CHOICE POLARIZATION ON A SOCIAL INFLUENCE NETWORK

Moses A. Boudourides

Department of Mathematics

University of Patras

265 00 Rio-Patras, Greece

* mboudour@upatras.gr

For presentation at the "Polarization and Conflict"
2nd Annual Summer Meeting of the PAC research group
Konstanz, June 2-5, 2005

An earlier version of this paper was presented at the XXIV International Sunbelt Social Network Conference, Portorož, Slovenia, May 12-16, 2004

Abstract: We are studying here a model of social influence in which actors are dwelling upon a social network and are influencing/influenced by each other through the recurrent and reproduced pattern of their relationships. Furthermore, we assume that influence outcomes represent preference orderings (among a number of alternatives) and, thus, they are located on a certain relational structure (a graph of orderings). Because of the structural arrangement of influence outcomes, actors' orderings may occupy antipodal positions. In this setting, after following a social influence process of random dyadic interactions among actors, the system is stabilized (absorbed) by two possible configurations of actors' influence outcomes: either a unanimous arrangement of orderings or a polarization across dipoles of antipodal orderings. We conduct computer simulations in order to study how this process of polarized social influence depends on certain parameters of the system (such as the size and connectivity of the social network) and the degree of its randomization. We find that polarization appears to be unexpectedly high over social networks of the 'small world' type. Furthermore, we discuss the relevance of this model of polarized social influence on relationally structured outcomes with Arrow's Impossibility Theorem and we outline some possible further extensions of this model.

Introduction

Among various phenomena of collective behavior, certain of them can be formally studied through a common theoretical perspective. These include phenomena of decision-making, opinion formation, social impact, voting in elections, consumers' purchasing decisions etc. (Coleman, 1990, pp. 237-239). What all these phenomena have in common is an underlying influence system affecting the outcomes and dynamics of the corresponding processes through regularized patterns of communication and exchange of information in an explicit network conception of structure (Laumann, 1973; Laumann & Pappi, 1976). According to the structurational 'duality' theory, the influence network – composed of the reproduced social relationships among the actors involved in such collective phenomena – both enables and constraints the corresponding processes (Giddens, 1984; Sewell, 1992). This means that, on the one hand, actors are influenced by each other throughout decision-making, opinion formation, attitudes shaping, voting, development of consumer choices etc. and, on the other hand, the recurrent social structure of these collective phenomena poses concrete opportunities and restrictions for the corresponding processes of influence.

Various types of formalized analyses and mathematical techniques have been applied in the study of such social influence models of collective behavior. Four characteristic formal methodologies typically employed in these studies are the following:[1]

- Social network analysis (French, 1956; Harary, 1959; Abelson, 1964; Friedkin & Johnsen, 1990, 1999; Friedkin, 1999, 2001; Stokman & Van den Bos, 1992; Kobayashi, 2001).

[1] This is not a complete list of references. We should add two remarks. First, in many of the above quoted papers, the level of technical (mathematical) formalization appears disproportional to the depth of substantial interpretations of the analyzed concepts and theories (this is especially true for the references from econophysics and sociophysics based on arguments from statistical mechanics). Furthermore, often more than one of the above four methodologies are combined in a study of an influence model of collective behavior.

- Game theory (Macy, 1990; Young, 1998; Durlauf & Young, 2001; Bramoullé, 2002; Young, 2005).
- Statistical mechanics and computer simulations (Latané, 1981; Nowak, Szamrej & Latané, 1990; Lewenstein, Nowak & Latané, 1992).
- Interacting particle systems[2] (Spitzer, 1971; Liggett, 1985, 1997, 1999; Durrett, 1988, 1999; Griffeath, 1988; Nakamaru & Levin, 2004).

The analyses in these studies are based on three categories of assumptions concerning the ways processes of social influence are formalized in these models. These are assumptions about (i) the 'social space' of actors, over which processes of social influence are embedded, (ii) the 'space of outcomes' that actors are co-developing in these phenomena of collective behavior and (iii) the set of 'rules of interaction' prescribing how social influence is conceptualized and implemented among actors in a social influence model. In what follows in this introduction, we are going to describe the common theoretical options about these assumptions and to specify the ones that we are following in the model of social influence, which we are discussing in the subsequent sections of the present paper.

As we have already explained, the social space of influence processes is usually conceived in relational structural terms. This means that these processes are sustained by a social network (a 'social influence network') connecting the set of the involved actors through certain (more or less stably) reproduced relationships, which disclose the relative influence or power or control that an actor is able to exert over another one. As it is the case in social network analyses, a directed or undirected valued graph is a typical formal representation of a configuration of a social influence system. The vertices (or nodes) of this graph represent actors, who are trying to influence each other according to the strength of their ties with others. The strength of such influence ties is represented by certain values that the links (edges or arcs) of the social influence network can take.

[2] Interacting particle systems might be considered as a specific class of statistical mechanical systems. However, we have treated them separately in the above list because they constitute a discrete methodology of stochastic systems. Furthermore, the references of the statistical mechanical systems that we have given refer almost exclusively to social impact systems analyzed by mean field theories.

These values are real numbers (with the understanding that the value 0 corresponds to a non-existing link), which are usually taken to be positive integers[3] (with dichotomous relationships being the ones taking just one of the two values 0 or 1). Symmetric influence relationships among actors give rise to a directed graph, while non-symmetric (non-reciprocated) relationships correspond to an undirected graph. We should add that in certain models of social influence developed in the tradition of statistical mechanics (in econophysics and sociophysics) and interacting particle systems (but often also in the context of game theory), more regular types of graphs, like lattices, are considered to constitute the social space of the sites that actors dwell upon and through the corresponding regular grid of links they tend to influence and be influenced by each other.[4]

The outcomes of social influence processes might be interpreted in a variety of ways depending on the specific context of the studied processes. For instance, the outcomes of decision-making might be conceived as "group norms" or "decision outcomes" or "policy decisions," the outcomes of opinion formation are just the "opinions," the outcomes of social impact are the "attitudes" or "attitudinal positions" and commonly the outcomes of election voting or consumers' purchasing decisions might be considered as "preference choices." In any case, these outcomes need to be formally represented in order to be analyzed in the context of a social influence model and this is done in one of two ways: Either the outcomes of a social influence process are represented as discrete entities or they are represented on a continuous scale (typically, an interval of real numbers or, more generally, a domain in a vector space).

Here we have to stress that social influence outcomes are considered to be disaggregated and distributed over the actors of the social influence network, who may change their individual outcomes according to the dynamics of the social influence process they are sustaining. This means that each actor in a social influence network might be affiliated

[3] Negative values over relationships are given in the context of the 'balance theory' – see Wasserman & Faust (1994, pp. 220-242).

[4] At the opposite end, we have the social space of actors dwelling upon a random graph or a scale-free network or a 'small world' (see Watts & Strogatz, 1998; Watts, 1999).

with a certain value of an outcome, which might change in time and possibly be transformed to another value in the course of the complex dynamics of the underlying social influence process. In other words, there is a dynamic pattern of values (the social influence outcomes) distributed over the vertices of the graph representing the underlying social space, which is dynamically updated according to the strength of the influences that are exerted over the existing pattern of linkages among actors. Thus, depending on whether an actor influences or is influenced by another one and on the strength of such influences, the outcome of an actor might be pulled from the outcome of another actor or might push the outcome of another actor towards his/her outcome. The total shift in an actor's outcome results after considering all the social influence receptivities and dispositions that can be channeled over the links of an actor with his/her neighbors. Given that a graph representing the social influence space is typically assumed to be connected, we can understand that it is a complex adaptive process what shapes the social influence outcomes that actors possess during the different instances of the time evolution of such collective phenomena.

In the particular study of social influence that we are doing in this paper, we consider the case that the actors' influence outcomes are discrete. Moreover, we are assuming that there is a certain structure in the set of all possible social influence outcomes. In fact, we consider the case that the influence outcomes are embedded in a certain graph. This assumption is not a technical one but it arises in a natural way in the specific context of the social influence processes that we are studying in this paper. This is the setting of 'preference choices' that actors in a social influence network might possess, when they are ordering or ranking a number of discrete alternatives. As we are going to explain in the next section, this is exactly the theoretical formalization used in theories of discrete social choice (Arrow, 1951; Sen, 1970; Mas-Colell *et al*., 1995; Schofield, 2003). In this way, interpreting a social influence outcome (at the individual level) as an actor's preference ordering, we find a proper relational structure among all the possible orderings. This structure is derived from the fundamental assumptions of social choice theory and from a set-theoretic notion of proximity among preferences, which quantifies dissimilarities in orderings through the 'Kemeny distance' among them. In fact, using the old graph-theoretic construction of Kemeny (1959), the set of orderings is represented by a certain graph (that we call 'graph of orderings'), which provides the relational structure encompassing all the outcomes of processes of ordering (or ranking) – such as relational voting or consumers' preferences.

However, by subsuming our analysis inside the perspectives of a social influence model, we are disengaged away from one of the fundamental assumptions of social choice theory, according to which the aggregation of individual choices should be devoid of any influences or manipulations that actors might exert to each other. In return to the abandonment of the main focus of social choice, we are gaining the insight and the methodological tool-kits that theories of social influence can provide. For example, if we follow a dynamic model of social influence, in which actors' outcomes are eventually stabilized[5] at certain absorbing states, then the resulting equilibrium positions might be interpreted as a dynamic aggregation of the influence outcomes, which are driven by the omnipresent processes of social influence. We should note that following such a 'social influence' model of social choice, it is possible that we end up to an equilibrium position of the influence outcomes, which is different from the initial configuration of outcomes that actors in the social influence network were holding in the beginning of the process (although actors were eventually pulled towards this equilibrium by the collective dynamics of the influence process). This is an emergent shift of outcomes towards an equilibrium position, which is produced by the mechanisms of social influence, it is adopted by every actor (unanimously) but it is not 'dictatorial' in the Arrowian sense (since no actor possessed it initially).

What is more interesting is that nonhomogeneous equilibrium positions can emerge too. These are equilibrium configurations composed of diverse influence outcomes, which are distributed over the vertices (actors) of the underlying social influence network. Of course, the existence of such discordant (and 'pluralist') absorbing states depends on the

[5] Of course, instability results implying an eventual fluctuation among different states of outcomes are equally interesting. However, not many such results exist in the literature up to now. Bramson & Griffeath (1989) have studied a fluctuating cyclic interacting particle system (with discrete outcomes) and Flache & Torenvlied (2004) have focused on the unstable dynamics of decision-making (with continuous outcomes).

rules of interactions that actors follow when they update their influence outcomes in a social influence process. Roughly speaking, what determines the absorbing states of an interacting system (as the social influence system that we are going to discuss in the sequel of this paper) depends upon under which conditions actors halt their interactions. In principle, we could think of two plausible ways of interaction halt. One is when all actors have reached the same outcome, after having shifted their initial outcomes, because of the influences they have received and they have exerted in their social influence network. Obviously, this is the case of a unanimous equilibrium outcome that all actors may adopt and subsequently they would be locked-in.

However, there is a second scenario of interaction halt. This is when two interacting actors reach two bridgeless opposite outcomes which constitute the two uncompromising antipodes of a polarized situation. From this point of view, we can understand where an appropriate relational structure of the set of influence outcomes might lead us to. If this structure affords the existence of polarized outcomes (for instance, antipodal positions), then when actors move inside this structure by alternating their positions among all possible outcomes, as they are subject to an influence process, then chances are that a pair of actors seizes two antipodal outcomes, in which case these two actors cannot further interact with each other. When this happens for all pairs of interacting actors, then the social influence network is polarized across two antipodal positions of outcomes, which means that the whole system is trapped (or locked-in) in a nonhomogeneous equilibrium state.

In fact, the last polarization scenario can be realized in the specific social influence model that we are developing here. Since the relational structure triggering polarization that we would like to consider here comes from the theoretical setting of social choice theory, in the following first section we intend to present the fundamental concepts of this theory. As a matter of fact, since our focus is on social networks and those working in the sociological field of social network analysis have rather limited exposure to theories of social choice, our presentation will be sufficiently detailed (at the cost of repeating very elementary facts from theoretical economics and formal political theory). In the second section, we are going to give an explicit construction of the graph of orderings (representing the relational structure of the social influence outcomes in our model), which is based in Kemeny's (1959) conception of a distance between orderings. In the third section, we are going to discuss the exact rules of interaction, upon which our social influence model is based. In fact, we are postulating the type of interactions which is followed in stochastic theories of interacting particle systems. Furthermore, in this section, we are going to classify the equilibrium states and to state a theorem which sets the conditions under which the equilibrium outcomes (orderings) will be unanimous or polarized. It turns out that these conditions refer to how the initial orderings are spread over the graph of orderings. Of course, a technical proof of our main theorem escapes the scope of the present presentation. In the fourth section, we are going to present the results of computer simulations that we have done in order to understand how probable the emergence of polarization is under this model of social influence. Our simulations are implemented on rather simple types of social networks: stars, rings and circular regular lattices, which are also rewired in order to produce 'small worlds' (Watts & Strogatz, 1998; Watts, 1999). In the last (fifth) section, we conclude by a discussion of our 'social influence' model of social choice and its relevance with the famous Impossibility Theorem of Arrow (1951). Furthermore, we are going to outline a number of possible extensions of the present work that we intend to pursue in the future.

Fundamental Concepts from Social Choice Theory

We start with a presentation of the fundamental definitions and concepts of the theory of social choice in the discrete case (Mas-Colell *et al.*, 1995; Schofield, 2003). Given a population of n **individuals** and a finite set X with $|X| = m \geq 3$, let us suppose that each individual has to order (or rank), in his/her own way, the elements of the set X, which are called **alternatives**. Sometimes, alternatives could be interpreted as **candidates**, in which case it would be proper that individuals might be called **voters**, with the understanding that this is not a voting of a single candidate/alternative among all others but a positioning of all candidates/alternatives in an **ordering** (or **ranking**) according to each voter's/individual's choice. Typically, individuals/voters will be denoted with symbols like i, j, k, …, while alternatives/candidates with x, y, z, ….

In particular, when an individual/voter is asked to order alternatives/candidates, he/she will do this through a preference ordering R, which is a **binary relation** $R \subseteq X \times X$. When we want to associate an ordering to a particular individual/voter, say, i, we will denote it by R_i but we are going to drop the subscript when it is clear to which individual/voter we are referring. Moreover, we might denote '$(x,y) \in R$' by 'xRy' and by this we mean that 'x is preferred (by an individual/voter) more or at least as much as y.' In particular, we assume that any preference ordering R is a **rational preference relation** (Mas-Colell *et al.*, 1995, p. 6), i.e., a **weak ordering**. By this, it is meant that R is a complete and transitive binary relation.[6] Furthermore, we are going to denote by P, I the asymmetric, symmetric (respectively) part of R, where P is a *strict preference* relation and I an *indifference* relation. Conceiving R, P, I as subsets of $X \times X$, one can easily see that $R = P \cap I$, while $P \cup I = \varnothing$ (Schofield, 2003, p. 28). The set of all weak orderings on X is denoted by $O(X)$.

The Graph of Orderings

Next, let us consider an ordering $R \in O(X)$ and k distinct alternatives $x_1, x_2, \ldots, x_k$, where $2 \leq k \leq m$. We say that $\{x_1, x_2, \ldots, x_k\}$ forms a **chain of alternatives** under R if, for all $i = 1, \ldots, k-1$, $(x_i,x_{i+1}) \in R$ and there exists no other alternative $z \in X$ such that both $(x_i,z) \in R$ and $(z,x_{i+1}) \in R$ are valid. In particular, when $k = 2$, a chain of two distinct alternatives $\{x,y\}$ will be called **pair of successive alternatives**. When $\{x_1, x_2, \ldots, x_k\}$ is a chain of alternatives under R such that $(x_i,x_{i+1}) \in I$, for all $i = 1, \ldots, k-1$, we say that $\{x_1, x_2, \ldots, x_k\}$ forms a **chain of indifferent alternatives**. Similarly, a pair of successive alternatives $\{x,y\}$ is called a **pair of successive indifferent alternatives**, whenever $(x,y) \in I$. Furthermore, if $\{x_1, x_2, \ldots, x_k\}$ (or $\{x,y\}$) is a chain of (or pair of successive) indifferent

[6] A binary relation R on X is *complete* when, for all $x, y \in X$, either xRy or yRx or both and *transitive* when, for all $x, y, z \in X$, xRy and yRz imply xRz. Completeness implies that R is *reflexive* too, i.e., for all $x \in X$, xRx (Mas-Colell *et al.*, 1995, p. 6).

alternatives, we write $R = \ldots(x_1x_2\ldots x_k)\ldots$ (or $R = \ldots(xy)\ldots$), while if they are chains of alternatives, both composed only of strict preference relations, we write $R = \ldots x_1x_2\ldots x_k\ldots$ (or $R = \ldots xy\ldots$). Finally, if $\{x_1,x_2, \ldots, x_p\}$, $\{y_1,y_2, \ldots, y_q\}$ (where $2 \leq p, q \leq m$) are two chains of alternatives, then they are called **disjoint** if $(x_i, y_j) \in P$, for all $i = 1, \ldots, p, j = 1, \ldots, q$.

We say that two distinct orderings $R, Q \in O(X)$ form an elementary alteration of the type of a **demi-transposition** of each other with respect to two alternatives x,y under both R, Q, when these orderings differ only in their relative ordering of x, y in a 'minimal' way. By this we mean that in order to transform one ordering into the other, we need just one single elementary or minimal change of preferences: either to change a strict preference (xPy) into an indifference (xIy) or the other way around.[7] In other words, such a demi-transposition can be implemented only as far as one of the orderings is of the form $\ldots xy\ldots$, while the other one is of the form $\ldots(xy)\ldots$. If we take into account that the remaining parts of these two orderings are exactly the same, it is obvious that two orderings R, Q are mutually transposable by such an elementary change if and only if their symmetric difference[8] $R \triangle Q$ includes only sets in the Cartesian product $X\times X$ of the form $\{(x,y)\}$ or $\{(y,x)\}$ but not both. Therefore, defining the **Hamming distance** $h(R,Q)$ between two binary relations R, Q as $h(R,Q) = |R \triangle Q|$ (van Deemen, 1997, p. 193), there is a demi-transposition with respect to some x, y between R, Q if and only if $h(R,Q) = 1$. In social choice theory, this metric is often called **Kemeny distance**, as it was originally introduced by Kemeny (1959).

We write '$R, Q \in E(x,y)$,' when this condition of elementary mutual alteration holds with respect to two alternatives x, y. Thus, one can define (Bossert & Storcken, 1992) the following binary relation $E(X)$ on $O(X)\times O(X)$:

$$E(X) = \{(R,Q) \in O(X)\times O(X): \exists\, x, y \in X, x \neq y, \text{ such that } R, Q \in E(x,y)\}.$$

[7] In this case, by definition, $\{x,y\}$ should necessarily be a pair of successive alternatives.

[8] The **symmetric difference** between two sets A, B,, denoted by $A \triangle B$, is defined as $A \triangle B = (A\backslash B)\cap(B\backslash A)$.

We should remark that not all R, $Q \in O(X)$ are such that $(R,Q) \in E(X)$. In fact, we can explicitly determine which are the pairs of weak orderings R, Q such that $(R,Q) \notin E(X)$. In these pairs, one of the two orderings should include a chain of three or more indifferent alternatives,[9] while the other should be composed of only strict preferences or it could also include disjoint pairs of indifferent alternatives. Then there is no single demi-transposition between such pairs of weak orderings. Essentially, this is because transitivity would necessitate at least two demi-transpositions to transform one of these orderings into the other, while the previous relation was defined through a single demi-transposition.

Let us denote by $O^*(X)$ the set of all weak orderings, any two of which can be related according to the relation $E(X)$. By the previous remark, $O^*(X)$ is composed of all orderings including either no indifferences at all or just disjoint pairs of indifferent alternatives. In other words, all orderings including chains of indifferent alternatives of length larger than two (i.e., cycles of three or more indifferent alternatives) are excluded from $O^*(X)$. Apparently, $L(X) \subset O^*(X) \subset O(X)$ (and both inclusions are strict). Here $L(X)$ denotes the set of all **linear orderings** or **strict-total preferences** (Mas-Colell *et al*., 1995, p. 6), where a binary relation R on X is called linear ordering if it is reflexive, transitive and weakly complete (or total).[10]

Next, let us consider the graph $G(X) = (O^*(X), E(X))$ with vertex set $O^*(X)$ and edge set $E(X)$ (cf., Bossert & Storcken, 1992, p. 247). For instance, in Figure 1 we show the graph of the 12 vertices/orderings of $O^*(X)$, when $m = 3$, and in Figure 2 we show the graph of the 66 vertices/orderings of $O^*(X)$, when $m = 4$.

<u>Figure 1 about here</u>

<u>Figure 2 about here</u>

[9] Amartya Sen (1966) calls ‘unconcerned’ any individual for whom all pairs of alternatives are indifferent.

[10] A binary relation R on X is *weakly complete* (or *total*) when, for all x, $y \in X$, $x \neq y$, xRy or yRx but not both. Equivalently (Sen 1970, p. 9), a binary relation R on X is a linear ordering if it is reflexive, transitive, complete and *anti-symmetric* (for all x, $y \in X$, xRy and yRx imply $x = y$). In other words, a weak ordering is linear if and only no distinct alternatives can be indifferent to each other.

In Table 1 we give the number of elements of $O^*(X)$ and $O(X)$ for different numbers of alternatives (from 3 to 11).

Table 1 about here

Consequently, for any two orderings R, $Q \in O^*(X)$, conceived as vertices in the graph $G(X)$, a **distance** $d(R,Q)$ can be defined as the length of any shortest path (geodesic) between the two vertices R, Q in the graph $G(X)$. Apparently, R, Q are adjacent vertices in $G(X)$ if and only if $d(R,Q) = h(R,Q) = 1$ (since adjacency is equivalent to the existence of a single demi-transposition between the two orderings, i.e., R, $Q \in E(X)$).

By a simple combinatorial calculation, one can easily obtain that the **maximum distance** between any two orderings in $O^*(X)$ is equal to $M = m(m - 1)$. In particular, any two orderings R, Q in $O^*(X)$ with $d(R,Q) = M$ will be called **antipodes** and the antipode (often called '**dual**' or '**inverse**' too) of an ordering R will be denoted by $\neg R$ (sometimes denoted as R^{-1}). We will say that a couple Δ of two preference orderings constitutes a **dipole** if it is an antipodal pair, i.e., if it is of the form $\Delta = \{Q, \neg Q\}$, for some $Q \in O^*(X)$,

Furthermore, for any two distinct orderings R, $Q \in O^*(X)$, following Bossert & Storcken (1992, p. 348), we define $[R,Q]$ to be the set of all vertices in $O^*(X)$ lying on any geodesic between R and Q. In other words,

$$[R,Q] = \{S \in O^*(X) \colon d(R,S) + d(S,Q) = d(R,Q)\}.$$

In general, for any finite non-empty set $A \subseteq O^*(X)$, we define $[A]$ to be the set of all vertices in $O^*(X)$ lying on any geodesic between any two elements of A. In other words,

$$[A] = \cap\{[R,Q] \colon R, Q \in A\}.$$

Thus, in terms of $[A]$, a set $A \subseteq O^*(X)$ is called:

(i) **homopolar** whenever $[A] \subset O^*(X)$ but $[A] \neq O^*(X)$;

(ii) **heteropolar** whenever $[A] = O^*(X)$; this happens if and only if there exist three distinct orderings $R, Q, S \in A$ such that $S \in [\neg R, \neg Q]$.

We have already said that each individual (or voter) is assigned to some ordering of preferences among all the alternatives (or candidates). This means that there is an assignment mapping π: $V \to O^*(X)$ such that the ordering $\pi(i) = R_i \in O^*(X)$ is the ordering of preferences of individual (voter) $i \in V$. A **preference profile** (or **configuration**) $\mathbf{R} = (\pi(1), \pi(2), \ldots, \pi(n)) = (R_1, R_2, \ldots, R_n)$ is an n-tuple of individual preference orderings of all individuals in V, i.e., $\mathbf{R} \in O(X)^n$.

As before, a preference profile $\mathbf{R} = (R_1, R_2, \ldots, R_n) \in O^*(X)^n$ is called homopolar/heteropolar if the set $\{R_1, R_2, \ldots, R_n\} \subseteq O^*(X)$ is such. Furthermore, $\mathbf{R}$ is called:

(i) **monopolar** whenever $R_1 = R_2 = \ldots = R_n$.

(ii) **bipolar** whenever there exists a partition of the set $V = \{1, 2, \ldots, n\}$ into two subsets U and W such that $R_j = Q$, for all $j \in U$, and $R_k = \neg Q$, for all $k \in W$, for some $Q \in O^*(X)$.

Interactions on a Social Influence Network

We first start with a description of the social influence network that we are going to consider in our model. Let us suppose that an **undirected graph** $G = (V,E)$ is given and that the **vertex** set $V = \{1,2, \ldots, n\}$, $|V| = n \geq 2$, represents a population of n **individuals** and the set E of **links** (**edges**) represents the existing **relationships** among these individuals. For the sake of simplicity, we are going to stick to the undirected graph case: a more general model of social influence based on a **valued directed graph** could be easily constructed by doing some direct modifications in the definition of the social space of the present model. Let us add that the graph G is supposed to be **simple** (without self-loops or multiple links) and **connected** in the sense that a chain of links connects any two vertices/individuals. Since we intend to adopt the terminology of social network analysis (Wasserman & Faust, 1994), 'individuals'/'voters' from now on we will be named 'actors.'

Following the formalization of interacting particle systems used in the context of stochastic systems (the standard reference is Liggett, 1985; see also Axelrod, 1997), our social influence model can be described as an iterated two-stage process of updating actors' orderings through successive random dyadic interactions among actors.[11] First, an actor, say, i, is chosen together with one of his/her neighbors, say, j. Next, with probability $p_{i,j} \in [0,1]$, an interaction between i and j occurs and it results a possible change of the ordering $\pi(i)$ of actor i because of the influence from actor j (this is what we call 'dyadic interaction'). In this way, actors i, j are 'interacting'; in particular, i is the 'influenced' and j is the 'influencing' actor.[12]

Now, the exact rules of how the dyadic interaction between actors i and j proceeds are the following: Assuming that the orderings of all actors are in $O^*(X)$, we compare the ordering $\pi(i) = R_i$ of the influenced actor i (who was chosen first, either randomly or arbitrarily) with the ordering $\pi(j) = R_j$ of the influencing actor j (who was chosen second, again either randomly or arbitrarily, among all i's neighbors). If the two orderings R_i, R_j coincide or they are antipodal, then no interaction occurs; this means that the conditions of interaction halt are $d(R_i,R_j) = 0$ or $d(R_i,R_j) = M = m(m - 1)$. Under any other circumstances ($0 < d(R_i,R_j) < M$), the interaction leads to a relative approaching of the ordering R_i of the influenced actor i towards the ordering R_j of the influencing actor j in the sense that R_i is transformed to an ordering R_i' such that $d(R_i', R_j) = d(R_i,R_j) - 1$. Unless $m = 3$, the transformed ordering R_i' could possibly be non unique but we could always have chosen randomly one among all the orderings which reduce $d(R_i,R_j)$ by 1.

[11] If we wanted to be more accurate in the formalization of interacting particle systems, we should have said that the updating times follow a Poisson process with rate 1.

[12] In the case that the social influence network was represented by a valued directed graph, we could have considered that actor i could influence an adjacent actor j only as far as there was a directed link (arc) from i to j.

When a dyadic interaction occurs ($0 < d(R_i,R_j) < M$), the probability $p_{i,j}$ can be taken as one in the following three cases:

(i) $p_{i,j} = 1$, i.e., interaction is certain to occur,

(ii) $p_{i,j} = (M - d(R_i,R_j))/M$, i.e., the probability of interaction equals the normalized 'similarity' between the two orderings, where 'similarity' is understood in the graph-distance sense, and

(iii) $p_{i,j} = 1/d(R_i,R_j)$, i.e., the probability of interaction is inversely proportional to the graph-distance between the two orderings.[13]

Iterating (repeating successively) many times the above rules of interaction (but each time randomly or arbitrarily selecting the pair of interacting actors), obviously all actors are almost surely going to be selected to interact with their neighbors. This is an example of a stochastic process, which is called **Markov random field** (Spitzer, 1971; Brémaud, 1999). Typically, this process eventually reaches an **equilibrium** state. Because of the two conditions of interaction halt, it is obvious which configurations of orderings should survive in an equilibrium state: In equilibrium, all adjacent actors should necessarily have either the same or antipodal orderings. Thus, using the terminology of Brian Arthur (1989), we might say that the two possible equilibrium **lock-ins** are either a single (unanimous) ordering or a dipole of two antipodal orderings. Of course, which one of these two configurations is eventually going to emerge in equilibrium depends on the starting initial distribution of orderings over actors. In other words, again using Arthur's terminology (p. 117), the process we are studying is '**path-dependent**' or '**nonergodic**' (since there exist more than one equilibria).

Although the probabilities of lock-ins inside particular equilibrium states (single-unanimous orderings or dipoles of antipodal orderings) are extremely hard to be computed analytically, we can say something about the wider areas where equilibrium lock-ins should be located, when actors' orderings start from certain initial

[13] In the case of a valued directed graph, the above definitions of the interaction probabilities $p_{i,j}$ could have been modified by multiplication with the normalized value of the directed link (arc) from i to j.

configurations. For any actor $i \in V$, let us denote by $R_i^o \in O^*(X)$ his/her initial ordering and by $R_i^e \in O^*(X)$ his/her equilibrium ordering, i.e., the ordering where this actor's preferences are locked-in through the sustained social influence interactions with adjacent actors. We set $\mathbf{R}^o = \{R_i^o: i \in V\}$ and $\mathbf{R}^e = \{R_i^e: i \in V\}$ for the profiles of orderings initially and at equilibrium (respectively); apparently, $\mathbf{R}^o$, $\mathbf{R}^e \in O^*(X)^n$. Then our main result describing the way this process of social influence is stabilized reads as follows:

Theorem 1.

(i) If $\mathbf{R}^o$ is homopolar, then $\mathbf{R}^e$ is monopolar and $\{R_i^e: i \in V\} = \{R\}$, for some $R \in [\mathbf{R}^o]$.

(ii) If $\mathbf{R}^o$ is heteropolar, then either $\mathbf{R}^e$ is monopolar and $\{R_i^e: i \in V\} = \{R\}$, for some $R \in O^*(X)$, or $\mathbf{R}^e$ is bipolar.

Simulation Results

We have run a number of different computer simulation[14] experiments in order to understand statistically the structure of the equilibrium states of this social influence process. These simulations were implemented for the case of three alternatives (i.e., m = 3 and $|O^*(X)| = 12$). In each simulation experiment, we have started from a randomly (or arbitrarily) chosen initial profile of orderings and have proceeded with random selections of interacting dyads of actors until reaching an equilibrium (interaction halt). Furthermore, each simulation converges to an equilibrium profile of orderings consisting of either a single unanimous ordering or a dipole of two antipodal orderings. After having run a simulation N times (where N is a large number), let B_N be the total number of dipoles produced throughout these simulations. Moreover, let the average $\langle B \rangle = B_N / N \in [0,1]$ be called **average bipolarity index**. Then, by running our simulations over specific types of social influence networks, it is interesting to examine how the average bipolarity index varies on certain simple graphs in terms of the number and the degree of vertices/actors and the level of graph randomness.

[14] These simulation were implemented through the computer software CHABRA (Mantzoukas, 2004).

The first simulation experiment corresponds to the case that the social influence network is a star. Then we have found the following dependence between the average bipolarity index $<B>$ and the number n of actors:

Figure 3 about here

What explains this form of variation of $<B>$ with n is the fact that, as the number of actors increases, the probability of the formation of dipoles should also increase but this can go up to a limit (an upper bound) due to the bounded number of all possible dipoles (6 when $m = 3$).

Next, we have conducted a simulation experiment with the ring topology of the social influence network. Furthermore, we have generated the corresponding random graphs by the 'rewiring algorithm' (Watts & Strogatz, 1998). We have found the following dependence between the average bipolarity index $<B>$ and the number n of actors of both rings and rewired rings:

Figure 4 about here

In this case, it is also expected that the average polarization should diminish as the diameter of the rings increases, because by pumping in more actors in the ring but keeping fixed their connectivity (always degree 2 for each vertex in the ring), it would be increasingly more difficult to polarize the total population of actors across a dipole. However, the effect of the randomization caused by the process of rewiring would be to produce actors/vertices of higher degree and, thus, to cause polarization proliferate at a higher rate than in the regular ring case.

In the last simulation experiment, we have been rewiring a circular regular lattice composed of 50 vertices/actors, each one having degree equal to 10 (i.e., 5 links to the right and 5 links to the left). Denoting by β the 'rewiring parameter,' we have found the following dependence between the average bipolarity index $<B>$ and β:

Figure 5 about here

Of course, one would expect an increase in polarization as $\beta \to 1$ ($\beta = 1$ corresponds to a random graph) since the level of the graph randomization caused by the rewiring algorithm increases with β. This means that randomization increases the probability of building random bridges among actors/vertices and, thus, it facilitates the formation of dipoles (polarization). What is surprising is that in these simulations we observe an outburst of polarization just to the left of the random graph condition $\beta = 1$, although the polarization level is constantly and uniformly low in the biggest part of the interval (0,1) of the values of the parameter β (note the scale of the abscissa in this diagram is logarithmic). In fact, Watts & Strogatz (1998; Watts, 1999) have postulated that the realm of the "small world" network is exactly in the same region of values of β where the polarization outburst occurs. The apparent conclusion which can be drawn from this simulation says that polarization in the outcomes of a social influence process tends to be maximized over "small world" graphs, which many believe that they depict more realistic representations of the actual social networks constituted by the common social relationships among people (Watts, 2003).

Conclusions

First, let us recall the famous 'Impossibility Theorem' of social choice of Kenneth Arrow (1951), a result which investigates the logical consistency among a number of plausible properties of a 'social welfare function,' i.e., a mapping F: $O(X)^n \to O(X)$. In fact, Arrow (1951) has shown that, in the case $m \geq 3$, there is no social welfare function satisfying all the following three properties: 'independence of irrelevant alternatives,' a 'weak Pareto condition' and 'non-dictatorship.'

Let us next consider the possible equilibrium outcomes of our model of social influence. We have seen in our main result that such a process can stabilize in either a monopolar or a bipolar equilibrium profile of orderings, depending on how the initial profile of orderings is spread over the graph of orderings $O^*(X)$ (i.e., depending on whether the initial profile is homopolar or heteropolar).

In particular, we have found that homopolar initial profiles are always stabilized into a monopolar equilibrium state (representing a unanimous ordering by all actors in the social influence network). In case that the monopolar equilibrium profile of orderings happens to be one of the profiles of orderings that one of the actors of the system was holding initially, then this actor might be called 'dictator' and his/her profile of orderings 'dictatorial' in accordance to Arrow's Theorem. However, in this case, dictatorship is generated by a process of social influence, which was explicitly neglected in the fundamental result of social choice theory. If the monopolar equilibrium profile of orderings was not held by any actor initially, then we could call it 'emergently dictatorial.'

Moreover, we have seen that in our model of social influence there is a second possibility giving rise to polarized configurations. If the initial profile of actors' orderings was heteropolar, then the equilibrium outcome could be either monopolar (a unanimous ordering, which was interpreted as dictatorial or emergently dictatorial above) or a dipole of two antipodal orderings. Again there are two cases depending on whether or not the poles of the equilibrium dipole were held initially by certain actors. We say that in the former case a 'dictatorial dipole' dominates through the dynamics of this social influence process and in the latter case that an 'emergently dictatorial dipole' is established through the conflictual mechanisms of this process.

So, what we have found in our model of social choice, driven by social influence, is still in accordance with Arrow's inevitability of dictatorial outcomes but, furthermore, in our case we have been able to derive a 'dictatorial polarization.' Such a polarization could either survive from the beginning and eventually dominate through all actors or could emerge in the middle of the process of social influence. In other words, when the relational structure of the space of influence outcomes (the graph of orderings) allows the existence of antipodal orderings, then a simple mechanism of social influence based on random dyadic interactions is shown to produce either unanimity or polarization depending on the initial configuration of actors' orderings.

Of course, our assumption of dyadic interactions was not really necessary and we have followed it just for the sake of simplicity in the statement and the proof of our main arguments. We could work the same social influence model but based on more complex and realistic modes of actors' interaction (for instance, threshold models, taking into account various protocols of aggregation of orderings, i.e., types of votings). Moreover, as we have already remarked, a more general social influence model would need to be based on a social influence network, which would be a valued directed graph, and it could also predict fluctuating unstable (even chaotic) collective behavior. One could even consider the case of a 'signed' valued graph, where links among actors/vertices with negative sign could be interpreted as channels of interaction producing the opposite effect to convergence, i.e., a divergence or a relative remotion of the orderings of two interacting actors, when the two actors are related through a 'negative' relationship. Another further extension of our model would be to allow a dynamic 'reshuffling' of the links (relationships) of the social influence network, driven by a propensity towards homophilous interactions that actors might have. Of course, in this case, we would expect to find an eventual segmentation of the starting social influence network into components of unanimous orderings with a number of them being antipodal to each other. These are among the lines of research through which we intend to advance our future work in order to generalize further the present model of social influence based on a structured space of influence outcomes.

Appendix

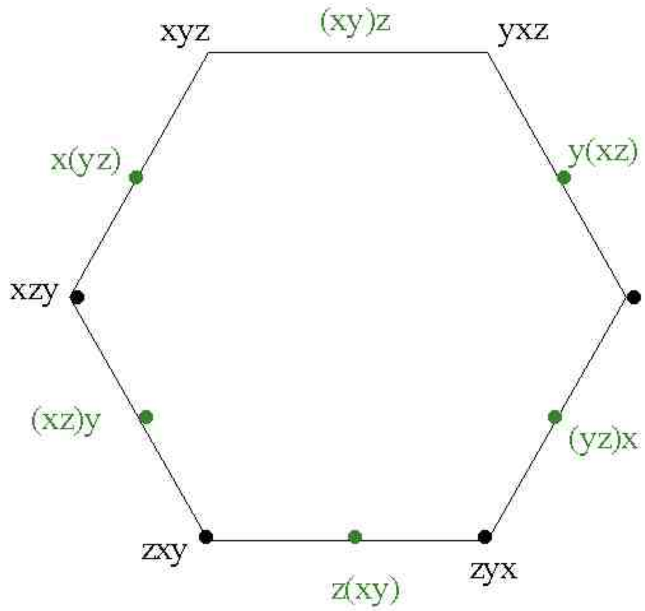

Figure 1. $O^*(X)$ for $m = 3$.

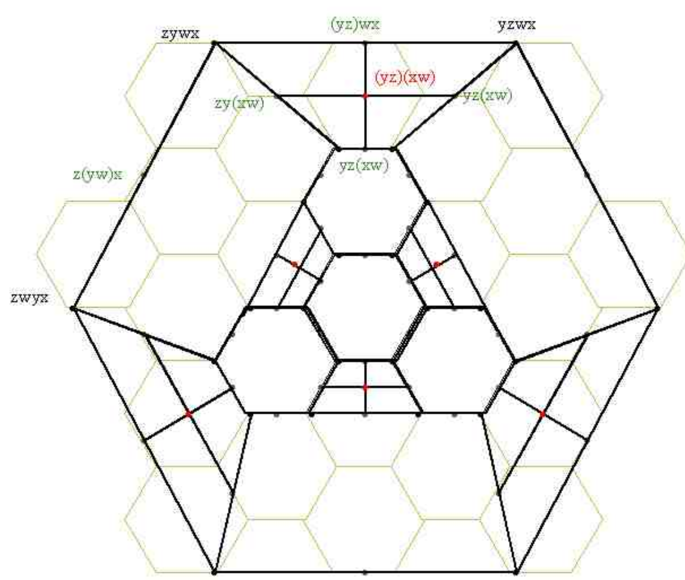

Figure 2. $O^*(X)$ for $m = 4$.

m	$\|O^*(X)\|$	$\|O(X)\|$
3	12	13
4	66	75
5	450	521
6	3690	4293
7	35280	40881
8	385560	442599
9	4740120	5375941
10	64751400	72475263
11	972972000	1074691949

Table 1. Cardinality of $O^*(X)$ and $O(X)$ for different numbers of alternatives.

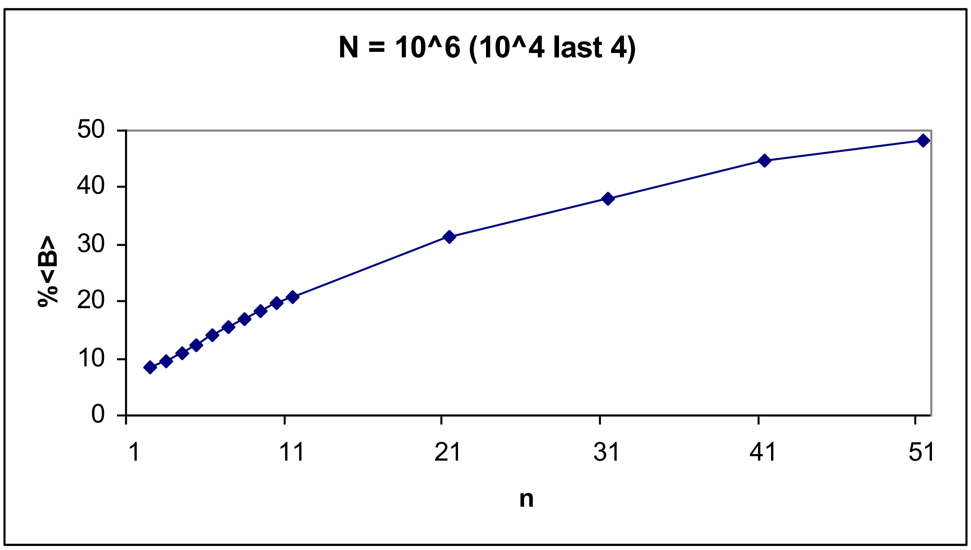

Figure 3. Polarization over stars.

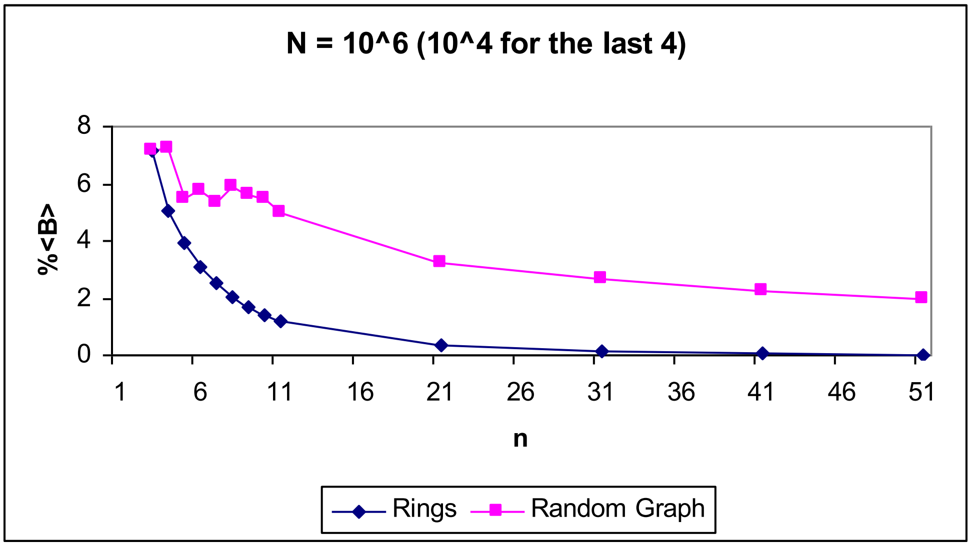

Figure 4. Polarization over rings and random graphs generated by rewiring rings.

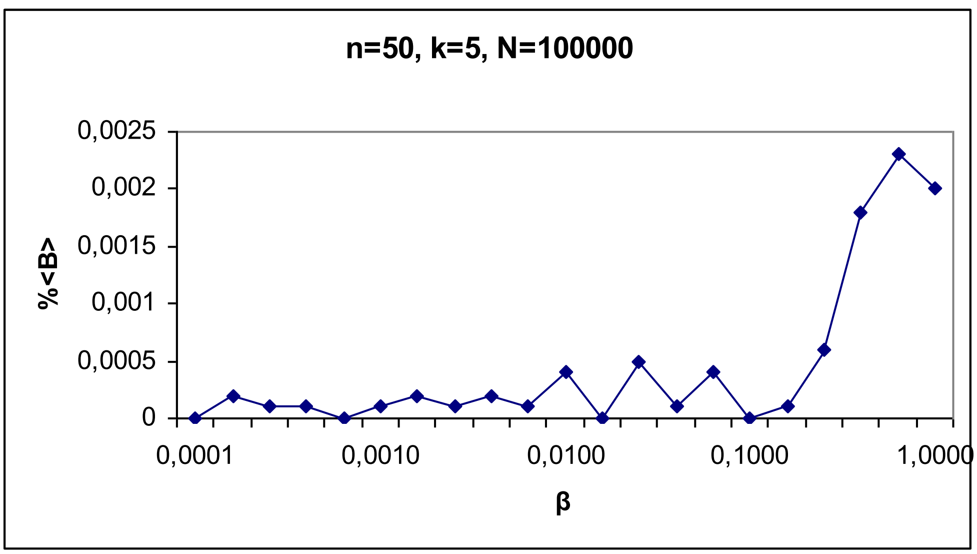

Figure 5. Polarization on a cyclical regular lattice in terms of the rewiring parameter.

References

Abelson, R.P. (1964). Mathematical models of the distribution of attitudes under controversy. In N. Frederiksen & H. Gulliksen (eds.), *Contributions to Mathematical Psychology*, pp. 142-160. New York: Holt, Rinehart & Winston.

Arrow, K.J. (1951). *Social Choice and Individual Values*. New Haven & London: Yale University Press.

Arthur, W.B. (1989). Competing technologies, increasing returns, and lock-in by historical small events. *Economic Journal*, vol. 99, pp. 116-131.

Axelrod, R. (1997). The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution*, vol. 41, pp. 203-226.

Bossert, W., & Storcken, T. (1992). Strategy-proofness of social welfare functions: The use of the Kemeny distance between preference orderings. *Social Choice and Welfare*, vol. 9, pp. 345-360.

Bramoullé, Y. (2002). Complementarity and social networks. CSED Working Paper No. 24. The Brookings Institution.

Bramson, M., & Griffeath, D. (1989). Flux and fixation in cyclic particle systems. *The Annals of Probability*, vol. 17, no. 1, pp. 26-45.

Brémaud, P. (1999). *Markov Chains. Gibbs Fields, Monte Carlo Simulation, and Queues*. New York: Springer-Verlag.

Coleman, J.S. (1990). *Foundations of Social Theory*. Cambridge, MA, & London: The Belknap Press of Harvard University Press.

Durlauf, S.N., & Young, H.P. (eds.) (2001). *Social Dynamics. Cambridge*, MA, & London: The MIT Press.

Durrett, R. (1988). *Lecture Notes on Particle Systems and Percolation*. Pacific Grove, CA: Wadsworth & Brooks/Cole.

Durrett, R. (1999). Stochastic spatial models. *SIAM Review*, vol. 41, pp. 677-718.

Flache, A., & Torenvlied, R. (2004). When will they ever make up their minds? The social structure of unstable decision making. *Journal of Mathematical Sociology*, vol. 28, pp. 171-196.

French, J.R.P., Jr. (1956). A formal theory of social power. *Psychological Review*, vol. 63, pp. 181-194.

Friedkin, N.E. (1999). Choice shift and group polarization. *American Sociological Review*, vol. 64, pp. 856-875.

Friedkin, N.E. (2001). Norm formation in social influence networks. *Social Networks*, vol. 23, pp. 167-189.

Friedkin, N.E., & Johnsen, E.C. (1990). Social influence and opinions. *Journal of Mathematical Sociology*, vol. 15, no. 3-4, pp. 193-205.

Friedkin, N.E., & Johnsen, E.C. (1999). Social influence networks and opinion change. *Advances in Group Processes*, vol. 16, pp. 1-29.

Giddens, A. (1984). *The Constitution of Social Life: Outline of a Theory of Structuration*. Berkeley: University of California Press.

Griffeath, D. (1988). Cyclic random competition: A case history in experimental mathematics. *Notices of the American Mathematical Society*, vol. 35, pp. 1472-1480.

Harary, F. (1959). A criterion for unanimity in French's theory of social power. In D. Cartwright (ed.), *Studies in Social Power*, pp. 168-182. Ann Arbor: Institute for Social Research.

Kemeny, J.G. (1959). Mathematics without numbers. *Daedalus*, vol. 88, pp. 577-591.

Kobayashi, J. (2001). Unanimous opinion in social influence networks. *Journal of Mathematical Sociology*, vol. 25, pp. 285-297.

Latané, B. (1981). The psychology of social impact. *American Psychologist*, vol. 36, pp. 343-365.

Laumann, E.O. (1973). *Bonds of Pluralism: The Form and Substance of Urban Social Networks*. New York: Wiley Interscience.

Laumann, E.O., & Pappi, F.U. (1976). *Networks of Collective Action*. New York: Academic Press.

Lewenstein, M., Nowak, A., & Latané, B. (1992). Statistical mechanics of social impact. *Physical Review A*, vol. 45, no. 2, pp. 763-776.

Liggett, T.M. (1985). *Interacting Particle Systems*. New York: Springer Verlag.

Liggett, T.M. (1997). Stochastic models of interacting systems. *The Annals of Probability*, vol. 25, no. 1, pp. 1-29.

Liggett, T.M. (1999). *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Berlin: Springer Verlag.

Macy, M.W. (1990). Learning theory and the logic of critical mass. *American Sociological Review*, vol. 55, pp. 809-826.

Mantzoukas, A. (2004). Computer software CHABRA for agent-based simulations. University of Patras, Department of Mathematics.

Mas-Colell, A., Whinston, M.D., & Green, J.R. (1995). *Microeconomic Theory*. New York & Oxford: Oxford University Press.

Nakamaru, M., & Levin, S.A. (2004). *Journal of Theoretical Biology*, vol. 230, pp. 57-64.

Nowak, A., Szamrej, J., & Latané, B. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review*, vol. 97, no. 3, pp. 362-376.

Schofield, N. (2003). *Mathematical Methods in Economics and Social Choice*. Berlin, Heidelberg, New York: Springer-Verlag.

Sen, A.K. (1966). A possibility theorem on majority decisions. *Econometrica*, vol. 34, no. 2, pp. 491-499.

Sen, A.K. (1970). *Collective Choice and Social Welfare*. San Francisco: Holden Day, Inc.

Sewell, W.H., Jr. (1992). A theory of structure: Duality, agency, and transformation. *American Journal of Sociology*, vol. 98, no. 1, pp. 1-29.

Spitzer, F. (1971). *Random Fields and Interacting Particle Systems*. MAA Summer Seminar at Williams College. Washington, DC: Mathematical Association of America.

Stokman, F.N., & Van den Bos, J.M.M. (1992). A two-stage model of collective decision making with an empirical test in the U.S. energy policy domain. *Research in Politics and Society*, vol. 4, pp. 219-253.

van Deemen, A.M.A. (1997). *Coalition Formation and Social Choice*. Boston, Dordrecht, London: Kluwer Academic Publishers.

Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge, UK: Cambridge University Press.

Watts, D.J. (1999). *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton, NJ: Princeton University Press.

Watts, D.J. (2003). *Six Degrees. The Science of a Connected Age*. New York & London: W.W. Norton & Company.

Watts, D.J., & Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks. *Nature*, vol. 393 (June 4), pp. 440-442.

Young, H.P. (1998). *Individual Strategy and Social Structure. An Evolutionary Theory of Institutions*. Princeton & Oxford: Princeton University Press.

Young, H.P. (2005). The diffusion of innovations in social networks. In L.E. Blume & S.N. Durlauf (eds.), *The Economy As an Evolving Complex System*, vol. III. New York & Oxford: Oxford University Press.