hills and feed the queen, and winged males mate with the queen and die. Another example of a complex system is a bee colony, where a queen produces eggs; workers clean out the cells, remove debris, feed the brood, care for the queen, build combs, guard the entrance, ventilate the hive, and forage for nectar, pollen, propolis, and water; and drones fertilize the queen and die upon mating.

There are three interconnected problems in the description of complex social systems: how to model a society, how to investigate the model, and how it would be possible to regulate the society's behavior. A model is to be simple enough and, at the same time, realistic. Too simple models may not describe reality, but too complicated models can distort reality because of the accumulation of errors in excessively complicated descriptions. Sometimes it happens that "less is more" [6]. So, it is desirable that models are neither trivial nor overcomplicated.

In this review, first, the basic ideas are described, allowing one to understand the general principles of constructing society models, so that, from one side, the model is not overcomplicated and, from another side, can identify the characteristic features of the considered society. Second, the main methods of investigating the system's behavior are studied. Third, conditions are discussed, making it possible to find the desired properties of the society and the related optimal parameters of the models.

The layout of the present review is as follows. In Section 2, the principle of minimal information is discussed, which plays a pivotal role in establishing probability distributions for equilibrium and quasi-equilibrium systems. Section 3 discusses some quite simple, typical models of equilibrium social systems. Section 4 gives the basics of collective decision-making in a society. Finally, Section 5 concludes.

The material of this paper is based on the lectures that the author has been giving for several years at the Swiss Federal Institute of Technology in Zürich (ETH Zürich). Overall, the content constitutes a one-year course consisting of two parts, one devoted to equilibrium (or quasi-equilibrium) systems and the other to nonequilibrium systems. Following this natural separation, the presentation of the material is also split into the corresponding two parts. The present review deals with equilibrium social systems. The forthcoming second part will consider nonequilibrium social systems [7]. Being bound by the lecture frames, the content of the review is limited accordingly. The choice of the presented material is based on the research interests of the author.

2. Principle of Minimal Information

Actually, almost all models that intend to describe the behavior of equilibrium or quasi-equilibrium complex systems start with some extremization principles. These principles are widely spread in science as well as in life. Just as a joke, we may say that all life follows the extremization principle, which can be formulated as a "minimum of labor and maximum of pleasure".

To model social systems in complicated situations where not all information is available, it is customary to resort to probabilistic descriptions, defining the probability distribution based on the principle of minimal information. This principle helps to develop the optimal description of a social system with limited information on its properties. The information is always limited since there are many agents and, in addition, their actions are often not absolutely rational. Experimental studies of brain activity well demonstrate that only a finite amount of information can be successfully processed by living beings [8]. The rationality of social individuals is always bound [9]. In addition, there exists random influences from the environment. Moreover, not all information is even necessary for a correct description, but excessive information can lead to the accumulation of errors and incorrect conclusions.

2.1. Information Entropy

Let us consider a social system of N agents enumerated by the index $j = 1, 2, \ldots, N$. Generally, the number of agents depends on time. Each j-th agent is marked by a set of characteristics,

$$\sigma_j = \{ \sigma_{j\alpha} : \ \alpha = 1, 2, \ldots, M_j \} \; . \tag{1}$$

The collection of the characteristics of all agents makes up the *society configuration set*,

$$X = \{ \sigma_j : \ j = 1, 2, \ldots, N \} \; . \tag{2}$$

The probability distribution, $\rho(\sigma)$, over the variables $\sigma \in X$ is to be normalized,

$$\sum_\sigma \rho(\sigma) = 1 \; , \qquad 0 \leq \rho(\sigma) \leq 1 \; , \tag{3}$$

where the sum over σ implies

$$\sum_\sigma = \sum_{j=1}^N \sum_{\alpha=1}^{M_j} \; .$$

If all variables are equiprobable, then the distribution is uniform,

$$\rho(\sigma) = \frac{1}{N_{\rm tot}} \qquad \left(N_{\rm tot} = \sum_{j=1}^N M_j \right) \; .$$

A *statistical social system* is the triplet,

$$\{ N, \ X, \ \rho(\sigma) \} \; . \tag{4}$$

However, the probability distribution needs to be defined. For this purpose, one introduces the information entropy, S, which is a measure of the system uncertainty. The entropy is a measure of the inaccessible information, representing a functional of the distribution over the society's characteristics,

$$S = S[\, \rho(\sigma) \,] \tag{5}$$

benefiting from the following properties.

(i) *Continuity*. It is a continuous function of the distribution, such that, for a small variation, $\delta\rho$, in the latter, the entropy variation, δS, is small:

$$S[\, \rho(\sigma) + \delta\rho(\sigma) \,] - S[\, \rho(\sigma) \,] \simeq \frac{\delta S[\, \rho(\sigma) \,]}{\delta\rho(\sigma)} \; \delta\rho(\sigma) \; . \tag{6}$$

(ii) *Monotonicity*. For a uniform distribution,

$$S\left[\frac{1}{N_{\text{tot}}}\right] < S\left[\frac{1}{N_{\text{tot}}+1}\right] . \tag{7}$$

(iii) *Additivity*. If a statistical system with the set of variables σ can be separated into two mutually independent subsystems with the variable sets σ_1 and σ_2, such that

$$\rho(\sigma) = \rho(\sigma_1)\,\rho(\sigma_2)$$

and

$$\sum_{\sigma_1}\rho(\sigma_1) = 1 , \qquad \sum_{\sigma_2}\rho(\sigma_2) = 1 ,$$

then the entropy of the total system is a sum of the subsystems entropies,

$$S[\,\rho(\sigma_1)\,\rho(\sigma_2)\,] = S[\,\rho(\sigma_1)\,] + S[\,\rho(\sigma_2)\,] . \tag{8}$$

Theorem 1 ([10])**.** *The unique functional, satisfying the above conditions, up to a positive constant factor, has the form:*

$$S[\,\rho(\sigma)\,] = -\sum_{\sigma}\rho(\sigma)\;\ln\;\rho(\sigma) . \tag{9}$$

Here, the natural logarithm is used, but, generally, it is not essential which type of the logarithm base is employed since the entropy is defined up to a constant factor. The information entropy coincides with Gibbs' entropy in statistical mechanics [11,12]. For the uniform distribution, one has:

$$S\left[\frac{1}{N_{\text{tot}}}\right] = \ln N_{\text{tot}} , \tag{10}$$

which, visibly, is a monotonic increasing function of N_{tot}. Generally, the entropy is in the range,

$$0 \le S[\,\rho(\sigma)\,] \le \ln N_{\text{tot}} . \tag{11}$$

The entropy is zero when there is just a single agent with a single characteristic, so that if $N_{\text{tot}} = 1$ and the distribution is trivial, then

$$S[\,1\,] = 0 \qquad (N_{\text{tot}} = 1) . \tag{12}$$

The information entropy is a measure of the system's uncertainty. In other words, one can say that entropy is a measure of inaccessible information.

When a statistical system with a distribution $\rho(\sigma)$ is initially characterized by a trial likelihood distribution, $\rho_0(\sigma)$, then the form,

$$I_{KL}[\,\rho(\sigma)\,] = \sum_{\sigma}\rho(\sigma)\;\ln\;\frac{\rho(\sigma)}{\rho_0(\sigma)} , \tag{13}$$

is called the Kullback–Leibler *information gain*, relative entropy, or Kullback–Leibler divergence [13,14].

2.2. Information Functional

The information functional has to include the information gain (13). In addition, one should not forget that the probability distribution is to be normalized, as in Equation (3), hence

$$\sum_{\sigma}\rho(\sigma) - 1 = 0 . \tag{14}$$

There can exist other information defining some average quantities, C_i, by the condition,

$$\sum_\sigma \rho(\sigma)C_i(\sigma) - C_i = 0 \; . \tag{15}$$

Overall, the information functional takes the form,

$$I[\,\rho(\sigma)\,] = \sum_\sigma \rho(\sigma)\; \ln\,\frac{\rho(\sigma)}{\rho_0(\sigma)} + \lambda_0 \left[\sum_\sigma \rho(\sigma) - 1\right] + \sum_i \lambda_i \left[\sum_\sigma \rho(\sigma)C_i(\sigma) - C_i\right] \; . \tag{16}$$

The use of the Kullback–Leibler information gain (13) for deriving probability distributions is justified by the Shore–Johnson theorem [15].

Theorem 2 ([15])**.** *There exists only one distribution satisfying consistency conditions, and this distribution is uniquely defined by the minimum of the Kullback–Leibler information gain (13) under given constraints.*

In the information functional (16), the coefficients λ_0 and λ_i are the Lagrange multipliers, whose variation,

$$\frac{\partial I[\,\rho(\sigma)\,]}{\partial\lambda_0} = 0 \; , \qquad \frac{\partial I[\,\rho(\sigma)\,]}{\partial\lambda_i} = 0 \; ,$$

yields the normalization condition (14) and the expectation conditions (15).

Principle of minimal information. *The probability distribution of an equilibrium statistical system (4) is defined as the minimizer of the information functional (16), thus satisfying the conditions*:

$$\frac{\delta I[\,\rho(\sigma)\,]}{\delta\rho(\sigma)} = 0 \; , \qquad \frac{\delta^2 I[\,\rho(\sigma)\,]}{\delta\rho(\sigma)^2} > 0 \; . \tag{17}$$

The minimization conditions give:

$$\frac{\delta I[\,\rho(\sigma)\,]}{\delta\rho(\sigma)} = 1 + \lambda_0 + \ln\,\frac{\rho(\sigma)}{\rho_0(\sigma)} + \sum_i \lambda_i C_i(\sigma) \; , \qquad \frac{\delta^2 I[\,\rho(\sigma)\,]}{\delta\rho(\sigma)^2} = \frac{1}{\rho(\sigma)} \; ,$$

which leads to the distribution,

$$\rho(\sigma) = \frac{\rho_0(\sigma)\exp\{-\sum_i \lambda_i C_i(\sigma)\}}{\sum_\sigma \rho_0(\sigma)\exp\{-\sum_i \lambda_i C_i(\sigma)\}} \; . \tag{18}$$

If there is no preliminary information on the trial distribution properties, so that all states are equiprobable, this implies that the trial distribution is uniform:

$$\rho_0(\sigma) = \frac{1}{N_{\rm tot}} \; .$$

Then, the sought distribution reads:

$$\rho(\sigma) = \frac{1}{Z}\; \exp\left\{ -\sum_i \lambda_i C_i(\sigma) \right\} \; , \tag{19}$$

with the normalization factor, called partition function,

$$Z = \sum_\sigma \exp\left\{ -\sum_i \lambda_i C_i(\sigma) \right\} \; . \tag{20}$$

In statistical mechanics, this is called the Gibbs' distribution [11,12]. The principle of minimal information is equivalent to the conditional maximization of entropy. As is stated by Jaynes, the Gibbs' distribution can be used for any complex system whose description is based on information theory [16,17].

It may happen that the probability distribution $\rho(\sigma, x)$ depends on a parameter x that has not been uniquely prescribed. In such a case, this parameter can be chosen such that it follows the principle of minimal information. To this end, substituting the distribution (19) into the information function (16) results in

$$I[\,\rho(\sigma, x)\,] = R(x) - \sum_i \lambda_i C_i \; , \tag{21}$$

where

$$R(x) = -\ln Z(x) \tag{22}$$

is *relative information*. Keeping in mind that the average quantities C_i are fixed, one can see that the minimization of the information function with respect to the parameter x is equivalent to the minimization of the relative information (22), since

$$\frac{\partial I[\,\rho(\sigma, x)\,]}{\partial x} = \frac{\partial R(x)}{\partial x} = 0 \; , \qquad \frac{\partial^2 I[\,\rho(\sigma, x)\,]}{\partial x^2} = \frac{\partial^2 R(x)}{\partial x^2} > 0 \; . \tag{23}$$

In this way, the probability distribution can be uniquely defined [18].

2.3. Representative Ensembles

When the probability distribution is defined by the principle of minimal information, this does not mean that it is useful to have little information on the system. Vice versa, it is necessary to include in the information function all available relevant information on the system corresponding to the expected-value conditions. The principle of minimal information shows how to obtain an optimal description of a complex system while possessing minimal information on the latter. All available important information on the system must be taken into account.

Thus, a statistical system is described by the triplet (4), consisting of N system members, society set X, and the probability distribution $\rho(\sigma)$. The pair $\{X, \rho(\sigma)\}$ is called a *statistical ensemble*. Observables are represented by real functions $A(\sigma) = A^*(\sigma)$ whose averages,

$$\langle\, A(\sigma)\, \rangle = \sum_\sigma \rho(\sigma) A(\sigma) \; , \tag{24}$$

are the observable quantities that can be measured. The collection of all available observable quantities $\{\langle A(\sigma)\rangle\}$ is a *statistical state*.

A statistical ensemble is termed *representative* when it provides the correct description for the system's statistical state, so the theoretical expectation values of observable quantities accurately describe the corresponding measured values. To this end, it is necessary to include in the information function all relevant information in the form of additional constraints. Only then will the minimization of the information function produce a correct probability distribution.

The idea of representative ensembles goes back to Gibbs [11,12], and their importance is emphasized by Ter Haar [19,20]. The necessity of employing representative ensembles for obtaining reliable theoretical estimates has been analyzed, and the explanation of how the use of non-representative ensembles leads to incorrect results has been discussed in detail [21–24].

2.4. Arrow of Time

In general, the probability distribution, $\rho(\sigma, t)$, can depend on time t. An important question is: Why is time assumed to always increase? One often connects this with the second law of thermodynamics, according to which the entropy of an isolated system left to spontaneous evolution cannot decrease [25]. However, strictly speaking, by Liouville's theorem, the entropy of a closed system remains constant in time (see, e.g., [26]). It is possible to infer that the arrow of time appears in quasi-isolated systems due to their stochastic instability [27–31]. Here, we show that there is quite a straightforward way of proving that the irreversibility of time can be connected with the non-decrease in information gain. First, one nees to recall the Gibbs' inequality.

Gibbs' inequality. *For two non-negative functions, $A(\sigma) \geq 0$ and $B(\sigma) \geq 0$, one has*:

$$\sum_\sigma [\, A(\sigma) \ln A(\sigma) - A(\sigma) \ln B(\sigma) \,] \geq \sum_\sigma [\, A(\sigma) - B(\sigma) \,] \; . \tag{25}$$

Proof. The proof follows from the inequality

$$\ln \frac{A\sigma)}{B(\sigma)} \geq 1 \; - \; \frac{B(\sigma)}{A(\sigma)} \; .$$

□

Consequence. The information gain caused by the transition from the probability $\rho_0(\sigma)$ to $\rho(\sigma)$ is semi-positive:

$$\sum_\sigma \rho(\sigma) \; \ln \frac{\rho(\sigma)}{\rho_0(\sigma)} \geq 0 \; . \tag{26}$$

More examples of useful inequalities in information theory can be found in Ref. [32].

Now, let us consider the natural change in the probability distribution with time, starting from the initial value, $\rho(\sigma, 0)$, to the final $\rho(\sigma, t)$ at time t. Then, the following statement is valid.

Non-decrease in information gain. *The information gain, due to the evolution of the probability distribution from the initial distribution $\rho(\sigma, 0)$ to the final $\rho(\sigma, t)$, does not decrease*:

$$\sum_\sigma \rho(\sigma, t) \; \ln \frac{\rho(\sigma, t)}{\rho(\sigma, 0)} \geq 0 \; . \tag{27}$$

Thus, the non-decrease in the information gain with time can be connected with the direction of time.

3. Equilibrium Social Systems

This Section introduces the main notions required for modeling equilibrium social systems and studies some simple enough statistical models. In the long run, social systems are, strictly speaking, nonequilibrium. However, if the society does not experience external shocks during a period of time much longer than the typical interaction time between the society members, this society can be treated as equilibrium for that period of time. More details on the so-called social physics can be found in Refs. [33–38].

3.1. Free Energy

Equilibrium social systems can be treated by statistical theory as a particular application of the above notions of information theory and the principle of minimal information. Following the notation of Section 2, let us consider a society of N members, where each member is associated with characteristics (1). The considered society occupies a volume V.

Our aim is to study almost isolated societies with self-organization due to their internal properties. Of course, no real society can be absolutely isolated from its surroundings. The influence of surroundings is treated as stationary random perturbations or stationary noise. The perturbations can be produced by other societies and by natural causes, such as earthquakes, floods, droughts, epidemics, etc. The noise is considered stationary, which implies that an equilibrium situation is assumed. The influence of noise on the society is measured by temperature, T, which is a measure of noise intensity. In limiting cases, the absence of noise implies $T = 0$, and extremely strong noise means $T = \infty$. One often uses the inverse temperature,

$$\beta \equiv \frac{1}{T} \; , \tag{28}$$

which can be interpreted as the measure of a society's isolation from random noise. Here, $\beta = 0$ means no isolation and extreme noise, while $\beta = \infty$ implies complete isolation and no noise.

An equilibrium system is conveniently characterized by a functional $H(\sigma)$, termed a Hamiltonian or harm. The expected value of the Hamiltonian is society's energy,

$$E = \sum_\sigma \rho(\sigma) H(\sigma) \equiv \langle\, H(\sigma) \,\rangle \; , \tag{29}$$

which is also called society cost. In some economic and financial applications, it is termed disagreement or dissatisfaction, since higher energy assumes a more excited society.

The energy of noise is TS, with S being the information entropy. The free energy is part of the society's energy, due to the society itself without the noise energy:

$$F = E - TS \; . \tag{30}$$

In the limiting case of a zero temperature, when there is no noise, the free energy coincides with the society's energy,

$$F = E \qquad (T = 0) \; . \tag{31}$$

The probability distribution over the society characteristics is defined by the principle of minimal information, which, in the present notation, gives

$$\rho(\sigma) = \frac{1}{Z} \, e^{-\beta H(\sigma)} \; , \tag{32}$$

with the partition function

$$Z = \sum_\sigma e^{-\beta H(\sigma)} \; . \tag{33}$$

Substituting distribution (32) into entropy (9) leads to

$$S = \beta E + \ln Z \; . \tag{34}$$

Comparing Equation (34) with Equation (30) results in the expression,

$$F = -T \; \ln Z \; . \tag{35}$$

It is straightforward to check that the entropy S can be represented as

$$S = -\,\frac{\partial F}{\partial T} \; . \tag{36}$$

The society pressure is defined as

$$P = -\,\frac{\partial F}{\partial V} \; . \tag{37}$$

The relation (37) implies that the region occupied by a society can be changed because of the pressure. Population density is

$$\rho \equiv \frac{N}{V} \; . \tag{38}$$

Combining Equation (22) and (35) gives the equality

$$F(x) = -T \ln Z(x) = TR(x) \; . \tag{39}$$

Hence, the information function (21) and the free energy (39) are connected:

$$I[\, \rho(\sigma, x) \,] = \beta F(x) - \sum_i \lambda_i C_i \; . \tag{40}$$

Since the values C_i are fixed, one has:

$$\frac{\partial I[\, \rho(\sigma, x) \,]}{\partial x} = \frac{1}{T} \, \frac{\partial F(x)}{\partial x} \; . \tag{41}$$

Therefore, the minimization of the information function over a parameter x is equivalent to the minimization of free energy over this parameter:

$$\min_x I[\, \rho(\sigma, x) \,] \longleftrightarrow \min_x F(x) \; . \tag{42}$$

3.2. Society Stability

Considering society models, it is important to make sure that the society is stable. A society is stable if small variations in parameters do not drive it far from its initial state. Between several admissible statistical states, the system chooses that which provides the minimal free energy. The system is stable with respect to the variation in parameters if its free energy is minimal, so that

$$\delta F = 0 \; , \qquad \delta^2 F > 0 \; . \tag{43}$$

The variation in the intensity of noise, that is, of temperature, under a fixed volume, is characterized by specific heat,

$$C_V = \frac{1}{N} \left(\frac{\partial E}{\partial T} \right)_V = -\, \frac{T}{N} \left(\frac{\partial^2 F}{\partial T^2} \right)_V \; . \tag{44}$$

With the free energy (35), this takes the form,

$$C_V = \frac{\mathrm{var}(H)}{NT^2} \; , \tag{45}$$

where the variation in the Hamiltonian, $H = H(\sigma)$, means

$$\mathrm{var}(H) \equiv \langle\, H^2 \,\rangle - \langle\, H \,\rangle^2 \; . \tag{46}$$

The latter is non-negative; hence, the specific heat has to be non-negative.

The variation in volume, under a fixed temperature, is characterized by the compressibility,

$$\varkappa_T = -\, \frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_{TN} = \frac{1}{V} \left(\frac{\partial^2 F}{\partial V^2} \right)^{-1}_{TN} \; . \tag{47}$$

The condition of stability requires that the specific heat, as well as compressibility, be positive and finite:

$$0 \leq C_V < \infty \; , \qquad 0 \leq \varkappa_T < \infty \; . \tag{48}$$

In this way, the principle of minimal information agrees with the minimal free energy, which, in turn, assumes societal stability.

As an example of societal instabilities, one may mention the disintegration of a country as a result of a war or because of external economic pressures. Another example is the bankruptcy of a firm caused by changed financial conditions.

3.3. Practical Approaches

In order to accomplish quantitative investigations of society's properties, two different approaches are employed: the network approach and the typical-agent approach.

The network approach, also called multi-agent modeling, is based on the following assumptions:

(i) The considered society consists of agents, or nodes, that are fixed at the lattice sites of a spatial (often two-dimensional) lattice.

(ii) The agents interact with each other when they are close to each other; commonly, the nearest-neighbor interactions are considered.

(iii) Due to rather complicated calculations, generally, one has to resort to numerical modeling with computers.

There exist many examples of networks, such as electric-current networks, models of magnetic and ferroelectric materials, neuron networks in brains, computer networks, etc.

The results of the network approach depend on lattice dimensionality modeling the society (whether one-, two-, or three-dimensional lattices are considered), lattice geometry (cubic, triangular, or another structure), and on the interaction type and range (long-range, short-range, or mid-range). The network approach is appropriate for small or just structured systems, with agents that can be treated as fixed at spatial points.

However, complex societies, such as human and biological societies, are not composed of agents tied to a spatial lattice; commonly, they are not attached to any fixed spatial locations or sites. The agent interactions are not of the nearest-neighbor type and can be independent of the distance between the agents. Interactions in a society can be either direct, when interacting with fixed neighbors, or can involve changing neighbors. Nowadays, interactions that do not depend on distance are widespread, such as through phone, Skype, WhatsApp, Telegram, and the like. There exist indirect interactions through letters and e-mails, reading newspapers and books, listening to the radio, and watching television. The majority of interactions are long-range, but not solely with nearest neighbors. Other biological societies also interact at large distances by means of their voices and smells.

Summarizing, complex societies, such as human or other biological societies, are formed by agents that are not fixed at spatial locations and can interact over long distances. This kind of society is better described by the typical-agent approach. In this approach, one reduces the problem to the consideration of the behavior of typical agents, representing a kind of average member of the society. For example, the typical interaction of agents, having the characteristics σ_i and σ_j and described by the term $\sigma_i\sigma_j$, is transformed into the expression,

$$\sigma_i\sigma_j = \sigma_i\langle\, \sigma_j \,\rangle + \langle\, \sigma_i \,\rangle\sigma_j - \langle\, \sigma_i \,\rangle\langle\, \sigma_j \,\rangle \; , \tag{49}$$

with $i \neq j$. The average characteristic is

$$\langle\, \sigma_j \,\rangle = \sum_{\sigma} \rho(\sigma)\sigma_j \equiv s \; . \tag{50}$$

Thus, the representation (49) becomes

$$\sigma_i\sigma_j = s\sigma_i + s\sigma_j - s^2 \; , \tag{51}$$

and the average interaction reads:

$$\langle\, \sigma_i\sigma_j \,\rangle = \langle\, \sigma_i \,\rangle\langle\, \sigma_j \,\rangle = s^2 \; . \tag{52}$$

Then, the description of the interaction is reduced to the consideration of typical agents subject to the average influence of other typical agents. For complex social systems, the typical-agent approach is not merely simpler but also more correct.

3.4. Society Transitions

In general, one is interested not in the characteristics of single agents but in the general behavior of a society. For this purpose, one considers the mean arithmetic characteristic,

$$s(\sigma) = \frac{1}{N}\sum_{j=1}^{N}\sigma_j \; . \tag{53}$$

The observable quantity is the average characteristic

$$s \equiv \langle\, s(\sigma) \,\rangle = \frac{1}{N}\sum_{j=1}^{N}\langle\, \sigma_j \,\rangle \; . \tag{54}$$

The general behavior of a society can be associated with its average characteristics. When the property of a characteristic qualitatively changes, one speaks of a *social phase transition*. For example, it may happen that, under such a social transition, the average characteristic varies between zero and nonzero values. Conditionally, one can name the social state with $|s| > 0$ an ordered state, while the state where $s = 0$, a disordered state. In that case, the average characteristic (54) is termed an *order parameter*. There can occur the following types of transitions when a system parameter, for instance, temperature, varies.

First-order transition. This type of transition happens when the order parameter at some point T_0 changes from nonzero to zero by a discontinuous jump:

$$s(T_0 - 0) > 0 \; , \qquad s(T_0 + 0) = 0 \; , \qquad s(T_0 - 0) \neq s(T_0 + 0) \; . \tag{55}$$

Discontinuous social transitions can be associated with revolutions.

Second-order transition. In this case, the order parameter at a critical point T_c changes continuously from nonzero to zero:

$$s(T_c - 0) > 0 \; , \qquad s(T_c + 0) = 0 \; , \qquad s(T_c - 0) = s(T_c + 0) \; . \tag{56}$$

Continuous transitions describe evolutions.

Gradual crossover. The order parameter does not become zero at a finite point, but at some crossover point, T_c, it strongly diminishes and tends to zero only in the limit of large T:

$$s(T_c - 0) = s(T_c + 0) > 0 \; ,$$

$$s(T) \rightarrow 0 \qquad (T \gg T_c) \; . \tag{57}$$

This transition corresponds to a smooth evolution.

There can occur more unusual situations when a society is not completely equilibrium [39], but for the description of equilibrium societies, the above three types of social transitions are sufficient.

3.5. Yes–No Model

This Section considers a simple enough model that is known in statistical physics, where it is called the Ising model [40–43]. This model is also often used in different applications for financial and economic problems [44]. This model is needed in order to illustrate in action the notions introduced in the previous sections, to exemplify the terminology associated with social systems, and to have the ground to study more complicated models in the following sections.

Suppose each agent of a society can have just two features, which are opinions that can be termed "yes" and "no". Generally, this can be any decision with two alternatives. For instance, this can be voting for or against a candidate in elections, supporting or rejecting a suggestion in a referendum, buying or selling stocks in a market, etc. These two alternatives can be represented by the binary variable taking two possible values, e.g.:

$$\sigma_j = \begin{cases} -1\,, & \text{no}\,; \\ +1\,, & \text{yes}\,. \end{cases} \tag{58}$$

The interaction, or mutual influence, between two members of the society is expressed as

$$H_{ij} = -J_{ij}\sigma_i\sigma_j\,, \tag{59}$$

with the value $J_{ij} = J_{ji}$ being the intensity of the interaction. The case of agreement (collaboration) or disagreement (competition) among the members, respectively, corresponds to the values

$$J_{ij} > 0 \qquad (\text{agreement, collaboration})\,,$$

$$J_{ij} < 0 \qquad (\text{disagreement, competition})\,. \tag{60}$$

The terminology comes from the fact that, under agreement, when $J_{ij} > 0$, the interaction energy is minimal for coinciding σ_i and σ_j, while in the case of disagreement, when $J_{ij} < 0$, the interaction energy is minimal for opposite σ_i and σ_j. When there is mutual agreement and both agents vote in the same way, either both "yes" or both "no", the interaction energy is lower than when the agents vote differently.

The Hamiltonian of the system is

$$H = -\frac{1}{2}\sum_{i\neq j}^{N} J_{ij}\sigma_i\sigma_j\,. \tag{61}$$

In physics, this is called the Ising model; see the history of the model in Ref. [45].

In the typical-agent approach, the Hamiltonian reads:

$$H = -Js\sum_{j=1}^{N}\sigma_j + \frac{1}{2}\,Js^2N\,, \tag{62}$$

where

$$J \equiv \frac{1}{N}\sum_{i\neq j}^{N} J_{ij} = \sum_{i(\neq j)}^{N} J_{ij}\,. \tag{63}$$

Using the equalities,

$$\exp\left(-\beta Js\sum_{j=1}^{N}\sigma_j\right) = \prod_{j=1}^{N}\exp(-\beta Js\sigma_j)\,,$$

$$\sum_{\sigma}\prod_{j=1}^{N}\exp(-\beta Js\sigma_j) = [\,\exp(\beta Js) + \exp(-\beta Js)\,]^N\,,$$

one obtains the probability distribution,

$$\rho(\sigma) = \prod_{j=1}^{N}\rho(\sigma_j)\,, \tag{64}$$

where

$$\rho(\sigma_j) = \frac{\exp(\beta Js\sigma_j)}{\exp(\beta Js) + \exp(-\beta Js)}\,. \tag{65}$$

For the order parameter (54), one obtains the equation,

$$s = \tanh(\beta J s) \ . \tag{66}$$

When there is mutual disagreement between the agents, so that $J < 0$, the sole solution for the order parameter is $s = 0$, which means a disordered state:

$$s \equiv 0 \qquad (J < 0) \ . \tag{67}$$

However, when the members of the society are in mutual agreement, so that $J > 0$, then there are two solutions to Equation (65). One solution is $s = 0$, but the other solution is nonzero for $T < T_c = J$. In order to choose the stable solution, one needs to find out which of them minimizes the free energy.

It is convenient to work with dimensionless, reduced free energy,

$$\overline{F}(s) \equiv \frac{F}{NJ} \qquad (J > 0) \ , \tag{68}$$

and to measure temperature in units of J, keeping in mind a positive J. Then,

$$\overline{F}(s) = \frac{1}{2}\, s^2 - T \ln[\, 2\cosh(\beta s)\,] \ . \tag{69}$$

The condition of the free-energy extremum is

$$\frac{\partial \overline{F}(s)}{\partial s} = s - \tanh(\beta s) = 0 \ , \tag{70}$$

which gives the order parameter Equation (66). The condition of the free-energy minimum,

$$\frac{\partial^2 \overline{F}(s)}{\partial s^2} = \tanh^2(\beta s) > 0 \ , \tag{71}$$

holds true for nonzero s. It is also not complicated to check that

$$\overline{F}(s) < \overline{F}(0) \qquad (T < T_c) \ , \tag{72}$$

where $\overline{F}(0) = -T\ln 2$. Hence, for temperatures lower than the critical temperature $T_c = J > 0$, the stable state corresponds to the nonzero value of the order parameter s.

At zero temperature $(T = 0)$, with collaborating agents $(J > 0)$, the society is completely ordered:

$$s = \pm 1 \qquad (T = 0 \ , \ J > 0) \ . \tag{73}$$

Although with an unspecified decision: either all deciding "yes" or all choosing the option "no".

One can immediately notice that for any T below T_c there are two solutions, positive and negative, both corresponding to the same free energy $\overline{F}(s) = \overline{F}(-s)$ that does not depend on the sign of the order parameter. This means that, with equal probability, the society can vote "yes" as well as "no". In that sense, the situation is degenerate.

3.6. Enforced Ordering

The yes–no degeneracy can be lifted by imposing an ordering force that acts on the members of the society. The ordering force, or regulation force, acting on the members represents different regulations, such as governmental rules and laws, as well as the society's traditions and habits. The society Hamiltonian, including the ordering force, takes the form,

$$H = -\,\frac{1}{2}\sum_{i \neq j}^{N} J_{ij}\sigma_i\sigma_j - \sum_{j=1}^{N} B_j\sigma_j \ . \tag{74}$$

In what follows, we assume that the regulation is uniform, that is, that the same regulations are applied to all members of the society. This translates into the condition that the ordering force is the same for all agents, such that

$$B_j = B_0 \; . \tag{75}$$

In other words, the laws are the same for everyone.

Resorting to the typical-agent approach yields the Hamiltonian,

$$H = -J(s+h)\sum_{j=1}^{N} \sigma_j + \frac{1}{2}\, Js^2 N \; , \tag{76}$$

with the dimensionless force,

$$h \equiv \frac{B_0}{J} \qquad (J > 0) \tag{77}$$

and the order parameter,

$$s \equiv \langle\, \sigma_j \,\rangle \; . \tag{78}$$

We keep in mind the case of mutual agreement, where $J > 0$.

In the typical-agent approach, the reduced free energy becomes

$$\overline{F}(s) = \frac{1}{2}\, s^2 - T \ln\left[\, 2\cosh\left(\frac{s+h}{T}\right)\,\right] , \tag{79}$$

where temperature is measured in units of J. The order parameter (78) is given by the equation,

$$s = \tanh\left(\frac{s+h}{T}\right) . \tag{80}$$

The order parameter can also be defined as the derivative,

$$s = -\,\frac{\partial \overline{F}(s)}{\partial h} \; . \tag{81}$$

How the society responds to the imposed regulations is described by the susceptibility

$$\chi = \frac{\partial s}{\partial h} = -\,\frac{\partial^2 \overline{F}(s)}{\partial h^2} , \tag{82}$$

which can be represented as

$$\chi = \frac{1}{T}\,\mathrm{var}(s(\sigma)) \; , \tag{83}$$

with the variance

$$\mathrm{var}(s(\sigma)) = \langle\, s^2(\sigma)\,\rangle - \langle\, s(\sigma)\,\rangle^2 \; .$$

Thus, the susceptibility is positive, which means that the imposed regulations increase the order. Explicitly, one finds:

$$\chi = \frac{1 - s^2}{T + s^2 - 1} \; . \tag{84}$$

From the order parameter Equation (80), it is seen that the sign of the order parameter is prescribed by the sign of the ordering force, so that $s > 0$ for $h > 0$ and $s < 0$ for $h < 0$. By choosing the sign of h, it is possible to enforce either the ordering "yes" or the ordering "no". For concreteness, we consider $h > 0$. If the ordering force is extremely strong, then

$$s \simeq 1 \qquad (h \to \infty) \; . \tag{85}$$

In that way, the imposed ordering force increases the order in the society and makes it more stable. For illustration, let us consider the case of no noise; hence, $T = 0$. Then, the reduced free energy coincides with the reduced energy,

$$\overline{E} \equiv \frac{E}{JN} = -\frac{1}{2} - h \qquad (T = 0) . \tag{86}$$

As one can see, the imposed force diminishes the societal energy, which assumes that the society should be more stable.

3.7. Command Economy

From Section 3.6, one may conclude that the stricter the regulations, the more ordered the society. That is, the larger the force h, the larger the order parameter s. The seeming conclusion could be that it is profitable to make the regulations as stringent as possible. Is this so? Let us consider, as an example, an economic society. The most strictly regulated type of economic organization is command economy or centrally planned economy. Is such over-regulated economics the most efficient kind?

The basic points of command economy are:

(i) A centralized government owns most means of production and most businesses.
(ii) The government controls production levels and distribution quotas.
(iii) The government controls all prices and salaries.

The proclaimed advantages of command economy assume that regulatory decisions are made for the benefit of the whole society, there is no large economic inequality, and the economy is claimed to be more stable. However, in reality, the proclaimed catchwords contain many problems:

1. It is not always well defined what the benefits of the society are. Governmental decision-makers often make decisions in their own favor, but not in favor of the society, announcing their egoistic goals as the society's objective needs.
2. It is actually impossible to formulate correct plans for all goods in the long term. Constant shortages of necessary goods and surpluses of unnecessary goods are a rule. Economy is in a permanent crisis.
3. Since the conditions for long-term future production cannot be exactly predicted, the plans are never accomplished. There is a necessity for correcting the plans, with additional spending for the plans' corrections.
4. Since everything is planned in advance, it is almost impossible to introduce innovations that have not been planned. As a result, technological retardation is imminent.
5. The kinds of science to develop are also planned. Some sciences are mistakenly announced as unnecessary or wrong. Examples are cybernetics and genetics in the Soviet Union. This results in irreparable harm to the economy.
6. It is impossible to absolutely fairly distribute wealth. Those who distribute always take more for themselves. Consequently, people are dissatisfied. Bureaucratic corruption flourishes, leading to enormous economic losses.
7. Since wealth is rigidly distributed by the government, there is no reason to work hard. Labor becomes inefficient, with low productivity.
8. The necessity of having huge planning institutions consumes a large amount of economic means. Ineffective planning reduces the planners to the state of parasites, merely wasting resources.
9. The necessity of having a large number of controllers for implementing economic plans and controlling their accomplishment also reduces such controllers to parasites.
10. The suppression of economic freedom, ascribed to the needs of the state but often needed to protect the privileges of the country's rulers, kills the motivation of people to work well.
11. As soon as there appears frustration of the suppressed citizens, there is a necessity to have excessively large police and regulating services supervising the society.

12. To support the order and punish those who are against it, who hesitate, or who could even be potentially dangerous, the government organizes massive suppression of the people, which results in large economic losses.
13. To realize unreasonable plans, the government practices mass arrests of innocent people to create slave labor. However, slavery is not economically efficient.
14. To distract the discontentment of the population from economic failure, a necessity arises to invent enemies, which requires a large army, which consumes a substantial amount of the country's wealth.
15. In order to persuade people that everything is all right, it is necessary to organize propaganda through mass media, which results in ineffective spending of resources.

More details on societies with a command economy can be found in the literature [46]. Due to so many factors that make the economy of an over-regulated society ineffective, it does not seem feasible that enforcing regulations unlimitedly could result in an ideally ordered society. It seems that the model of a regulated society, described in the previous section, does not take into account some factors that prevent an indefinite increase in order by over-regulating the society.

3.8. Regulation Cost

The problem with regulation is that it requires the use of societal resources. Regulations are costly. The more strict regulations there are, the more resources required. All regulations imposed on a society are produced by a part of the same society, that is, by the society itself. Some members of the society act on other members. The regulation cost is the cost that the society has to pay in order to introduce the desired regulations. The regulation cost, keeping in mind the above, can be modeled by the expression

$$H_{\rm reg} = \frac{1}{2} \sum_{i \neq j}^{N} A_{ij} B_i B_j \sigma_i \sigma_j \; , \tag{87}$$

where $A_{ij} = A_{ji} > 0$. The coefficients A_{ij} describe the regulation efficiency. By the order of magnitude, they are proportional to the interaction, J_{ij}, between the agents since regulations need to overcome mutual interactions in order to impose restrictions. Note that J_{ij} and A_{ij} are different types of interactions: one is a direct interaction not related to the process of ordering, and the other is an interaction inducing the ordering in the presence of an additional force. Different types of interactions, in general, are different.

Thus, the total Hamiltonian becomes

$$H = -\frac{1}{2} \sum_{i \neq j}^{N} J_{ij} \sigma_i \sigma_j - \sum_{i=1}^{N} B_i \sigma_i + \frac{1}{2} \sum_{i \neq j}^{N} A_{ij} B_i B_j \sigma_i \sigma_j \; . \tag{88}$$

Let us again assume that the force acting on the society members is uniform in the sense of equality (75).

Now, the order parameter (78) differs from the expression,

$$M = -\frac{\partial \overline{F}}{\partial h} = s - \frac{1}{N} \sum_{i \neq j}^{N} \alpha_{ij} h \langle \, \sigma_i \sigma_j \, \rangle \; , \tag{89}$$

with the notation,

$$\alpha_{ij} \equiv A_{ij} J \; . \tag{90}$$

The expression (89) plays the role of the total order parameter, contrary to s, which, in the present case, is a partial order parameter.

In the typical-agent approach, Hamiltonian (88) reads:

$$H = -(Js + B_0 - AB_0^2 s) \sum_{i=1}^{N} \sigma_i + \frac{1}{2}\left(J - AB_0^2\right) s^2 N \; , \tag{91}$$

where

$$A \equiv \frac{1}{N} \sum_{i \neq j}^{N} A_{ij} \; . \tag{92}$$

The reduced free energy takes the form,

$$\overline{F} = \frac{1}{2}\left(1 - \alpha\, h^2\right) s^2 - T \ln\left\{ 2 \cosh\left[\frac{(1 - \alpha h^2) s + h}{T} \right] \right\} , \tag{93}$$

in which

$$\alpha \equiv \frac{1}{N} \sum_{i \neq j}^{N} \alpha_{ij} = AJ \; . \tag{94}$$

The parameters α and h are assumed to be positive.

The order parameter, $s = \langle \sigma_j \rangle$, can be obtained from the condition,

$$\frac{\partial \overline{F}}{\partial s} = 0 \tag{95}$$

leading to the equation,

$$s = \tanh\left[\frac{(1 - \alpha h^2) s + h}{T} \right] . \tag{96}$$

However, in the presence of a Hamiltonian of a quadratic, with respect to the ordering force term, the quantity s does not define the total order that is given by the form (89), which yields the total order parameter

$$M = -\frac{\partial \overline{F}}{\partial h} = s - \alpha h s^2 \; . \tag{97}$$

The susceptibility is defined by the derivative,

$$\chi \equiv \frac{\partial M}{\partial h} = -\frac{\partial^2 M}{\partial h^2} , \tag{98}$$

resulting in the expression,

$$\chi = \frac{(1 - s^2)(1 - 2\alpha h s)^2}{T + (1 - s^2)(\alpha h^2 - 1)} , \tag{99}$$

which differs from the derivative,

$$\frac{\partial s}{\partial h} = \frac{(1 - s^2)(1 - 2\alpha h s)}{T + (1 - s^2)(\alpha h^2 - 1)} . \tag{100}$$

The behavior of the social system at small and large regulating forces shows how the order parameters and free energy vary. The analysis of the asymptotic behavior demonstrates the relation between the system's stability and its order. This relation can be rather complicated depending on the system parameters.

With a small ordering force and weak noise, when $0 < T < 1$, the order parameter s behaves as

$$s \simeq s_0 + \frac{1 - s_0^2}{s_0^2 + T - 1}\, h \qquad (0 < T < 1 \; , \; h \to 0) \; , \tag{101}$$

where s_0 is the solution to the equation,

$$s_0 = \tanh\left(\frac{s_0}{T}\right) . \tag{102}$$

The total order parameter, M, with a small ordering force and weak noise behaves as

$$M \simeq s_0 + \left(\frac{1-s_0^2}{s_0^2+T-1} - \alpha s_0^2\right)h \qquad (0 < T < 1 \, , \; h \rightarrow 0) \, , \tag{103}$$

which shows that the order parameters can either increase or decrease with rising h, depending on the system parameters.

The free energy with a small ordering force and weak noise is

$$\overline{F} \simeq \frac{s_0^2}{2} - T\ln\left[2\cosh\left(\frac{s_0}{T}\right)\right] - s_0 h \qquad (0 < T < 1 \, , \; h \rightarrow 0) \, . \tag{104}$$

For $T > 1$ and weak h, the order parameters are

$$s \simeq \frac{1}{T-1}\, h - \frac{3\alpha(T-1)^2+T}{3(T-1)^4}\, h^3 \, ,$$

$$M \simeq \frac{1}{T-1}\, h - \left[\frac{3\alpha(T-1)^2+T}{3(T-1)^4} + \frac{\alpha}{(T-1)^2}\right]h^3 \qquad (T > 1 \, , \; h \rightarrow 0) \, . \tag{105}$$

The free energy reads:

$$\overline{F} \simeq -T\ln 2 - \frac{1}{2(T-1)}\, h^2 \qquad (T > 1 \, , \; h \rightarrow 0) \, . \tag{106}$$

When the regulating force is strong, then, at any temperature, the order parameters are

$$s \simeq \frac{1}{\alpha h} - \frac{T-1}{\alpha^2 h^3} \, , \qquad M \simeq \frac{2(T-1)}{\alpha^2 h^3} \qquad (h \rightarrow \infty) \, , \tag{107}$$

and the free energy,

$$\overline{F} \simeq -T\ln 2 - \frac{1}{2\alpha} + \frac{T-1}{2\alpha^2 h^2} \qquad (h \rightarrow \infty) \, . \tag{108}$$

In the absence of noise, that is, when $T = 0$, the free energy equals the energy. Considering the reduced energy,

$$\overline{E} \equiv \frac{\langle H \rangle}{NJ} = -\frac{1}{2}\, s^2 - hs + \frac{1}{2}\, \alpha h^2 s^2 \, , \tag{109}$$

at zero T, one obtains:

$$\overline{E} = -\frac{1}{2} - h + \frac{1}{2}\, \alpha h^2 \qquad (T = 0) \, . \tag{110}$$

As one can see, the energy (110) with switching on regulations, first decreases, making the society more stable,

$$\overline{E} \simeq -\frac{1}{2} - h \qquad (T = 0 \, , \; h \rightarrow 0) \, ; \tag{111}$$

then, it reaches the minimum,

$$\min_h \overline{E} = -\frac{1}{2}\left(1 + \frac{1}{\alpha}\right) \qquad \left(h = \frac{1}{\alpha}\right) , \tag{112}$$

after which it increases, making the society less and less stable:

$$\overline{E} \simeq \frac{1}{2}\,\alpha h^2 \qquad (T = 0\;,\; h \to \infty)\;. \tag{113}$$

Thus, in the absence of noise, the optimal regulation, making the society the most stable, corresponds to $h = 1/\alpha$, when the energy (112) is minimal. This implies that some amount of regulation is useful in stabilizing the society. However, the society should not be over-regulated. Over-regulation can make the society unstable.

In the presence of noise, the situation is much more complicated, depending on the noise strength and the value of the regulating force. The peculiarity of the system's behavior is illustrated in Figures 1 and 2. Figure 1 shows the behavior of the free energy $\overline{F}$ as a function of the regulation force h for different temperatures (noise intensity), when $T < 1$ in Figure 1a and when $T > 1$ in Figure 1b. Figure 2 shows the order parameter s as a function of the regulation force h for different temperatures: low $(T < 1)$ and high $(T > 1)$. As one can see, at low temperatures, the free energy first decreases and then increases, similarly to the behavior of energy (109). The order parameter first increases but later decreases. That is, there exists an optimal regulation strength when the society is the most stable and at the same time the most ordered. In the presence of strong noise, the free energy decreases with h; however, the order parameter first increases but then decreases. So, the regulating force stabilizes the society but improves the order only at a limited value of h, while at a quite strong force, the order diminishes.

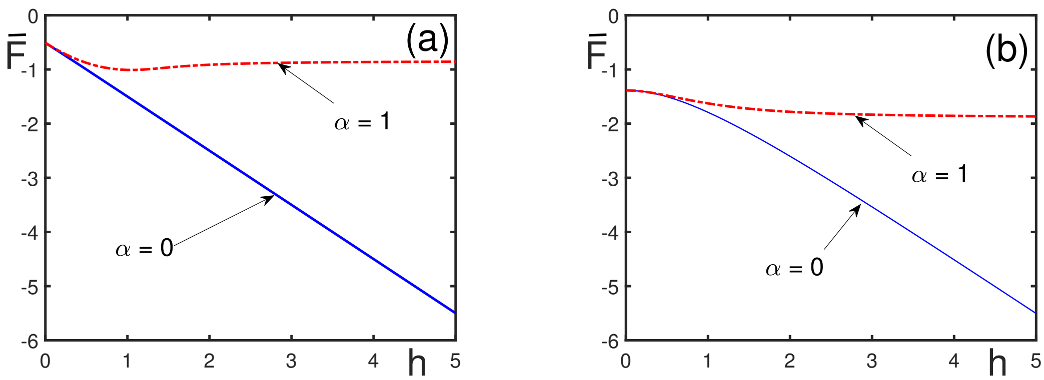

Figure 1. Yes–no model with regulation cost. Behavior of the free energy, F, as a function of the regulation force, h, for different temperatures (noise intensity): (**a**) $T < 1$ and (**b**) $T > 1$.

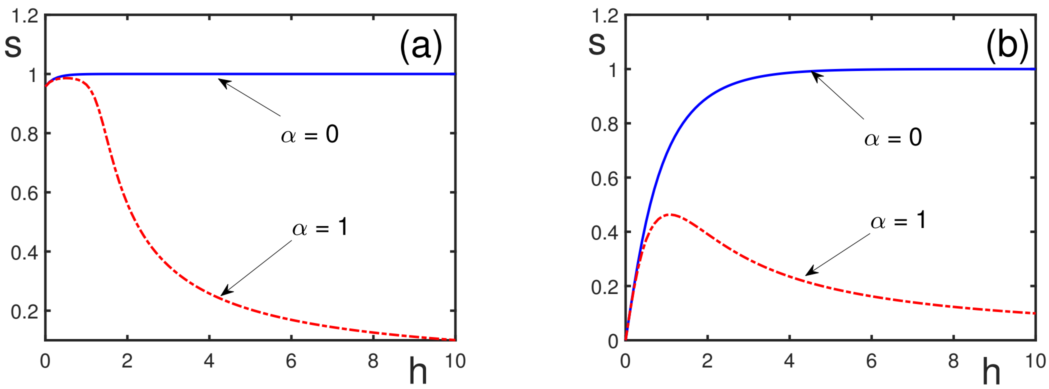

Figure 2. Yes–no model with regulation cost. The order parameter, s, as a function of the regulation force, h, for (**a**) low $(T < 1)$ and (**b**) high $(T > 1)$ temperatures.

3.9. Fluctuating Groups

It is customary for society members to separate into several groups with different properties, so that the number of members in a group is not constant but varies with time. For example, these could be groups of opponents to the government, workers on strike, opposing organizations, and the like. Different groups of agents can be described by different order parameters. The groups can be localized in some spatial parts of the society or may move through the whole societal space. The groups are not permanent in their agent numbers and do not necessarily exist forever, but they can arise and vanish. In that sense, the groups fluctuate; they can appear, change, and then disappear.

Let the number of agents in a fluctuating group be $N_{\rm f}$ and the characteristic time during which they do not essentially change be $t_{\rm f}$. Fluctuating groups are mesoscopic, at least in one of the following senses.

Mesoscopic in size:

$$1 \ll N_{\rm f} \ll N \; . \tag{114}$$

The number $N_{\rm f}$ is much larger than can be used to define the group order parameter, and $N_{\rm f}$ is much smaller than the total number of agents, N, in the society to be classified as a group inside that society.

Mesoscopic in time:

$$t_{\rm int} \ll t_{\rm f} \ll t_{\rm exp} \; , \tag{115}$$

where $t_{\rm int}$ is the characteristic interaction time between the agents and $t_{\rm exp}$ is the observation (experiment) time during which the society is studied. The time $t_{\rm f}$ has to be much larger than the interaction time in order that the groups as such could be formed. Additionally, $t_{\rm f}$ is much smaller than the observation time for the groups to be classified as fluctuating. The general method of describing this kind of society that contains fluctuating groups is presented in this section.

Let us consider a snapshot of a society picture, where the spatial location of a j-th agent of a society is denoted by a vector $\mathbf{a}_j$. The overall society containing all its members is given by the collection,

$$\mathbb{G} = \left\{ \mathbf{a}_j : \; j = 1,2,\ldots,N \right\} \; . \tag{116}$$

The society consists of several groups, each characterized by a specific feature. The features are enumerated by the index $f = 1,2,\ldots$ A group with an f-th feature is

$$\mathbb{G}_f = \left\{ \mathbf{a}_j \in \mathbb{G}_f : \; j = 1,2,\ldots,N_f \right\} \; . \tag{117}$$

The union of all groups forming the whole society is

$$\bigcup_f \mathbb{G}_f = \mathbb{G} \; , \qquad \sum_f N_f = N \; . \tag{118}$$

The spatial locations of the groups is represented by the manifold indicator functions,

$$\xi_f(\mathbf{a}_j) = \left\{ \begin{array}{ll} 1 \; , & \mathbf{a}_j \in \mathbb{G}_f \; , \\ 0 \; , & \mathbf{a}_j \not\in \mathbb{G}_f \; . \end{array} \right. \tag{119}$$

The collection of all indicator functions, showing the society configuration, is denoted as

$$\xi = \{ \xi_f(\mathbf{a}_j) : \; \mathbf{a}_j \in \mathbb{G} \; , \; f = 1,2,\ldots \} \; . \tag{120}$$

The indicator functions possess the properties:

$$\sum_f \xi_f(\mathbf{a}_j) = 1 \; , \qquad \sum_j \xi_f(\mathbf{a}_j) = N_f \; , \tag{121}$$

the first of which means that each agent pertains to some of the groups, while the second shows the number of agents in an f-th group.

In a realistic society, the agents can change their locations, so the location vector $\mathbf{a}_j = \mathbf{a}_j(t)$ depends on time t. Hence, the societal configuration is also a function of time,

$$\xi(t) = \{\xi_f(\mathbf{a}_j(t)) : \mathbf{a}_j(t) \in \mathbb{G} \; ; \; f = 1,2,\ldots\} \; . \tag{122}$$

Since $t_{\rm exp} \gg t_{\rm int}$, the observable quantities describing the society correspond to the double average over the system variables and over time,

$$\langle \, A(\sigma,\xi(t)) \, \rangle = \sum_\sigma \frac{1}{t_{\rm exp}} \int_0^{t_{\rm exp}} \rho(\sigma,\xi(t)) A(\sigma,\xi(t)) \, dt \; . \tag{123}$$

The motion of the groups inside the society is generally so complicated that it is neither possible nor reasonable to follow the detailed movements of all agents, but the agent locations can be treated as random. Then, it is possible to interchange the average over time with the average over the random society configurations by the rule,

$$\frac{1}{t_{\rm exp}} \int_0^{t_{\rm exp}} dt \longmapsto \int \mathcal{D}\xi \; . \tag{124}$$

Then, one needs to realize the functional integration over the manifold indicator functions. We do not describe the mathematical details of the integration that can be found in the reviews [47,48], but present the results.

The probability distribution $\rho(\sigma,\xi)$ of a heterophase society, depending on the configuration, ξ, can be derived following Section 2.2. The probability is to be normalized:

$$\sum_\sigma \int \rho(\sigma,\xi) \; \mathcal{D}\xi = 1 \; . \tag{125}$$

The average energy is given by the expression,

$$\sum_\sigma \int \rho(\sigma,\xi) H(\sigma,\xi) \; \mathcal{D}\xi = E \; . \tag{126}$$

The information functional takes the form,

$$I[\, \rho(\sigma,\xi) \,] = \sum_\sigma \int \rho(\sigma,\xi) \; \ln \; \frac{\rho(\sigma,\xi)}{\rho_0(\sigma,\xi)} \; \mathcal{D}\xi$$
$$+ \; \lambda_0 \left[\sum_\sigma \int \rho(\sigma,\xi) \; \mathcal{D}\xi - 1 \right] + \lambda \left[\sum_\sigma \int \rho(\sigma,\xi) \; H(\sigma,\xi) \; \mathcal{D}\xi - E \right] . \tag{127}$$

Minimizing the information functional and assuming the uniformity of the trial distribution, $\rho_0(\sigma,\xi) = {\rm const}$, yields the probability,

$$\rho(\sigma,\xi) = \frac{1}{Z} \, e^{-\beta H(\sigma,\xi)} \; , \tag{128}$$

with the partition function,

$$Z = \sum_\sigma \int e^{-\beta H(\sigma,\xi)} \; \mathcal{D}\xi \; . \tag{129}$$

The free energy reads:

$$F = -T \ln Z = -T \ln \sum_\sigma \int e^{-\beta H(\sigma,\xi)} \; \mathcal{D}\xi \; . \tag{130}$$

Observable quantities are given by the averages,

$$\langle\, A(\sigma,\xi)\,\rangle = \sum_\sigma \int \rho(\sigma,\xi) A(\sigma,\xi)\; \mathcal{D}\xi \; . \tag{131}$$

Accomplishing the average over configurations allows us to define an effective Hamiltonian, $\widetilde{H}$, by the relation,

$$e^{-\beta\widetilde{H}} = \int e^{-\beta H(\sigma,\xi)}\; \mathcal{D}\xi \; . \tag{132}$$

The effective Hamiltonian takes the form (with the direct summation $\bigoplus$),

$$\widetilde{H} = \bigoplus_f H_f \; , \qquad H_f = H_f(\sigma, w_f) \; , \tag{133}$$

where the probability that an agent pertains to a j-th group is

$$w_f \equiv \frac{N_f}{N} \; . \tag{134}$$

In other words, this is the fraction of the society agents belonging to the f-th group. Certainly, the normalization conditions are valid:

$$\sum_f w_f = 1 \; , \qquad 0 \leq w_f \leq 1 \; . \tag{135}$$

The group probabilities are the minimizers of the free energy,

$$F = -T \ln \sum_\sigma e^{-\beta\widetilde{H}} \; , \tag{136}$$

under the normalization conditions (135),

$$F = \text{abs}\min F(w_1, w_2, \ldots) \; . \tag{137}$$

The observable quantities are given by the averages

$$\langle\, A(\sigma,\xi)\,\rangle = \sum_f \langle\, A_f(\sigma, w_f)\,\rangle \; , \tag{138}$$

where

$$\langle\, A_f(\sigma,w_f)\,\rangle = \sum_\sigma \rho_f(\sigma,w_f) A_f(\sigma,w_f) \; , \tag{139}$$

with the effective probability distributions,

$$\rho_f(\sigma,w_f) = \frac{1}{Z_f}\; e^{-\beta H_f(\sigma,w_f)} \; , \tag{140}$$

and the partition functions,

$$Z_f = \sum_\sigma e^{-\beta H_f(\sigma,w_f)} \; . \tag{141}$$

The sum of direct summation $\bigoplus$ in Equation (133) is used instead of the simple sum in order to stress that the summands are defined on different configuration sets,

$$X_f = \{ \sigma_{fj} : \; j = 1, 2, \ldots, N_f \} \; , \tag{142}$$

with the total society configuration set being the tensor product,

$$X_f = \bigotimes_f X_f \; . \tag{143}$$

This is the general approach for describing heterophase societies with fluctuating groups. The approach reduces the consideration of quasi-equilibrium systems with mesoscopic group fluctuations to the description of effective equilibrium systems with a renormalized effective Hamiltonian. The mathematics of accomplishing functional integration over manifold indicator functions is described in reviews [21,47,48]. More mathematical details are presented in Refs. [49–51]. Section 3.10 just below gives an example of applying this approach.

3.10. Self-Organized Disorder

Let us consider the yes–no model of Section 3.5 modified to take account of fluctuating groups. For simplicity, two groups have to be kept in mind: one ordered and the other disordered. The ordered group consists of agents agreeing with each other and is characterized by a nonzero order parameter, while the disordered group is described by a zero order parameter to be defined in this Section below.

Each j-th agent from an f-th group is characterized by the features σ_{fj}. In the yes–no model, these features are described by the variable $\sigma_{fj} = \pm 1$. The total feature set for the whole system is the configuration set (143).

In addition to the interaction $J_{ij} > 0$, typical for self-organization in the yes–no model, one needs to include other interactions that would describe some disagreements between the agents. Let us denote the disordering interaction by U_{ij}.

In the snapshot picture, the Hamiltonian with ordering and disordering interactions reads as

$$H(\sigma,\xi) = H_1(\sigma,\xi_1) \bigoplus H_2(\sigma,\xi_2) \; , \tag{144}$$

with

$$H_f(\sigma,\xi_f) = \frac{1}{2}\sum_{i\neq j}^N \xi_f(\mathbf{a}_i)\xi_f(\mathbf{a}_j)U_{ij} - \frac{1}{2}\sum_{i\neq j}^N \xi_f(\mathbf{a}_i)\xi_f(\mathbf{a}_j)J_{ij}\sigma_{fi}\sigma_{fj} \; . \tag{145}$$

Here, the manifold indicator functions show the belonging of the agents to the corresponding groups.

After averaging over group configurations, one obtains the effective Hamiltonian,

$$\widetilde{H} = H_1 \bigoplus H_2 \; , \qquad H_f = H_f(\sigma,w_f) \; ,$$

$$H_f(\sigma,w_f) = \frac{1}{2}\, w_f^2 \sum_{i\neq j}^N (U_{ij} - J_{ij}\sigma_{fi}\sigma_{fj}) \; . \tag{146}$$

The order parameter for the f-th group is

$$s_f \equiv \frac{1}{N}\sum_{j=1}^N \langle\, \sigma_{fj}\, \rangle \; , \tag{147}$$

which translates into the equation,

$$s_f = \tanh\left(\frac{Jw_f^2 s_f}{T}\right) \; . \tag{148}$$

Let the first group be ordered, so that

$$s_1 \neq 0 \; , \tag{149}$$

while the second group be disordered, in the sense that

$$s_2 \equiv 0 \; . \tag{150}$$

In the case of two groups, it is convenient to denote the fraction of agents in the ordered group and disordered group as

$$w_1 \equiv w \; , \qquad w_2 = 1 - w \; . \tag{151}$$

Then, the necessary condition for the free-energy minimum is

$$\frac{\partial F}{\partial w} = \left\langle \frac{\partial \widetilde{H}}{\partial w} \right\rangle = 0 \; . \tag{152}$$

This results in the equation,

$$w = \frac{u - s_2^2}{2u - s_1^2 - s_2^2} \; , \tag{153}$$

in which the notation is used:

$$u \equiv \frac{U}{J} \; , \qquad U \equiv \frac{1}{N} \sum_{i \neq j}^{N} U_{ij} \; , \qquad J \equiv \frac{1}{N} \sum_{i \neq j}^{N} J_{ij} \; . \tag{154}$$

Taking into account that the disordered group is described by the zero-order parameter $s_2 = 0$, one has the fraction of agents in the ordered group:

$$w = \frac{u}{2u - s_1^2} \; . \tag{155}$$

Due to the probability definition $0 \leq w \leq 1$, and because $0 \leq s_1 \leq 1$, expression (155) exists only for sufficiently strong disordering interactions, such that $u \geq 1$.

Note that the second derivative of the free energy, being positive, leads to the inequality

$$\frac{\partial^2 F}{\partial w^2} = \left[\left\langle \frac{\partial^2 \widetilde{H}}{\partial w^2} \right\rangle - \beta \left\langle \left(\frac{\partial \widetilde{H}}{\partial w} \right)^2 \right\rangle \right] \; > \; 0 \; , \tag{156}$$

which yields the condition $u > 1/2$.

The societal reduced energy,

$$\overline{E} \equiv \frac{\langle \widetilde{H} \rangle}{JN} \; , \tag{157}$$

leads to

$$\overline{E} = \overline{E}_1 + \overline{E}_2 \; , \qquad E_f = \frac{1}{2} \, w_f^2 \left(u - s_f^2 \right) \; . \tag{158}$$

For simplicity, let us study the case of no noise, so that $T = 0$. Then, $s_1 = 1$, and the energy becomes

$$\overline{E} = w^2 \left(u - \frac{1}{2} \right) + \left(\frac{1}{2} - w \right) u \; , \tag{159}$$

while the fraction of ordered agents is

$$w = \frac{u}{2u - 1} \; . \tag{160}$$

Thus, the society's energy is

$$\overline{E} = \frac{u(u-1)}{2(2u-1)} \; . \tag{161}$$

Now, one has to decide which society is more stable: that which contains fluctuating disordered groups or that which does not contain them? The more stable is the society whose free energy is lower. In the case of small noise, one needs to compare the corresponding expected energies.

If no fluctuating groups were allowed, then the ordered fraction would be exactly one ($w \equiv 1$). The energy of a completely ordered society is

$$E_{\rm ord} \equiv \overline{E}(w=1) = \frac{u-1}{2} \; , \tag{162}$$

where $w \equiv 1$. If the whole society contained only disordered agents ($w \equiv 0$), then the societal energy would be

$$E_{\rm dis} \equiv \overline{E}(w=0) = \frac{1}{2}\, u \; . \tag{163}$$

The society with disordered group fluctuations is more stable when its energy is lower. Comparing the above expressions, one sees that

$$\overline{E} < E_{\rm ord} < E_{\rm dis} \qquad (u > 1) \; . \tag{164}$$

Thus, for competing interactions, satisfying the condition $u > 1$, the energy of the society with fluctuating disorder is lower than that of the perfectly ordered society, which, in turn, is lower than the disordered society. That is, a completely ordered society is more stable than a disordered society. However, the most stable is the society with a mixture of fluctuating disorder. Thus, disorder fluctuations make the society more stable. In that sense, the disorder is self-organized.

The total order parameter is defined as the average of the expression

$$M(\sigma,\xi) = \frac{1}{N}\sum_f \sum_{j=1}^N \xi_f(\mathbf{a}_j)\sigma_{fj} \; , \tag{165}$$

which gives

$$M = \langle\, M(\sigma,\xi)\, \rangle = \sum_f w_f s_f \; . \tag{166}$$

When there is no noise, that is at zero temperature, one finds:

$$M = w s_1 = \frac{u}{2u-1} \qquad (T=0) \; . \tag{167}$$

Concluding, under sufficiently strong competition between the society's agents, there spontaneously appear social disorder groups. When the disorder groups appear, they diminish the total society order. However, the existence of the disorder groups makes the society more stable. This is an example that shows that it is not always useful to try to realize a complete order in a society since the presence of some disorder can make the society more stable. Sometimes, societies that are more stable generate self-organized disorder.

3.11. Coexistence of Populations

Quite often, social systems consist of several groups of rather different people, so the groups are more or less stable and, on average, do not essentially change for a long time, contrary to fluctuating groups considered in the previous section. Each of the groups can be characterized by different typical agents. There are numerous examples of such societies. Thus, a society, forming a country, can often include groups of different nationalities or religions.

In a society composing a financial market, there are groups of fundamentalists and chartists. Fundamentalists base their decisions upon market fundamentals, such as interest rates, the growth or decline of the economy, companies' performance, etc. Fundamentalists expect asset prices to move towards their fundamental values, hence, they either buy or sell assets that are assumed to be undervalued or, respectively, overvalued. Chartists, or technical analysts, look for patterns and trends in past market prices and base their decisions upon the extrapolation of these patterns. There exist as well contrarian groups that buy or sell contrary to the trend followed by the majority.

Different groups of people or other living beings possessing differing features are called populations. Let a society, inhabiting the volume V, contain several different populations enumerated by the index $n = 1, 2, \ldots$. In an n-th population, there are N_n members. In what follows, we consider a general approach that describes whether the populations can or cannot coexist with each other. The populations can be of any origin. For concreteness, we can keep in mind the different population groups living in a country.

The populations can either peacefully live in the whole country, which can be called a mixed society, or can possess the tendency to separate, thus having no intention for joint coexistence. Our aim is to understand how one could quantify the conditions when different populations prefer to live together and when they wish to separate.

According to the general law, a society is more stable whose free energy is lower. That is, a mixed society living in the same country is more stable than separated populations existing separately from each other, if the free energy, $F_{\rm mix}$, of the mixed society is lower than the free energy, $F_{\rm sep}$, of the separated society:

$$F_{\rm mix} < F_{\rm sep} \; . \tag{168}$$

The free energies can be represented as

$$F_{\rm mix} = E_{\rm mix} - TS_{\rm mix} \; , \qquad F_{\rm sep} = E_{\rm sep} - TS_{\rm sep} \; . \tag{169}$$

Therefore, denoting the entropy of mixing

$$\Delta S_{\rm mix} \equiv \frac{1}{N} \left(S_{\rm mix} - S_{\rm sep} \right) \; , \tag{170}$$

one has the condition of stability of the united country:

$$E_{\rm mix} - E_{\rm sep} < NT\Delta S_{\rm mix} \; . \tag{171}$$

We need to write down the energy of the mixed and separated populations. To this end, let us denote the density of an n-th population by $\rho_n(\mathbf{r})$ and the interaction between the members of an m-th and n-th populations by $\Phi_{mn}(\mathbf{r})$, where $\mathbf{r}$ represents the member's location. Then, the interaction energy for the mixed populations can be represented as

$$E_{\rm mix} = \frac{1}{2} \sum_{mn} \int_V \rho_m(\mathbf{r}) \; \Phi_{mn}(\mathbf{r} - \mathbf{r}') \; \rho_n(\mathbf{r}') \; d\mathbf{r} d\mathbf{r}' \; . \tag{172}$$

Assuming a uniform distribution of the people across the country implies the unform densities,

$$\rho_m(\mathbf{r}) = \frac{N_m}{V} \; , \qquad \rho_n(\mathbf{r}) = \frac{N_n}{V} \; . \tag{173}$$

Thus, the interaction energy of a country with mixed populations reads as

$$E_{\rm mix} = \frac{1}{2} \sum_{mn} \Phi_{mn} \; \frac{N_m N_n}{V} \; , \tag{174}$$

where the quantity

$$\Phi_{mn} \equiv \int_V \Phi_{mn}(\mathbf{r}) \; d\mathbf{r} \tag{175}$$

describes the average interaction strength between the members of m-th and n-th populations.

Now, let us consider the case of separated populations, where each population lives in a separate location characterized by the volume V_n. Then, the energy of a separated country is the sum,

$$E_{\rm sep} = \frac{1}{2} \sum_n \int_{V_n} \rho_n^{\rm sep}(\mathbf{r}) \; \Phi_{nn}(\mathbf{r}-\mathbf{r}') \; \rho_n^{\rm sep}(\mathbf{r}') \; d\mathbf{r}d\mathbf{r}' \; . \tag{176}$$

Again keeping in mind that each population inside their part of the country is uniformly distributed gives the uniform densities,

$$\rho_n^{\rm sep}(\mathbf{r}) = \frac{N_n}{V_n} \; . \tag{177}$$

Then, the energy of a separated country is

$$E_{\rm sep} = \frac{1}{2} \sum_n \overline{\Phi}_{nn} \; \frac{N_n^2}{V_n} \; , \tag{178}$$

where

$$\overline{\Phi}_{nn} \equiv \int_{V_n} \Phi_{nn}(\mathbf{r}) \; d\mathbf{r} \; . \tag{179}$$

Using the identity,

$$\frac{V}{V_m} = \sum_n \frac{V_m}{V_n} \; ,$$

reduces the energy of a separated country to

$$E_{\rm sep} = \frac{1}{2} \sum_{mn} \overline{\Phi}_{mm} \; \frac{N_m^2 N_n}{V V_m} \; . \tag{180}$$

Any two different populations, although separated, exert pressure on each other through their common boundary. If the pressure of one of them is greater than that of the other, there is no equilibrium between the populations, but there can arise nonequilibrium movements, such as invasions and wars. The equilibrium coexistence of the two separated populations implies the equality of their pressures:

$$P_m = P_n \qquad \left(P_n = -\; \frac{\partial F_{\rm sep}}{\partial V_n} \right) \; . \tag{181}$$

By its meaning, this equality can be called a *no-war condition*. For not too strong noise, such that

$$\frac{\partial F_{\rm sep}}{\partial V_m} \cong \frac{\partial E_{\rm sep}}{\partial V_m} = \frac{1}{2} \; \overline{\Phi}_{mm} \left(\frac{N_m}{V_m} \right)^2 \; ,$$

one finds the no-war condition in the form,

$$\frac{\overline{\Phi}_{mm}}{\overline{\Phi}_{nn}} = \left(\frac{N_n V_m}{N_m V_n} \right)^2 \; . \tag{182}$$

Equation (182) shows that two neighboring populations have no war and are in equilibrium with each other only when the signs of the effective interactions of the members inside each population are the same and the interactions satisfy the above conditions. An m-th population and a n-th population can be in equilibrium only when their internal interactions Φ_{mm} and Φ_{nn} are both either positive or negative. If one population has $\Phi_{mm} > 0$, while the members of the other population have $\Phi_{nn} < 0$, then there can be no equilibrium between such neighboring populations.

Taking account of the no-war condition (182) makes it possible to rewrite the energy of the separated country in the form,

$$E_{\text{sep}} = \frac{1}{2} \sum_{mn} \frac{N_m N_n}{V} \sqrt{\overline{\Phi}_{mm} \overline{\Phi}_{nn}} \; . \tag{183}$$

Then, the condition of stability (171) for the mixed society yields the inequality,

$$\sum_{mn} \frac{N_m N_n}{2V} \left(\Phi_{mn} - \sqrt{\overline{\Phi}_{mm} \overline{\Phi}_{nn}} \right) \; < \; NT \Delta S_{\text{mix}} \; . \tag{184}$$

With the identity,

$$N \equiv \frac{1}{N} \sum_{mn} N_m N_n \; ,$$

one finds:

$$\sum_{mn} \frac{N_m N_n}{2V} \left(\Phi_{mn} - \sqrt{\overline{\Phi}_{mm} \overline{\Phi}_{nn}} - \frac{2T}{\rho} \, \Delta S_{\text{mix}} \right) \; < \; 0 \; , \tag{185}$$

where $\rho = N/V$ is the average density of the total population. From Equation (185), the sufficient condition of stability for the mixed country becomes

$$\Phi_{mn} \; < \; \sqrt{\overline{\Phi}_{mm} \overline{\Phi}_{nn}} \; + \; \frac{2T}{\rho} \, \Delta S_{\text{mix}} \; . \tag{186}$$

The entropy of mixture can be represented as

$$\Delta S_{\text{mix}} = - \sum_m n_m \ln n_m \qquad \left(n_m \equiv \frac{N_m}{N} \right) \; . \tag{187}$$

Thus, one obtains the condition of stability for the country with the mixed population, as compared to the country where the populations are separated,

$$\Phi_{mn} \; - \; \sqrt{\overline{\Phi}_{mm} \overline{\Phi}_{nn}} \; < \; - \; \frac{2T}{\rho} \sum_m n_m \ln n_m \; . \tag{188}$$

When this condition is not met, the country with the mixed population is not stable and will disintegrate into several separate countries, each composed of just one population type. There exist numerous examples of countries that have disintegrated into several smaller countries. This destiny, for example, has happened with the Roman Empire, Makedonsky Empire, British Empire, French Empire, Austrian Empire, Ottoman Empire, Soviet Union, Yugoslavia, and Czechoslovakia.

3.12. Forced Coexistence

From history, it is known that empires are generally kept united not merely by the economic advantage of the countries composing them but also by force. Why is this possible, and what happens when the required force becomes too strong? Is it possible to keep different populations together inside one empire by increasing force? To answer these questions, it is necessary to consider the case of imposed regulations, including regulation cost.

Let the regulation force, acting on a member of an n-th population be $f_n(\mathbf{r})$. The energy of a society with mixed populations, including the force applied to keep different populations together, can be written as

$$E_{\text{mix}} = \frac{1}{2} \sum_{mn} \int_V \rho_m(\mathbf{r}) \; \Phi_{mn}(\mathbf{r} - \mathbf{r}') \; \rho_n(\mathbf{r}') \; d\mathbf{r} d\mathbf{r}' - \sum_n \int_V f_n(\mathbf{r}) \; \rho_n(\mathbf{r}) \; d\mathbf{r} +$$

$$+\frac{1}{2}\sum_{mn}\int_V \rho_m(\mathbf{r})\; f_m(\mathbf{r})\; A_{mn}(\mathbf{r}-\mathbf{r}')\; f_n(\mathbf{r}')\; \rho_n(\mathbf{r}')\; d\mathbf{r}d\mathbf{r}'\;, \tag{189}$$

where the last term characterizes the cost of supporting the regulation force, with $A_{mn}(\mathbf{r})$ being the strength of the interaction between the members of the society caused by the force applied to them.

It is possible to assume that the population densities are uniform and the force is constant, so that

$$\rho_n(\mathbf{r}) = \frac{N_n}{V}\;, \qquad f_n(\mathbf{r}) = f_n\;. \tag{190}$$

Then, the societal energy reads:

$$E_{\rm mix} = \frac{1}{2}\sum_{mn}(\Phi_{mn} + A_{mn}f_m f_n)\;\frac{N_m N_n}{V} - \sum_n f_n N_n\;, \tag{191}$$

where

$$A_{mn} \equiv \int_V A_{mn}(\mathbf{r})\; d\mathbf{r}\;. \tag{192}$$

When the society is disintegrated into several independent countries, there is no need to apply force; hence, the energy of a society with separated populations is Equation (183). Using the identity,

$$\sum_n f_n N_n = \frac{1}{2N}\sum_{mn} N_m N_n(f_m + f_n)$$

and following the same way as in Section 3.11, one obtains the sufficient condition for the society with mixed populations

$$\Phi_{mn} + A_{mn}f_m f_n \;<\; \sqrt{\overline{\Phi}_{mm}\overline{\Phi}_{nn}} \;+\; \frac{2}{\rho}\left(T\Delta S_{\rm mix} + \frac{f_m + f_n}{2}\right)\;. \tag{193}$$

For brevity, condition (193) can be named the condition of empire stability. If the condition (193) does not hold, the empire disintegrates into separate countries.

The condition (193) shows that switching on the regulating force first increases the right-hand side of the inequality, thus making the mixed society more stable. However, quite a strong force makes the left-hand side of the inequality larger, thus leading to the instability of the empire and its disintegration. Hence, by enforcing reasonable regulations, it is possible to maintain the coexistence of populations even when, without such an enforcement, the country would disintegrate. However, the regulating force should not be quite strong. This explains why, for some time, an empire can exist as a united family of different populations. However, when enforcement is too costly, it becomes unsupportable, and disintegration becomes inevitable. Empires do not exist forever.

An empire with forced coexistence of different populations is almost always less stable than several independent countries with their own populations and weaker regulation forces. However, the unions of different populations can also exist when they are kept together not by force but by their collaborative interactions. A prominent example is Switzerland, where the German, French, and Italian cantons peacefully coexist, not being forced but due to the advantage of their mutual interactions.

4. Collective Decisions in a Society

The models considered above allow us to describe the general structures and states of social systems whose members exhibit specific behaviors. Thus, in the yes–no model, the members are assumed to make decisions, either "yes" or "no", with regard to problems, say voting in elections or buying something. As one can see, the underlying process of any action is the process of taking decisions. One can say that actually all properties of a social system are caused by the decisions made by its members. This is why it is so important to have a general understanding of how decisions are made. In this Section, some models are considered describing how members of a society make decisions. Collective decisions are based on decisions made by individuals interacting with each other. Therefore, first of all, it is necessary to understand how decisions are made by individuals and then to model how they interact amongst themselves to reach a collective decision. Moreover, there are deep parallels between the processes of making decisions by individuals and by groups since even individual decision-making is a kind of collective decision-making accomplished by the neurons of a brain. Since the notions of decision-making theory are less known to the physically oriented audience, we give an overview below of the main literature and describe the basic points of the theory.

4.1. General Overview

Nowadays, the dominant theory describing the individual behaviors of decision-makers is expected utility theory. This theory was introduced by Bernoulli [52] when investigating the so-called St. Petersburg paradox. Von Neumann and Morgenstern [53] axiomatized this theory. Savage [54] integrated into the theory the notion of subjective probability. The power of the theory was demonstrated by Arrow [55], Pratt [56], and Rothschild and Stiglitz [57,58] in their studies on risk aversion. The flexibility of the theory for characterizing the attitudes of decision-makers toward risk was illustrated by Friedman and Savage [59] and Markowitz [60]. The expected utility theory has provided a mathematical basis for several fields of economics, finance, and management, including the theory of games, the theory of investment, and the theory of search [61–70].

Despite many successful applications of the expected decision theory, quite a number of researchers have discovered a large body of evidence that decision-makers, both human as well as animal, often do not obey the prescriptions of the theory and depart from this theory in a rather systematic way [71]. There then started to appear numerous publications, beginning with Allais [72], Edwards [73,74], and Ellsberg [75], that experimentally confirmed systematic deviations from the prescriptions of expected utility theory, leading to a number of paradoxes.

As examples of such paradoxes, it is possible to mention the Allais paradox [72], independence paradox [72], Ellsberg paradox [75], Kahneman–Tversky paradox [76], Rabin paradox [77], Ariely paradox [71], disjunction effect [78], conjunction fallacy [79,80], isolation effects [81], planning paradox [82], and dynamic inconsistency [83,84].

In order to avoid paradoxes, there have been many attempts to change the expected utility theory, which have been named non-expected utility theories. There are a number of such non-expected utility theories, among which we mention just a few of the best-known approaches. These are the prospect theory [73,76,85], weighted-utility theory [86–88], regret theory [89], optimism–pessimism theory [90], dual-utility theory [91], ordinal-independence theory [92], and quadratic-probability theory [93]. More detailed information on this topic can be found in the recent reviews by Camerer, Loewenstein, and Rabin [94] and by Machina [95].

Despite numerous attempts to resolve the expected utility theory, none of the suggested modifications can explain all of the paradoxes, as has been shown by Safra and Sigal [96]. The best that could be achieved is to fit the interpretation of just one or two of the paradoxes, with the other paradoxes remaining unexplained. Moreover, spoiling the structure of the expected utility theory results in the appearance of several inconsistencies. Accomplishing a detailed analysis, Al-Najjar and Weinstein [97,98] concluded that any variation in the expected utility theory "ends up creating more paradoxes and inconsistences than it resolves".

An original interpretation was advanced by Bohr [99–101], who suggested that psychological processes could be described by resorting to quantum notions, such as interference and complementarity. Von Neumann [102] mentioned that the theory of quantum measurements could be interpreted as decision theory. These ideas were developed in quantum decision theory [103,104], on the basis of which the classical decision-making paradoxes could be explained. It has also been shown [105,106] that quantum decision theory can be reformulated into classical language without employing quantum formulas.

When decision-makers interact with each other, the process of decision-making becomes collective [107–112]. Then, each individual takes decisions based not solely on personal deliberations but also, to some extent, on imitating the actions of other members of the society. Sometimes, this imitation grows to the level of a herding effect. In the present chapter, we describe the main steps of how the process of decision-making develops, starting from individual decisions and progressing to collective decisions taken by a network of society members.

4.2. Utility Function

The primary elements in a problem of choice are some events, outcomes, consequences, or payoffs denoted as x. One considers a set of outcomes, a set of payoffs, a consumer set, or a field of events,

$$X = \{x_i : i = 1, 2, \ldots, N\}. \tag{194}$$

Generally, payoffs, x_i, can be either finite or infinite, and the payoff set as well can be finite or infinite. A mathematically correct definition of infinity is the limit of a sequence.

Payoffs or outcomes have to be measured in a common system of units, accomplished through a utility function $u(x)$, also called an elementary utility function, a pleasure function, a satisfaction function, or a profit function $u(x) : \quad X \rightarrow \mathbb{R}$. The utility function has to satisfy the following properties:

(i) It has to be nondecreasing, so that

$$u(x_1) \geq u(x_2) \qquad (x_1 \geq x_2) \; . \tag{195}$$

If Equation (195) is a strict inequality, the function is termed strictly increasing.

(ii) It is often (although not always) taken to be concave, when

$$u(\alpha_1 x_1 + \alpha_2 x_2) \geq \alpha_1 u(x_1) + \alpha_2 u(x_2), \tag{196}$$

where

$$\alpha_1 \geq 0 \; , \qquad \alpha_2 \geq 0 \; , \qquad \alpha_1 + \alpha_2 = 1 \; .$$

It is called strictly concave, if the inequality (196) represents a strict inequality.

A twice differentiable function is nondecreasing if

$$u'(x) \equiv \frac{du(x)}{dt} \geq 0$$

and it is concave when

$$u''(x) \equiv \frac{d^2u(x)}{dx^2} \leq 0 \; .$$

The derivative $u'(x)$ defines the marginal utility function. According to the above properties, the marginal utility does not increase.

An important property of a utility function is its *risk aversion*. The degree of absolute risk aversion is measured [56] by the quantity,

$$r(x) \equiv -\, \frac{u''(x)}{u'(x)} \; . \tag{197}$$

The coefficient of relative risk aversion [55,56] is defined as

$$\tilde{r}(x) \equiv x r(x) = -x \, \frac{u''(x)}{u'(x)}. \tag{198}$$

For a concave utility function, the degree of risk aversion is non-negative; hence, the utility function is risk averse, $r(x) \geq 0$. This means that with increasing x, the growth of the utility function $u(x)$ does not increase.

Actually, one often employs a *linear utility function*,

$$u(x) = kx \qquad (k > 0) \; . \tag{199}$$

For function (199),

$$u'(x) = k \; , \qquad u''(x) = 0 \; ;$$

hence, the degree of risk aversion is zero, $r(x) = 0$.

Another used form is a *power-law utility function*,

$$u(x) = kx^{\gamma} \qquad (k > 0, \; 0 \leq \gamma < 1) \; , \tag{200}$$

assuming $x \geq 0$. For this case,

$$u'(x) = \frac{k\gamma}{x^{1-\gamma}} \; , \qquad u''(x) = -\frac{k\gamma(1-\gamma)}{x^{2-\gamma}} \; .$$

The degree of risk aversion diminishes with increasing x as

$$r(x) = \frac{1-\gamma}{x} > 0 \; .$$

One more example is the *logarithmic utility function*,

$$u(x) = c \ln(1 + kx) \qquad (c > 0, \; k > 0) \; , \tag{201}$$

again, keeping in mind $x \geq 0$. Then,

$$u'(x) = \frac{ck}{1 + kx} \; , \qquad u''(x) = -\frac{ck^2}{(1 + kx)^2} \; .$$

Hence, the degree of risk aversion diminishes with the increasing payoff x,

$$r(x) = \frac{ck}{1 + kx} \; .$$

Sometimes, one uses an *exponential utility function*,

$$u(x) = c\left(1 - e^{-kx}\right) \qquad (c > 0, \; k > 0) \; , \tag{202}$$

for which

$$u'(x) = cke^{-kx} \; , \qquad u''(x) = -ck^2 e^{-kx} \; ,$$

and the degree of risk aversion is constant, $r(x) = k$.

In the above examples, the utility function is twice differentiable, such that $u'(x)$ and $u''(x)$ exist; it is non-decreasing and concave, hence risk-averse; it is also non-negative for non-negative x; and it is normalized to zero, so that $u(0) = 0$, implying that the utility of nothing is zero.

Some of the utility functions exemplified above are defined only for positive payoffs $x > 0$. In the case of losses, payoffs are negative. Then, one defines different utility functions for gains and losses, for example, $u_{\rm g}(x)$ for gains and $u_{\rm l}(x)$ for losses. Certainly, these should coincide at zero, so that $u_{\rm g}(0) = u_{\rm l}(0)$.

It is also useful to mention that there are two types of utility: cardinal and ordinal. Cardinal utility can be precisely measured, and the magnitude of the measurement is meaningful, similar to how distance is measured in meters, time in hours, or weight in kilograms. For ordinal utility, its precise magnitude is not important, but the magnitude of the ratios between different utilities only is meaningful.

4.3. Expected Utility

The notion of expected utility was introduced by Bernoulli [52], and axiomatic utility theory was formulated by von Neumann and Morgenstern [53]. The definitions below follow the classical exposition of von Neumann and Morgenstern [53].

One defines a probability measure over the set of payoffs,

$$\{ p(x_i) \in [0,1] : \; i = 1,2,\ldots,N \} \; , \tag{203}$$

with the probabilities, $p(x_i)$, normalized to one,

$$\sum_{i=1}^{N} p(x_i) = 1 \; . \tag{204}$$

The probabilities (203) can be objective, prescribed by a rule [53], or they can be subjective probabilities, evaluated by a decision-maker [54].

The probability distribution over a set of payoffs is termed a *lottery*,

$$L = \{ x_i, p(x_i) : \; i = 1,2,\ldots,N \} \; . \tag{205}$$

A compound lottery is a lottery whose outcomes are other lotteries. For two lotteries

$$L_1 = \{ x_i, p_1(x_i) \} \; , \qquad L_2 = \{ x_i, p_2(x_i) \} \; ,$$

the compound lottery is the linear combination,

$$\alpha_1 L_1 + \alpha_2 L_2 = \{ x_i, \; \alpha_1 p_1(x_i) + \alpha_2 p_2(x_i) \} \; , \tag{206}$$

where

$$\alpha_1 \geq 0 \; , \qquad \alpha_2 \geq 0 \; , \qquad \alpha_1 + \alpha_2 = 1 \; .$$

The *lottery mean* is the lottery expected value,

$$E(L) \equiv \sum_{i=1}^{N} x_i p(x_i) \; . \tag{207}$$

The lottery volatility, lottery spread, or lottery dispersion is the variance,

$$\mathrm{var}(L) \equiv \sum_{i=1}^{N} x_i^2 p(x_i) - E^2(L) \; . \tag{208}$$

The lottery volatility, or lottery dispersion, is a measure of lottery uncertainty.

The expected utility of a lottery is the function,

$$U(L) \equiv \sum_{i=1}^{N} u(x_i) p(x_i) \; . \tag{209}$$

The expected utility (209) is proportional to the lottery mean in the case of a linear utility function. Lotteries are ordered, implying the following definitions.

Lotteries L_1 and L_2 are indifferent if, and only if,

$$U(L_1) = U(L_2) \qquad (L_1 = L_2) \; . \tag{210}$$

A lottery L_1 is preferred to L_2 if, and only if,

$$U(L_1) > U(L_2) \qquad (L_1 > L_2) \; . \tag{211}$$

A lottery L_1 is preferred or indifferent to L_2 if, and only if,

$$U(L_1) \geq U(L_2) \qquad (L_1 \geq L_2) \; . \tag{212}$$

The expected utility satisfies the following properties.

(i) *Completeness*. For any two lotteries L_1 and L_2, one of the conditions is valid:

$$L_1 = L_2 \; , \quad L_1 < L_2 \; , \quad L_1 > L_2 \; , \quad L_1 \leq L_2 \; , \quad L_1 \geq L_2 \; . \tag{213}$$

(ii) *Transitivity*. For any three lotteries, such that $L_1 \leq L_2$ and $L_2 \leq L_3$, it follows that $L_1 \leq L_3$.

(iii) *Continuity*. For any three lotteries, ordered so that $L_1 \leq L_2 \leq L_3$, there exists $\alpha \in [0,1]$ for which

$$L_2 = \alpha L_1 + (1-\alpha) L_3 \; . \tag{214}$$

(iv) *Independence*. For any $L_1 \leq L_2$ and arbitrary L_3, there exists $\alpha \in [0,1]$, such that

$$\alpha L_1 + (1-\alpha) L_3 \leq \alpha L_2 + (1-\alpha) L_3 \; . \tag{215}$$

The properties (213)–(215) follow directly from the properties of the utility function described above.

The standard decision-making process proceeds through the following steps. For the same set of payoffs $\{x_n : n = 1, 2, ..., N\}$, there exist several lotteries,

$$L_n = \{x_i, \; p_n(x_i)\} \qquad (n = 1, 2, \ldots) \; . \tag{216}$$

Defining the utility function $u(x)$, one calculates the expected lottery utilities, $U(L_n)$. Comparing the values $U(L_n)$, one selects the largest among them.

A lottery L^* is named *optimal* when it is preferred to all others from the given choice of lotteries. A lottery is optimal if, and only if, its expected utility is the largest:

$$U(L^*) \equiv \sup_n U(L_n) \; . \tag{217}$$

The aim of a decision-maker is assumed to choose an optimal lottery maximizing the related expected utility.

4.4. Time Preference

Betweenwhiles, people need to compare *present goods* that are available for use at the present time with *future goods* that are defined as present expectations of goods becoming available at some date in the future. *Time preference* is the insight that people prefer present goods to future goods [84,113].

A mathematical description of the time preference effect can be performed as follows. Let us consider at time $t = 0$ a lottery,

$$L = \{x_i, \; p(x_i)\} \qquad (t = 0) \; . \tag{218}$$

With a utility function $u(x)$ at time $t = 0$, the expected utility of the lottery is

$$U(L) = \sum_i u(x_i) p(x_i) \qquad (t = 0) \; . \tag{219}$$

The lottery, expected at time $t > 0$, has the form,

$$L(t) = \{x_i(t),\ p(x_i(t))\} \qquad (t > 0)\ . \tag{220}$$

Denoting the utility function at time $t > 0$ as $u(x(t), t)$, we have the expected utility of $L(t)$ as

$$U(L(t)) = \sum_i u(x_i(t), t) p(x_i(t))\ . \tag{221}$$

According to the meaning of time preference, the same goods at a future time are valued lower than at the present time just because any goods can be used during the interval of time $[0, t]$. In the case of money, its value increases with time since it can bring additional profit through an interest rate; hence, the utility of a fixed amount of money decreases with time. This can be formalized as the inequality,

$$u(x_i, t) < u(x_i) \qquad (t > 0)\ . \tag{222}$$

One may introduce a *discount function*, $D(x, t)$, by the relation,

$$u(x_i, t) = u(x_i) D(x_i, t)\ , \tag{223}$$

with the apparent condition,

$$D(x_i, 0) = 1\ . \tag{224}$$

From the time preference condition (222):

$$D(x_i, t) < 1 \qquad (t > 0)\ . \tag{225}$$

In this way, if the payoffs at the present time $t = 0$ and at a future time $t > 0$ are the same, then the present lottery is preferable over the future one,

$$U(L(t)) < U(L) \qquad (x_i(t) = x_i)\ . \tag{226}$$

In particular, if the discount function is uniform with respect to the payoffs,

$$D(x_i, t) = D(t)\ , \tag{227}$$

then

$$U(L(t)) = U(L) D(t)\ , \tag{228}$$

where $D(0) = 1$.

In decision-making, one uses the following discount functions.

Power-law discount function,

$$D(t) = \frac{1}{(1+r)^t}\ . \tag{229}$$

Exponential discount function,

$$D(t) = \exp(-\gamma t)\ . \tag{230}$$

Note that the power-law discount function (229) is equivalent to the exponential form (230), since they are related by the reparametrization,

$$1 + r = e^{\gamma}\ , \qquad \gamma = \ln(1 + r)\ .$$

Hyperbolic discount function,

$$D(t) = \frac{1}{(1 + t/\tau)^{\gamma\tau}} \,, \tag{231}$$

with positive parameters γ and τ.

A detailed review on the effect of time preference is given in Ref. [84].

4.5. Stochastic Utility

There exists a number of factors that influence decision-making. These factors are random or stochastic, and because of this, the related approach to decision-making is named stochastic. These factors, for instance, are:

(i) external conditions under which the decision is made, such as weather, situation in the country, relations with people, opinions of other people, etc.;

(ii) internal physical state of the decision-maker, such as fatigue, illness, pain, influence of alcohol, etc.;

(iii) internal psychological state, including biases, emotions, mood, impatience, etc.;

(iv) previous experience, information, prejudices, religion, principles, character, etc.;

(v) occasional errors in decisions.

Stochastic decision theory assumes that all these interrelated factors can be characterized by some variables, for example, ξ, called *states of nature*. The total collection of the states of nature forms the nature set, $\{\xi\}$. The variables ξ are random and stochastic. It is supposed that there should be a probability measure, $\mu(\xi)$, on the nature set, $\{\xi\}$. The utility function, $u(x,\xi)$, becomes a random variable. The probability of payoffs, $p(x,\xi)$, is also a random variable.

In this way, a lottery becomes a *stochastic lottery*

$$L(\xi) = \{x_i, \; p(x_i,\xi) : \;\; i = 1,2,\ldots,N\} \,. \tag{232}$$

Respectively, one comes to a *stochastic utility,*

$$U(L,\xi) = \sum_{i=1}^{N} u(x_i,\xi) p(x_i,\xi) \,. \tag{233}$$

As far as the states on nature are random, one needs to average over these states; thus, coming to the expected utility,

$$U(L) \equiv \int U(L,\xi) \, d\mu(\xi) \,. \tag{234}$$

The process of stochastic decision-making consists in the choice between several stochastic lotteries,

$$L_n(\xi) = \{x_i, \; p_n(x_i,\xi) : \;\; i = 1,2,\ldots,N\} \,. \tag{235}$$

Technically, this implies the choice between several expected utilities,

$$U(L_n) = \int U(L_n,\xi) \, d\mu(\xi) \,. \tag{236}$$

One needs to choose a lottery with the largest expected utility $U(L^*)$. This approach, however, confronts several quite serious difficulties:

(i) it is not clear how the random variables should be incorporated into lotteries;

(ii) calculations become rather complicated;

(iii) the nature set is not fixed, generally, depending on time;

(iv) the nature set also can depend on the set of payoffs;

(v) the explicit form of the nature states probability is not known and needs to be postulated.

More details on stochastic utility theory can be found, for example, in the books [114,115].

4.6. Affective Decisions

People make decisions based not merely on rational grounds by calculating utility, but also when affected by emotions, which are irrational. Several attempts have been made to take account of emotions in decision-making by modifying expected utility, which is equivalent to some variants of non-expected utility models [91,116,117]. Here, we present the main points of a probabilistic approach to taking emotions into account. First, this approach was formulated by resorting to techniques of quantum theory [103,104]; however, later, it was shown [105,106] that it can be reformulated into classical terms without invoking any quantum expressions. The basics of the *probabilistic affective decision theory* are as follows.

The aim of any decision-making process is to choose from a set of several alternatives. Let the set of alternatives be denoted as

$$\mathbb{A} = \{A_n : \; n = 1, 2, \ldots, N_A\} \; . \tag{237}$$

Each alternative from this set is assumed to be equipped with a probability, $p(A_n)$, with the normalization condition,

$$\sum_n p(A_n) = 1 \; , \qquad 0 \leq p(A_n) \leq 1 \; . \tag{238}$$

This probability shows how it is probable that the alternative A_n can be chosen.

An alternative, A_1, is said to be stochastically preferable to A_2 if, and only if,

$$p(A_1) > p(A_2) \; . \tag{239}$$

Two alternatives are called stochastically indifferent if, and only if,

$$p(A_1) = p(A_2) \; . \tag{240}$$

An alternative, $A_{\rm opt}$, is stochastically optimal if its probability is maximal:

$$p(A_{\rm opt}) = \sup_n p(A_n) \; . \tag{241}$$

The usefulness of an alternative is characterized by a *utility factor*, $f(A_n)$, whose form is to be prescribed by normative rules. This factor shows the probability of choosing an alternative, A_n, based on a rational understanding of its utility. The standard probability normalization is applied:

$$\sum_n f(A_n) = 1 \; , \qquad 0 \leq f(A_n) \leq 1 \; . \tag{242}$$

One says that an alternative A_1 is more useful than A_2 if, and only if,

$$f(A_1) > f(A_2) \; . \tag{243}$$

Two alternatives are equally useful if, and only if,

$$f(A_1) = f(A_2) \; . \tag{244}$$

Affective features of an alternative are represented by an *attraction factor*, $q(A_n)$, with the normalization,

$$\sum_n q(A_n) = 0 \; , \qquad -1 \leq q(A_n) \leq 1 \; . \tag{245}$$

Positive attraction factors imply attractive alternatives, while negative attraction factors mean that the corresponding alternatives are repulsive. An alternative A_1 is more attractive than A_2 if, and only if,

$$q(A_1) > q(A_2) \; . \tag{246}$$

Two alternatives are equally attractive if, and only if,

$$q(A_1) = q(A_2) \; . \tag{247}$$

The probability, $p(A_n)$, of an alternative A_n is a function of the related utility factor, $f(A_n)$, and of attraction factor, $q(A_n)$, such that the *rational boundary condition* is valid:

$$p(A_n) = f(A_n) \; , \qquad q(A_n) = 0 \; , \tag{248}$$

when, in the absence of emotions, the probability of an alternative coincides with the rational utility factor.

The simplest form of the probability function satisfying the rational boundary condition (248) is prescribed by the *superposition axiom*,

$$p(A_n) = f(A_n) + q(A_n) \; . \tag{249}$$

An explicit expression for the utility factor can be derived from the principle of minimal information in Section 2. Assume that for each alternative, A_n, there exists a utility functional, $U(A_n)$, so that an average utility is given by the standard condition,

$$\sum_n f(A_n) U(A_n) = U \; . \tag{250}$$

Then, the information function, under conditions (242) and (250), reads:

$$I[\, f(A_n)\,] = \sum_n f(A_n) \; \ln \, \frac{f(A_n)}{f_0(A_n)} + \alpha \left[1 - \sum_n f(A_n) \right] + \beta \left[U - \sum_n f(A_n) U(A_n) \right] , \tag{251}$$

where $f_0(A_n)$ is a trial distribution. Minimizing the information function with respect to the utility factor $f(A_n)$, yields

$$f(A_n) = \frac{f_0(A_n) e^{\beta U(A_n)}}{\sum_n f_0(A_n) e^{\beta U(A_n)}} \; . \tag{252}$$

The trial distribution, $f_0(A_n)$, can be taken following the Luce rule [118–120],

$$f_0(A_n) = \frac{a_n}{\sum_n a_n} \; , \tag{253}$$

where a_n is the attribute of the alternative A_n, having the form,

$$a_n = U(A_n) \; , \qquad U(A_n) \geq 0 \; , \tag{254}$$

for semi-positive utility functions and

$$a_n = \frac{1}{\mid U(A_n) \mid} \; , \qquad U(A_n) < 0 \; , \tag{255}$$

for negative utility functions [106,121,122].

4.7. Wisdom of Crowds

Wisdom of crowds is the notion that assumes that large groups of people are collectively smarter than individual members of the same groups. This concerns any kind of problemsolving and decision-making. The justification of this idea is based on the understanding that the viewpoint of an individual can inherently be biased, being influenced by the individual's emotions and prejudices, whereas taking the average knowledge of a crowd results in the elimination of the noise of subjective biases and emotions; thus, producing a wiser aggregated result [123–126]. When discussing the wisdom of crowds, one keeps in mind the following crowd features: (i) The crowd should have a diversity of opinions; (ii) Each personal opinion should remain independent to those around them, not being influenced by anyone else; (ii) Each individual from the crowd should make their own decision based solely on their individual knowledge. These conditions exclude the situation where crowd members consult with each other and mimic the actions of their neighbors, which can lead to herding effects. The latter is considered in Section 4.8 below.

Let us enumerate the members of a crowd or a society by the index $j = 1, 2, \ldots N$. According to the conisderation in Section 4.6, each member of the considered group chooses an alternative A_n with the probability,

$$p_j(A_n) = f_j(A_n) + q_j(A_n) \; , \tag{256}$$

under the standard normalization condition,

$$\sum_{n=1}^{N_A} p_j(A_n) = 1 \; , \qquad 0 \leq p_j(A_n) \leq 1 \; .$$

The aggregate opinion implies the arithmetic averaging over the society members, which yields the average probability,

$$p(A_n) \equiv \frac{1}{N} \sum_{j=1}^{N} p_j(A_n) \; , \tag{257}$$

composed of the superposition of the average utility factor,

$$f(A_n) \equiv \frac{1}{N} \sum_{j=1}^{N} f_j(A_n) \; , \tag{258}$$

and the average attraction factor,

$$q(A_n) \equiv \frac{1}{N} \sum_{j=1}^{N} q_j(A_n) \; . \tag{259}$$

Thus, coming to expression (249).

The utility factor is prescribed by a rational evaluation of the utility of the considered alternatives; hence, it is weakly dependent on subjective emotions. This means that $f_j(A_n)$ is approximately the same for any group member, which is equivalent to the condition,

$$f_j(A_n) = f(A_n) \; . \tag{260}$$

Notice that the normalization conditions,

$$\sum_{n=1}^{N_A} f_j(A_n) = 1 \; , \qquad 0 \leq f_j(A_n) \leq 1 \tag{261}$$

remain valid.

The attraction factor, on the contrary, is subjective, being essentially influenced by the agent's emotions. In that sense, the attraction factor is a random quantity. It is random because of several causes. First, its randomness is due to the variability of emotions experienced by rather different people. Second, emotions, even of the same person, vary at different times. Third, emotions randomly influence choice due to the generic variability and local instability of neural networks in the brain, as has been found in numerous psychological and neurophysiological studies [127–135].

Nevertheless, despite the intrinsic randomness of the attraction factor, some of its aggregate properties can be well defined. First of all, the *alternation law* is satisfied:

$$\sum_{n=1}^{N_A} q_j(A_n) = 0 \; . \tag{262}$$

Equation (262) follows straightforwardly from Equations (256) and (261). The expression (256) also tells us that the attraction factor for a j-th society member is in the range

$$-f_j(A_n) \leq q_j(A_n) \leq 1 - f_j(A_n) \; . \tag{263}$$

Respectively, the aggregate attraction factor varies in the interval

$$-f(A_n) \leq q(A_n) \leq 1 - f(A_n) \; . \tag{264}$$

The set of alternatives can be separated into three classes, depending on the signs of the attraction factors, the class of positive attraction factors,

$$q_+(A_n) = q(A_n) > 0 \; , \tag{265}$$

negative attraction factors,

$$q_-(A_n) = q(A_n) < 0 \; , \tag{266}$$

and zero attraction factors when $q(A_n) = 0$. According to the limits (264), positive attraction factors are in the interval

$$0 < q_+(A_n) < 1 - f(A_n) \; , \tag{267}$$

while negative attraction factors are in the range

$$-f(A_n) < q_-(A_n) < 0 \; . \tag{268}$$

Non-informative priors for the attraction factors can be estimated by means of the related arithmetic averages. Recall that, if a quantity y lies in an interval $[a, b]$, its arithmetic average is $\overline{y} = (a + b)/2$. If the interval limits are in the ranges $a_1 \leq a \leq a_2$ and $b_1 \leq b \leq b_2$, then they are expressed through their averages, so that

$$\overline{y} = \frac{1}{2}\left(\overline{a} + \overline{b}\right) = \frac{1}{2}\left(\frac{a_1 + a_2}{2} + \frac{b_1 + b_2}{2}\right) .$$

Keeping in mind that $0 < q_+(A_n) < 1 - f(A_n)$ and $-f(A_n) < q_-(A_n) < 0$, while $0 \leq f(A_n) \leq 1$, one obtains the *quarter law*,

$$q_+(A_n) = \frac{1}{4} \; , \qquad q_-(A_n) = -\,\frac{1}{4} \; . \tag{269}$$

Employing the non-informative priors for the aggregate attraction factors, one can estimate the probability of alternatives, averaged over the crowd, as

$$p(A_n) = f(A_n) \, \pm \, \frac{1}{4} \; , \tag{270}$$

depending on whether the alternatives on average are attractive or not. The quarter law has been found to be in quite a good agreement with empirical data [106,136].

4.8. Herding Effect

In Section 4.7, the process of decision-making by a crowd of independent agents was considered. However, generally, the members of a society interact with each other, which can result in drastic changes in the agents' behavior, such as the occurrence of the herding effect [137–143]. There can exist two kinds of interactions between the society members, which can lead to collective effects such as herding. First, the members can mimic the actions of others, replicating their behavior. Second, the members communicate through information exchange. Strictly speaking, the correct description of collective effects arising in the process of agents' interactions requires the study of temporal processes, just because collective effects need time to form and develop. The consideration of temporal collective effects is beyond the scope of the present article. However, due to their importance for social systems, we delineate the principal points of how collective interactions can be incorporated into affective decision theory. Further details can be found in Refs. [144–146].

Multistep decision theory deals with quantities depending on time. Then, the probability of choosing by an agent j an alternative A_n at time t is $p_j(A_n,t)$, with the normalization,

$$\sum_{n=1}^{N_A} p_j(A_n,t) = 1\,, \qquad 0 \le p_j(A_n,t) \le 1\,. \tag{271}$$

The utility factor becomes $f_j(A_n,t)$, with the normalization,

$$\sum_{n=1}^{N_A} f_j(A_n,t) = 1\,, \qquad 0 \le f_j(A_n,t) \le 1\,, \tag{272}$$

and the attraction factor reads as $q_j(A_n,t)$, satisfying the conditions,

$$\sum_{n=1}^{N_A} q_j(A_n,t) = 0\,, \qquad -1 \le q_j(A_n,t) \le 1\,. \tag{273}$$

Taking into account the tendency of society members to replicate the actions of others defines the probability dynamics,

$$p_j(A_n,t+1) = (1-\varepsilon_j)[\, f_j(A_n,t) + q_j(A_n,t)\,] + \frac{\varepsilon_j}{N-1}\sum_{i(\neq j)}^{N}[\, f_i(A_n,t) + q_i(A_n,t)\,]\,, \tag{274}$$

where ε_j is a replication parameter satisfying the condition,

$$0 \le \varepsilon_j \le 1 \qquad (j = 1,2,\ldots,N)\,. \tag{275}$$

As is explained in Section 4.6 above, there are two types of agent interactions in a society: the replication of the actions of other members and the exchange of information. The latter is known to attenuate the influence of emotions, which results in the attenuation of the attraction factor. The attenuation of emotional influence due to agent interactions is well confirmed by empirical observations [147–158]. The attenuation of the attraction factor is described [144–146,158] by the form,

$$q_j(A_n,t) = q_j(A_n)\exp\{-M_j(t)\}\,, \tag{276}$$

where $q_j(A_n)$ is the attraction factor of an agent j in the absence of social interactions, and $M_j(t)$ is the amount of information received at the moment of time t by an agent j. The quantity $M_j(t)$ that, for brevity, can be called the *memory*, is expressed as

$$M_j(t) = \sum_{t'=0}^{t}\sum_{i=1}^{N} J_{ji}(t,t')\mu_{ji}(t')\,. \tag{277}$$

Here, $J_{ji}(t, t')$ is the information transfer function from agent i to agent j in the interval of time from t' to t, and $\mu_{ji}(t)$ is the Kullback–Leibler information gain (13) received by agent j from agent i at time t,

$$\mu_{ji}(t) = \sum_{n=1}^{N_A} p_j(A_n, t) \ \ln \ \frac{p_j(A_n, t)}{p_i(A_n, t)} \ . \tag{278}$$

The interaction of the society members through replication and information exchange results in a rich variety of behavioral types, including herding, periodic cycles, and chaotic fluctuations, which depend on the society's parameters. A detailed analysis of multistep decision-making by intelligent members of a society is given in Refs. [144–146].

5. Conclusions

In this review, the principle of minimal information has been formulated, which gives the key for constructing probability distributions for equilibrium and quasi-equilibrium social systems. Several simple enough examples, based on the yes–no model, have been considered. Despite their seeming simplicity, the models allow one to describe rather nontrivial effects, including the role of regulation costs and the existence of fluctuating groups inside a society. Quite surprisingly, it turns out that the occurrence of self-organized disorder can make a society more stable. The peculiarity of coexisting populations explains when these populations can mix and live peacefully in the same country and when this coexistence becomes unstable, so that the country separates with different populations in different locations. Since, behind all actions of any society, there are decisions of its members, the basics of the probabilistic affective decision theory are delineated.

The author hopes that the material in the above survey has given the reader a feeling of the wide possibilities of applying mathematical models of physics to describe social systems, even being limited to equilibrium systems. The other part [7] of the review will be devoted to nonequilibrium systems studying various evolution equations and dynamical effects.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: The author very much appreciates the hospitality of and numerous discussions with Didier Sornette, without whose influence this work would not have been done. I am grateful for the discussions and help of Elizaveta P. Yukalova.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Waldrop, M.M. *Complexity: The Emerging Science at the Edge of Order and Chaos*; Touchstone/Simon & Schuster: New York, NY, USA, 1993. Available online: https://uberty.org/wp-content/uploads/2017/04/Waldrop-M.-Mitchell-Complexity-The-Emerging-Science-at-Edge-of-Order-and-Chaos.pdf (accessed on 12 May 2023).
2. Bar-Yam, Y. *Dynamics of Complex Systems*; Addison-Wesley: Reading, MA, USA, 1997. Available online: https://necsi.edu/dynamics-of-complex-systems (accessed on 12 May 2023).
3. Mitchell, M. *Complexity: A Guided Tour*; Oxford University Press: New York, NY, USA, 2009. Available online: https://dokumen.tips/documents/complexity-a-guided-tour-56b13a81a2c4f.html (accessed on 12 May 2023).
4. Thurner, S.; Hanel, R.; Klimekl, P. *Introduction to the Theory of Complex Systems*; Oxford University Press: New York, NY, USA, 2018. [CrossRef]
5. Ladyman, J.; Wiesner, K. *What Is a Complex System*; Yale University Press: New Haven, CT, USA, 2020. [CrossRef]
6. Browning, R. *Browning's Complete Poetical Works (The Complete Poetic and Dramatic Works of Robert Browning. Cambridge Edition)*; Houghton, Mifflin and Company: Boston, MA, USA, 1895.
7. Yukalov, V.I. Selected topics of social physics: Nonequilibrium systems. *Physics* **2023**, *5*, *to be published*.
8. Loewenstein, G.; Rick, S.; Cohen, J.D. Neuroeconomics. *Annu. Rev. Psychol.* **2008**, *59*, 647–672. [CrossRef] [PubMed]
9. Simon, H.A. A behavioral model of rational choice. *Quart. J. Econ.* **1955**, *69*, 99–118. [CrossRef]
10. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Techn. J.* **1948**, *27*, 379–423. [CrossRef]

11. Gibbs, J.W. *The Collected Works*; Longmans, Green and Co.: New York, NY, USA, 1928; Volume 1. Available online: https://gallica.bnf.fr/ark:/12148/bpt6k95192s.image (accessed on 12 May 2023).
12. Gibbs, J.W. *The Collected Works*; Yale University Press: New Haven, CT, USA, 1948; Volume 2. Available online: https://archive.org/details/collectedworksof02gibb/page/n5/mode/2up (accessed on 12 May 2023).
13. Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [CrossRef]
14. Kullback, S. *Information Theory and Statistics*; Peter Smith: Gloucester, MA, USA, 1978. Available online: https://doc.lagout.org/Others/Information%20Theory/Information%20Theory/Information%20theory%20and%20statistics%20-%20Solomon%20Kullback.pdf (accessed on 12 May 2023).
15. Shore, J.E.; Johnson, R.W. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inf. Theory* **1980**, *26*, 26–37. [CrossRef]
16. Janes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620–630. [CrossRef]
17. Jaynes, E.T. *Probability Theory: The Logic of Science*; Cambridge University Press: New York, NY, USA, 2003. Available online: http://www.med.mcgill.ca/epidemiology/hanley/bios601/GaussianModel/JaynesProbabilityTheory.pdf (accessed on 12 May 2023).
18. Yukalov, V.I. Nonequilibrium representative ensembles for isolated quantum systems. *Phys. Lett. A* **2011**, *375*, 2797–2801. [CrossRef]
19. Ter Haar, D. *Elements of Statistical Mechanics*; Reed Elsevier plc group/Butterworth-Heinemann Ltd: Oxford, UK, 1995. . [CrossRef]
20. Ter Haar, D. Theory and applications of the density matrix. *Rep. Prog. Phys.* **1961**, *24*, 304–362. [CrossRef]
21. Yukalov, V.I. Phase transitions and heterophase fluctuations. *Phys. Rep.* **1991**, *208*, 395–492. [CrossRef]
22. Yukalov, V.I. Fluctuations of composite observables and stability of statistical systems. *Phys. Rev. E* **2005**, *72*, 066119. [CrossRef] [PubMed]
23. Yukalov, V.I. Representative ensembles in statistical mechanics. *Int. J. Mod. Phys. B* **2007**, *21*, 69–86. [CrossRef]
24. Yukalov, V.I. Representative statistical ensembles for Bose systems with broken gauge symmetry. *Ann. Phys.* **2008**, *323*, 461–499. [CrossRef]
25. Kubo, R. *Thermodynamics*; North-Holland Publishing Company: Amsterdam, The Netherlands, 1968. Available online: https://www.scribd.com/doc/279300920/Kubo-R-Thermodynamics-NH-1968-T-310s-pdf (accessed on 12 May 2023).
26. Yukalov, V.I. Theory of cold atoms: Basics of quantum statistics. *Laser Phys.* **2013**, *23*, 062001. [CrossRef]
27. Yukalov, V.I. Stochastic instability of quasi-isolated systems. *Phys. Rev. E* **2002**, *65*, 056118. [CrossRef] [PubMed]
28. Yukalov, V.I. Irreversibility of time for quasi-isolated systems. *Phys. Lett. A* **2003**, *308*, 313–318. [CrossRef]
29. Yukalov, V.I. Expansion exponents for nonequilibrium systems. *Phys. A* **2003**, *320*, 149–168. [CrossRef]
30. Yukalov, V.I. Equilibration of quasi-isolated quantum systems. *Phys. Lett. A* **2012**, *376*, 550–554. [CrossRef]
31. Yukalov, V.I. Decoherence and equilibration under nondestructive measurements. *Ann. Phys.* **2012**, *327*, 253–263. [CrossRef]
32. Dembo, A.; Cover, T.M.; Thomas, J.A. Information theoretic inequalities. *IEEE Trans. Inform. Theory* **1991**, *37*, 1501–1517. [CrossRef]
33. Weidlich, W. Physics and social science: The approach of synergetics. *Phys. Rep.* **1991**, *204*, 1–163. [CrossRef]
34. Parsons, T. *The Social System*; Taylor & Francis Group/Routledge: London, UK, 2005. Available online: https://voidnetwork.gr/wp-content/uploads/2016/10/The-Social-System-by-Talcott-Parsons.pdf (accessed on 12 May 2023).
35. Galam, S. *Sociophysics: A Physicist's Modeling of Psycho-Political Phenomena*; Springer Science+Business Media, LLC: New York, NY, USA, 2012. [CrossRef]
36. Perc, M.; Gomez-Gardenes, J.; Szolnoki, A.; Floria, L.M.; Moreno, Y. Evolutionary dynamics of group interactions on structured populations: A review. *J. Roy. Soc. Interface* **2013**, *10*, 20120997. [CrossRef]
37. Perc, M.; Jordan, J.J.; Rand, D.G.; Wang, Z.; Boccaletti, S.; Szolnoki, A. Statistical physics of human cooperation. *Phys. Rep.* **2017**, *687*, 1–51. [CrossRef]
38. Jusup, M.; Holme, P.; Kanazawa, K.; Takayasu, M.; Romic, I.; Wang, Z.; Gecek, S.; Lipic, T.; Podobnik, B.; Wang, L.; et al. Social physics. *Phys. Rep.* **2022**, *948*, 1–148. [CrossRef]
39. Yukalov, V.I.; Yukalova, E.P. Zeroth-order nucleation transition under nanoscale phase separation. *Symmetry* **2021**, *13*, 2379 [CrossRef]
40. Huang, K. *Statistical Mechanics*; John Wiley & Sons, Inc.: New York, NY, USA, 1987. Available online: https://physicsgg.files.wordpress.com/2016/09/huang-kerson-1987-statistical-mechanics-2ed-wileyt506s.pdf (accessed on 12 May 2023).
41. Kubo, R. *Statistical Mechanics*; North-Holland Publishing Company: Amsterdam, The Netherlands, 1965. Available online: https://emineter.files.wordpress.com/2018/08/kubo_zbirka.pdf (accessed on 12 May 2023).
42. Isihara, A. *Statistical Physics*; Academic Press, Inc.: New York, NY, USA, 1971. [CrossRef]
43. Kadanoff, L.P. *Statistical Physics. Statics, Dynamics and Renormalization*; World Scientific Co. Ltd: Singapore, 2000. [CrossRef]
44. Sornette, D. Physics and financial economics (1776–2014): Puzzles, Ising and agent-based models. *Rep. Prog. Phys.* **2014**, *77*, 062001. [CrossRef] [PubMed]
45. Brush, S.G. History of the Lenz-Ising model. *Rev. Mod. Phys.* **1967**, *39*, 883–893. [CrossRef]
46. Hayek, F. *The Road to Serfdom*; Taylor & Francis Group/Routledge: New York, NY, USA, 2006; Reprint of 1944 edition. Available online: https://ctheory.sitehost.iu.edu/img/Hayek_The_Road_to_Serfdom.pdf (accessed on 12 May 2023).
47. Yukalov, V.I. Mesoscopic phase fluctuations: General phenomenon in condensed matter. *Int. J. Mod. Phys. B* **2003**, *17*, 2333–2358. [CrossRef]

48. Yukalov, V.I. Systems with symmetry breaking and restoration. *Symmetry* **2010**, *2*, 40–68. [CrossRef]
49. Yukalov, V.I. Renormalization of quasi-Hamiltonians under heterophase averaging. *Phys. Lett. A* **1987**, *125*, 95–100. [CrossRef]
50. Yukalov, V.I. Procedure of quasi-averaging for heterophase mixtures. *Physica A* **1987**, *141*, 352–374. [CrossRef]
51. Yukalov, V.I. Lattice mixtures of fluctuating phases. *Physica A* **1987**, *144*, 369–389. [CrossRef]
52. Bernoulli, D. Specimen Theoriae Novae de Mensura Sortis. *Comment. Academ. Scient. Imper. Petropolit. [Proc. Imper. Acad. Sci. St. Petersburg]* **1738**, *V*, 175–192. English translation: Exposition of a new theory on the measurement of risk. In *The Kelly Capital Growth Investment Criterion. Theory and Practice*; MacLean, L.C., Thorp, E.O., Ziemba, W.T., Eds.; World Scientific Co. Ltd.: Singapore, 2011; pp. 11–24. [CrossRef]
53. von Neumann, J.; Morgenstern, O. *Theory of Games and Economic Behavior*; Princeton University Press: Princeton, NJ, USA, 1953. [CrossRef]
54. Savage, L.J. *The Foundations of Statistics*; Dover Publications, Inc.: New York, NY, USA, 1972. Available online: https://gwern.net/doc/statistics/decision/1972-savage-foundationsofstatistics.pdf (accessed on 12 May 2023).
55. Arrow, K.J. *Essays in the Theory of Risk Bearing*; North Holland Publishing Company: Amsterdam, The Netherlands, 1974. Available online: https://archive.org/details/essaysintheoryof0000arro (accessed on 12 May 2023).
56. Pratt, J.W. Risk aversion in the small and in the large. *Econometrica* **1964**, *32*, 122–136. [CrossRef]
57. Rothschild, M.; Stiglitz, J. Increasing risk: A definition. *J. Econ. Theory* **1970**, *2*, 225–243. [CrossRef]
58. Rothschild, M.; Stiglitz, J. Increasing risk: Its economic consequences. *J. Econ. Theory* **1971**, *3*, 66–84. [CrossRef]
59. Friedman, M.; Savage, L. The utility analysis of choices involving risk. *J. Polit. Econ.* **1948**, *56*, 279–304. [CrossRef]
60. Markovitz, H. The utility of wealth. *J. Polit. Econ.* **1952**, *60*, 151–158. [CrossRef]
61. Lindgren, B.W. *Elements of Decision Theory*; Macmillan: New York, NY, USA, 1971.
62. White, D.I. *Decision Theory*; Taylor & Francis Group/Routledge: New York, NY, USA, 2006. [CrossRef]
63. Rivett, P. *Model Building for Decision Analysis*; John Wiley & Sons: Chichester, UK, 1980.
64. Berger, J.O. *Statistical Decision Theory and Bayesian Analysis*; Springer Science+Business Media New York: New York, NY, USA, 1985. [CrossRef]
65. Marshall, K.T.; Oliver, R.M. *Decision Making and Forecasting*; McGraw-Hill, Inc.: New York, NY, USA, 1995.
66. Bather, J. *Decision Theory: An Introduction to Dynamic Programming and Sequential Decisions*; John Wiley & Sons, Ltd: Chichester, UK, 2000. Available online: https://archive.org/details/decisiontheoryin0000bath (accessed on 12 May 2023).
67. French, S.; Rios Insua, D. *Statistical Decision Theory. Kendall's Library of Statistics 9*; John Wiley & Sons Ltd/Hodder Arnold: London, UK, 2000.
68. Raiffa, H.; Schlaifer, R. *Applied Statistical Decision Theory*; John Wiley & Sons: New York, NY, USA, 2000.
69. Weirich, P. *Decision Space*; Cambridge University Press: Cambridge, UK, 2001. [CrossRef]
70. Gollier, C. *Economics of Risk and Time*; The MIT Press: Cambridge, MA, USA, 2001. [CrossRef]
71. Ariely, D. *Predictably Irrational*; Harper Collins Publishers: New York, NY, USA, 2008. Available online: https://radio.shabanali.com/predictable.pdf (accessed on 12 May 2023).
72. Allais, M. Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'ecole Americaine. *Econometrica* **1953**, *21*, 503–546. [CrossRef]
73. Edwards, W. The prediction of decision among bets. *J. Experim. Psychol.* **1955**, *50*, 200–204. [CrossRef]
74. Edwards, W. Subjective probabilities inferred from decisions. *Psychol. Rev.* **1962**, *69*, 109–135. [CrossRef]
75. Ellsberg, D. Risk, ambiguity, and the Savage axioms. *Quart. J. Econ.* **1961**, *75*, 643–669. [CrossRef]
76. Kahneman, D.; Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica* **1979**, *47*, 263–291. [CrossRef]
77. Rabin, M. Risk aversion and expected-utility theory: A calibration theorem. *Econometrica* **2000**, *68*, 1281–1292. [CrossRef]
78. Tversky, A.; Shafir, E. The disjunction effect in choice under uncertainty. *Psychol. Sci.* **1992**, *3*, 305–309. [CrossRef]
79. Tversky, A.; Kahneman, D. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgement. *Psychol. Rev.* **1983**, *90*, 293–315. [CrossRef]
80. Shafir, E.B.; Smith, E.E.; Osherson, D.N. Typicality and reasoning fallacies. *Mem. Cognit.* **1990**, *18*, 229–239. [CrossRef]
81. McCaffery, E.J.; Baron, J. Isolation effects and the neglect of indirect effects of fiscal policies. *J. Behav. Decis. Mak.* **2006**, *19*, 289–302. [CrossRef]
82. Kydland, F.E.; Prescott, E.C. Rules rather than discretion: The inconsistency of optimal plans. *J. Polit. Econ.* **1977**, *85*, 473–492. [CrossRef]
83. Strotz, R.H. Myopia and inconsistency in dynamic utility maximization. *Rev. Econ. Stud.* **1955**, *23*, 165–180. [CrossRef]
84. Frederick, S.; Loewenstein, G.; O'Donoghue, T. Time discounting and time preference: A critical review. *J. Econ. Literat.* **2002**, *40*, 351–401. [CrossRef]
85. Tversky, A.; Kahneman, D. Advances in prospect theory: Cumulative representations of uncertainty. *J. Risk Uncert.* **1992**, *5*, 297–323. [CrossRef]
86. Karmarkar, U. Subjectively weighted utility: A descriptive extension of the expected utility model. *Org. Behav. Hum. Perform.* **1978**, *21*, 61–72. [CrossRef]
87. Karmarkar, U. Subjectively weighted utility and the Allais paradox. *Org. Behav. Hum. Perform.* **1979**, *24*, 67–72. [CrossRef]
88. Chew, S.H. A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox. *Econometrica* **1983**, *51*, 1065–1092. [CrossRef]

89. Loomes, G.; Sugden, R. Regret theory: An alternative theory of rational choice under uncertainty. *Econ. J.* **1982**, *92*, 805–824. [CrossRef]
90. Hey, J. The economics of optimism and pessimism: A definition and some applications. *Kyklos* **1984**, *37*, 181–205. [CrossRef]
91. Yaari, M. The dual theory of choice under risk. *Econometrica* **1987**, *55*, 95–115. [CrossRef]
92. Green, J.; Jullien, B. Ordinal independence in nonlinear utility theory. *J. Risk Uncert.* **1988**, *1*, 355–387. [CrossRef]
93. Chew, S.; Epstein, L.; Segal, U. Mixture symmetry and quadratic utility. *Econometrica* **1991**, *59*, 139–163. [CrossRef]
94. Camerer, C.F.; Loewenstein, G.; Rabin, R. *Advances in Behavioral Economics*; Princeton University Press: Princeton, NJ, USA, 2004. . [CrossRef]
95. Machina, M.J. Non-expected utility theory. In *The New Palgrave Dictionary of Economics. Second Edition*; Durlauf, S.N., Blume, L.E., Eds.; Macmillan Publishers Ltd/Palgrave Macmillan: Hampshire, UK, 2008; pp. 4586–4596. [CrossRef]
96. Safra, Z.; Segal, U. Calibration results for non-expected utility theories. *Econometrica* **2008**, *76*, 1143–1166. [CrossRef]
97. Al-Najjar, N.I.; Weinstein, J. The ambiguity aversion literature: A critical assessment. *Econ. Philos.* **2009**, *25*, 249–284. [CrossRef]
98. Al-Najjar, N.I.; Weinstein, J. Rejoinder: The ambiguity aversion literature: A critical assessment. *Econ. Philos.* **2009**, *25*, 357–369. [CrossRef]
99. Bohr, N. Light and life. *Nature* **1933**, *131*, 421–423. [CrossRef]
100. Bohr, N. Light and life. *Nature* **1933**, *131*, 457–459. [CrossRef]
101. Bohr, N. *Atomic Physics and Human Knowledge*; John Wiley & Sons, Inc.: New York, NY, USA, 1958. Available online: https://www.holybooks.com/wp-content/uploads/Atomic-Physics-and-Human-Knowledge.pdf (accessed on 12 May 2023).
102. von Neumann, J. *Mathematical Foundations of Quantum Mechanics, New Edition*; Wheeler, N.A., Ed.; Princeton University Press: Princeton, NJ, USA, 2018. [CrossRef]
103. Yukalov, V.I.; Sornette, D. Quantum decision theory as quantum theory of measurement. *Phys. Lett. A* **2008**, *372*, 6867–6871. [CrossRef]
104. Yukalov, V.I.; Sornette, D. Mathematical structure of quantum decision theory. *Adv. Compl. Syst.* **2010**, *13*, 659–698. [CrossRef]
105. Yukalov, V.I. A resolution of St. Petersburg paradox. *J. Math. Econ.* **2021**, *97*, 102537. [CrossRef]
106. Yukalov, V.I. Quantification of emotions in decision making. *Soft Comput.* **2022**, *26*, 2419–2436. [CrossRef]
107. Krause, J.; Ruxton, G.D. *Living in Groups*; Oxford University Press: Oxford, UK, 2002.
108. Seeley, T.D. *Honeybee Democracy*; Princeton University Press: Princeton, NJ, USA, 2010. Available online: https://hadinur1969.files.wordpress.com/2017/10/thomas_d_seeley-honeybee_democracy_-princeton_univ.pdf (accessed on 12 May 2023).
109. Easley, D.; Kleinberg, J. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*; Cambridge University Press: New York, NY, USA, 2010. [CrossRef]
110. Bourke, A.F.G. *Principles of Social Evolution*; Oxford University Press Inc.: New York, NY, USA, 2011. . acprof:oso/9780199231157.001.0001 [CrossRef]
111. Marshall, J.A.R. *Social Evolution and Inclusive Fitness Theory: An Introduction*; Princeton University Press: Princeton, NJ, USA, 2015.
112. Hamann, H. *Swarm Robotics: A Formal Approach*; Springer International Publishing AG: Cham, Switzerland, 2018. [CrossRef]
113. Samuelson, P. A note on measurement of utility. *Rev. Econ. Stud.* **1937**, *4*, 155–161. [CrossRef]
114. Pratt, J.W.; Raiffa, H.; Schlaifer, R. *Introduction to Statistical Decision Theory*; The MIT Press: Cambridge, MA, USA, 2008.
115. Liese, F.; Miescke, K.-J. *Statistical Decision Theory: Estimation, Testing, and Selection*; Springer: New York, NY, USA, 2008. [CrossRef]
116. Reynaa, V.F.; Brainer, C.J. Dual processes in decision making and developmental neuroscience: A fuzzy-trace model. *Developm. Rev.* **2011**, *31*, 180–206. [CrossRef] [PubMed]
117. Bracha, A.; Brown, D.J. Affective Decision Making: A Theory of Optimism Bias. *Games Econ. Behav.* **2012**, *1*, 67–80. [CrossRef]
118. Luce, R.D. *Individual Choice Behavior: A Theoretical Analysis*; Wiley: New York, NY, USA, 1959. Available online: https://www.scribd.com/book/271527880/Individual-Choice-Behavior-A-Theoretical-Analysis (accessed on 12 May 2023).
119. Luce, R.D.; Raiffa, R. *Games and Decisions: Introduction and Critical Survey*; Dover Publications, Inc.: New York, NY, USA, 1989. Available online: https://www.scribd.com/book/271636770/Games-and-Decisions-Introduction-and-Critical-Survey (accessed on 12 May 2023).
120. Gul, F.; Natenzon, P.; Pesendorfer, W. Random choice as behavioral optimization. *Econometrica* **2014**, *82*, 1873–1912. [CrossRef]
121. Yukalov, V.I.; Sornette, D. Manupulating decision making of typical agents. *IEEE Trans. Syst. Man Cybern. Syst.* **2014**, *44*, 1155–1168. [CrossRef]
122. Yukalov, V.I.; Sornette, D. Quantitative predictions in quantum decision theory. *IEEE Trans. Syst. Man Cybern. Syst.* **2018**, *48*, 366–381. [CrossRef]
123. Surowiecki, J. *The Wisdom of Crowds*; Random House, Inc./Anchor Books: New York, NY, USA, 2005.
124. Sunstein, C. *Infotopia: How Many Minds Produce Knowledge*; Oxford University Press: Oxford, UK, 2006.
125. Page, S.E. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*; Princeton University Press: Princeton, NJ, USA, 2007. Available online: https://www.scribd.com/book/232949830/The-Difference-How-the-Power-of-Diversity-Creates-Better-Groups-Firms-Schools-and-Societies-New-Edition (accessed on 12 May 2023).
126. Fiechter, J.; Kornell, N. How the wisdom of crowds, and of the crowd within, are affected by expertise. *Cogn. Res.* **2021**, *6*, 5. [CrossRef]
127. Werner, G.; Mountcastle, V.B. The variability of central neural activity in a sensory system and its implications for the central reflection of sensory events. *J. Neurophysiol.* **1963**, *26*, 958–977. [CrossRef]

128. Arieli, A.; Sterkin, A.; Grinvald, A.; Aertsen, A. Dynamics of ongoing activity: Explanation of the large variability in evoked cortical responses. *Science* **1986**, *273*, 1868–1871. [CrossRef]
129. Gold, J.I.; Shadlen, M.N. Neural computations that underlie decisions about sensory stimuli. *Trends Cognit. Sci.* **2001**, *5*, 10–16. [CrossRef]
130. Glimcher, P.W. Indeterminacy in brain and behaviour. *Annu. Rev. Psychol.* **2005**, *56*, 25–56. [CrossRef]
131. Schumacher, J.F.; Thompson, S.K.; Olman, C.A. Contrast response functions for single Gabor patches: ROI-based analysis over-represents low-contrast patches for GE BOLD. *Front. Syst. Neurosci.* **2011**, *5*, 19. [CrossRef]
132. Shadlen, M.N.; Shohamy, D. Perspective decision making and sequential sampling from memory. *Neuron* **2016**, *90*, 927–939. [CrossRef]
133. Webb, R. The neural dynamics of stochastic choice. *Manag. Sci.* **2019**, *65*, 230–255. [CrossRef]
134. Kurtz-David, V.; Persitz, D.; Webb, R.; Levy, D.J. The neural computation of inconsistent choice behaviour. *Nature Commun.* **2019**, *10*, 1583. [CrossRef] [PubMed]
135. Woodford, M. Modeling imprecision in perception, valuation, and choice. *Annu. Rev. Econ.* **2020**, *12*, 579–601. [CrossRef]
136. Yukalov, V.I.; Sornette, D. Preference reversal in quantum decision theory. *Front. Psychol.* **2015**, *6*, 1538. [CrossRef]
137. Martin, E.D. *The Behavior of Crowds: A Psychological Study*; Harper & Brothers: New York, NY, USA, 1920. Available online: https://www.gutenberg.org/files/40914/40914-h/40914-h.htm (accessed on 12 May 2023).
138. Sherif, M. *The Psychology of Social Norms*; Harper & Brothers Publishers: New York, NY, USA, 1936. Available online: https://archive.org/details/in.ernet.dli.2015.264611 (accessed on 12 May 2023).
139. Smelser, N.J. *Theory of Collective Behavior*; Macmillan Company/The Free Press: New York, NY, USA, 1965. Available online: https://archive.org/details/theoryofcollecti00smel/mode/2up (accessed on 12 May 2023).
140. Merton, R.K. *Social Theory and Social Structure*; Macmillan Publishing Co., Inc./The Free Press: New York, NY, USA, 1968. Available online: https://edisciplinas.usp.br/pluginfile.php/4250035/mod_folder/content/0/Textos/Merton%2C%20Social%20Theory%20and%20Social%20Structure.pdf (accessed on 12 May 2023).
141. Turner, R.H.; Killian, L.M. *Collective Behavior*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1993.
142. Hatfield, E.; Cacioppo, J.T.; Rapson, R.L. *Emotional Contagion*; Cambridge University Press: New York, NY, USA, 1993. [CrossRef]
143. Brunnermeier, M.K. *Asset Pricing under Asymmetric Information: Bubbles, Crashes, Technical Analysis, and Herding*; Oxford University Press: New York, NY, USA, 2001. [CrossRef]
144. Yukalov, V.I.; Yukalova, E.P.; Sornette, D. Information processing by networks of quantum decision makers. *Physica A* **2018**, *492*, 747–766. [CrossRef]
145. Yukalov, V.I.; Yukalova, E.P.; Sornette, D. Role of collective information in networks of quantum operating agents. *Physica A* **2022**, *598*, 127365. [CrossRef]
146. Yukalov, V.I.; Yukalova, E.P. Self-excited waves in complex social systems. *Physica D* **2022**, *433*, 133188. [CrossRef]
147. Kühberger, A.; Komunska, D.; Perner, J. The disjunction effect: Does it exist for two-step gambles? *Org. Behav. Hum. Decis. Process.* **2001**, *85*, 250–264. [CrossRef] [PubMed]
148. Charness, G.; Rabin, M. Understanding social preferences with simple tests. *Quart. J. Econ.* **2002**, *117*, 817–869. [CrossRef]
149. Blinder, A.; Morgan, J. Are two heads better than one? An experimental analysis of group versus individual decision-making. *J. Money Credit Bank.* **2005**, *37*, 789–811.
150. Cooper, D.; Kagel, J. Are two heads better than one? Team versus individual play in signaling games. *Am. Econ. Rev.* **2005**, *95*, 477–509. [CrossRef]
151. Sutter, M. Are four heads better than two? An experimental beauty-contest game with teams of different size. *Econ. Lett.* **2005**, *88*, 41–46. [CrossRef]
152. Tsiporkova, E.; Boeva, V. Multi-step ranking of alternatives in a multi-criteria and multi-expert decision making environment. *Inform. Sci.* **2006**, *176*, 2673–2697. [CrossRef]
153. Charness, G.; Karni, E.; Levin, D. Individual and group decision making under risk: An experimental study of bayesian updating and violations of first-order stochastic dominance. *J. Risk Uncert.* **2007**, *35*, 129–148. [CrossRef]
154. Charness, G.; Rigotti, L.; Rustichini, A. Individual behavior and group membership. *Am. Econ. Rev.* **2007**, *97*, 1340–1352. [CrossRef]
155. Chen, Y.; Li, S. Group identity and social preferences. *Am. Econ. Rev.* **2009**, *99*, 431–457. [CrossRef]
156. Liu, H.H.; Colman, A.M. Ambiguity aversion in the long run: Repeated decisions under risk and uncertainty. *J. Econ. Psychol.* **2009**, *30*, 277–284. [CrossRef]
157. Charness, G.; Karni, E.; Levin, D. On the conjunction fallacy in probability judgement: New experimental evidence regarding Linda. *Games Econ. Behav.* **2010**, *68*, 551–556. [CrossRef]
158. Yukalov, V.I.; Sornette, D. Role of information in decision making of social agents. *Int. J. Inf. Technol. Decis. Mak.* **2015**, *14*, 1129–1166. [CrossRef]