Dynamics of Conformist Bias

Brian Skyrms

Abstract: We compare replicator dynamics for some simple games with and without the addition of conformist bias. The addition of conformist bias can create equilibria, it can change the stability properties of existing equilibria, it may leave the equilibrium structure intact but change the relative size of basins of attraction, or it may do nothing at all. Examples of each of the foregoing are given.

1. **Replicator Dynamics**. The most thoroughly studied dynamic model for cultural evolution is the replicator dynamics. [Taylor and Jonker (1978) See Hofbauer and Sigmund (1998) for a comprehensive treatment.] It was originally proposed as an account of differential reproduction based on haploid genetics. But it also can be motivated as a model of cultural evolution based differential imitation [Schlag (1997)] - with the more successful strategies being imitated more often. The population is assumed to be large enough so that a deterministic dynamics can be a useful approximation to the true noisy behavior. For strategies A_i, we write $U(A_i)$ for the average value to A_i in the population, $P(A_i)$ for the proportion of the population using strategy A_i and $Ubar = \Sigma_i P(A_i) U(A_i)$ for the average value in the population. Then the replicator dynamics is given by the system of differential equations:

$$dP(A_i)/dt = P(A_i)\,[U(A_i)-Ubar]$$

Above-average value (fitness, utility) leads to positive growth in population proportion; below-average value leads to diminishing population proportion.

2. **Value**. Average value to a strategy may come from average payoffs to that strategy when played against randomly chosen opponents from the population. We will be interested in some well-studied two-person games- The Winding Road, Stag Hunt, Prisoner's Dilemma, Rock-Scissors-Paper. Payoff to one player is determined not only by his own strategy, but also that of the player with whom he is paired. We assume that the identities of the players are not important, only the strategies played. Thus payoffs are specified by a matrix V_{ij}, which gives the value of A_i when played against A_j. In a large population with random encounters, where all value comes from the game interactions, we can take the average value of a strategy, A_i, to be:

$$\text{Payoff Value} = \Sigma_j\, P(A_j)\, V_{ij}$$

This is the analysis familiar from evolutionary game theory.

But in a society with conformist bias, the value of a strategy - that which leads it to be differentially imitated - may have two components: the payoff component and the conformist component. Overall value can be gotten as a weighted average:

$$U(A_i) = (1\text{-}c) \text{ Payoff Value} + c \text{ Conformist Value}$$

where the constant c determines the strength of conformist bias. If the conformism constant c=1, conformism counts for everything; if c=0 it counts for nothing.

Conformist value of a strategy increases as the strategy becomes more common. We should have the conformist value of a strategy be some monotonically increasing function of the population proportion of the strategy. Here we will take the simplest choice and let the conformist value simply be the population proportion of the strategy at issue. We can then rewrite overall value as:

$$U(A_i) = (1\text{-}c) \Sigma_j P(A_j) V_{ij} + c P(A_i)$$

Overall value then feeds into the replicator dynamics.

For a given type of interaction, we can start with c=0 and then see how the dynamics changes when we add conformist bias. The most dramatic effects that we might encounter consist in the creation of new equilibria or the destruction of old ones. Short of this, we might see changes in the stability of equilibria, with stable equilibria becoming unstable or unstable ones being stabilized. Even if the equilibrium structure is unaltered, the size of basins of attraction for existing equilibria might be changed. Or, perhaps, addition of conformist bias might leave the dynamical system essentially unchanged.

3. **Interactions**. The effect of conformist bias will be different relative to different kinds of interactions. Here we survey the evolutionary dynamics of a number of paradigmatic interactions with and without conformist bias.

3.1 Winding Road I

	Right	Left
Right	1	0
Left	0	1

This is a classic pure coordination game that nicely illustrates the virtues of conformism. The choice is to drive on the left or on the right. The conventions of everyone driving on the left or e·'eryone on the right are equally good. It is only important that all choose the same side of the road. Applying the replicator dynamics, we find that the two possible conventions - 100% left and 100% right - are the only stable equilibria, with and additional unstable equilibrium at the point where exactly half the population drives on the left and half on the right. If more than half the population drives on the right, then the dynamics carries it to the equilibrium when all drive on the right; that is to say that the basic of attraction of this equilibrium consists of all population proportions Pr(R) > .5. Likewise, the basin of attraction of 100% left equilibrium consists of populations with Pr(L) < .5.

When we add conformist bias to any degree, nothing changes. The dynamics is exactly the same. The overall value of Left is a weighted average of the payoff value of left [= Pr(L)] and the conformist value of left [=Pr(L)], likewise with right, so the addition of conformist bias adds nothing. The structure of the interaction by itself generates conformism, with or without conformist bias.

3.2 The Stag Hunt.

	Stag	Hare
Stag	4	0
Hare	3	3

The Stag Hunt is a coordination game, but not a pure coordination game. There are, as before, three equilibria. The two population monomorphisms, All hunt stag and All hunt hare, are stable attractors in the replicator dynamics. There is also an unstable polymorphic equilibrium at Pr(Stag) = .75. The basin of attraction of the Hare hunting equilibrium [Pr (Stag) < .75] is three times as large as that of the Stag hunting equilibrium [Pr (Stag) > .75]. From the point of view of social welfare this is a shame, because everyone is better off at the Stag Hunting equilibrium.

If we add in some conformist bias, the basic equilibrium structure remains the same - the two stable monomorphisms and the one unstable polymorphism - but the polymorphic equilibrium moves toward the center and the basin of attraction of Hare hunting is diminished. With 100% conformism, the picture would look just like The Winding Road.

Conformist bias has some socially positive effect by decreasing the riskiness of Stag Hunting, and - in a sense - making the dynamical picture more favorable to the socially efficient equilibrium. We can't really generalize from this example, even in coordination games. Consider the following.

3.3 Winding Road II.

	Right	Left
Right	3	0
Left	0	1

This is the winding road for a population who are all blind in the right eye. There must be a better story, but the point is that although it is still a pure coordination game one equilibrium is better for everyone than the other. Consequently, the equilibrium where all drive on the right has a greater basin of attraction. If $Pr(R) > .25$, the dynamics carries the population to All Right; if $Pr(R) < .25$, the dynamics carries the population to All Left. The dynamical picture looks like that for the Stag Hunt, with Right for Hare and Left for Stag.

If we add conformist bias, the basin of attraction for All Right shrinks and approaches .5 as overall value approaches pure conformism. But here, unlike in the Stag Hunt game, conformism works against mutual benefit rather that for it.

3.4 Rock-Scissors-Paper I

	R	S	P
R	1	2	0
S	0	1	2
P	2	0	1

Rock breaks scissors, scissors cuts paper, paper covers rock. [For a more interesting example of this kind of cyclic structure in a public goods provision game with optional participation, see Hauert et al.(2002).] The three possible monomorphisms of the population (All Rock, All Scissors, All Paper) do not correspond to Nash equilibria. They are, or course, dynamical equilibria under the replicator dynamics[other types are extinct], but they are dynamically unstable [at each, another type could invade]. Thus, at the population state All Rock, a few mutants who play Paper could invade.

There is a unique polymorphic equilibrium with 1/3 of the population playing each of the strategies. The first place to look for information about the stability of this equilibrium is the Jacobian matrix of partial derivatives for the dynamics. If all eigenvalues of the Jacobian have negative real part, then the equilibrium is an attractor. If there there is at least one eigenvalue with positive real part, it is unstable. [See Hofbauer and Sigmund (1998)] The dynamics can be written in terms of Pr(R) and Pr(S) since the population proportions must sum to one. Writing the dynamics this way, and evaluating the eigenvalues of the Jacobian at Pr(R)=Pr(S)=1/3, we get [-SQRT(-1/3), SQRT(-1/3)]. The imaginary eigenvalues indicate a rotating motion. Since the real parts of eigenvalues are zero, they do not answer the stability question and other means must be used.

The quantity Pr(R)*Pr(S)*(1-Pr(R)-Pr(S) is a constant of motion of the system - its time derivative is zero. The orbits of the dynamics must keep this value constant. It assumes its maximum of 1/27 only at the polymorphic equilibrium, Pt(R) = Pr(S) = 1/3. Off the equilibrium, constant values of the conserved quantity correspond to closed curves around the equilibrium. This is illustrated in the contour plot shown in figure 1. These closed curves are the orbits of the dynamics. The equilibrium is dynamically stable because populations near to it cycle around and stay near to it. However it is not asymptotically stable. Populations near to it are attracted to it.

(fig. 1 here)

If we add even the smallest bit of conformist bias to the dynamics, the polymorphic equilibrium at <1/3,1/3,1/3> is destabilized. The general expression for the eigenvalues of the Jacobian at this point, when conformist bias is included, is:

{c/3 - SQRT(-1/3 + 2c/3 - c^2/3), c/3 + SQRT(-1/3 + 2c/3 - c^2/3)}

If c>0, then these eigenvalues have positive real part, which indicate that the equilibrium has become unstable. The time derivative of the product of the proportions of the strategies is no longer a constant of motion. Now this quantity decreases along all orbits in the interior of the space of population proportions. If you start arbitrarily near to the equilibrium, the orbit will spiral outward and approach the boundary.

The stability characteristics of the monomorphic equilibria, however, are not changed by a little conformist bias. They remain dynamically unstable saddle points, with Rock, for example, attracting on the edge connecting it with Scissors but repelling along the edge connecting it with paper.

Adding considerably more conformist bias, however, produces another qualitative change (a bifurcation) in the dynamics. The eigenvalues of the Jacobian at each of the monomorphic equilibria are:

{-c - SQRT(1 - 2c + c^2), -c - SQRT(1 - 2c + c^2)}

with no conformist bias, c=0, these are {1,-1} indicating the unstable saddle. At c=.5, there is a bifurcation, and these values are {-1,0}. With c> .5, both eigenvalues become negative, indicating that the monomorphisms have changed from (unstable) saddles to (strongly stable) attractors. At the same time, continuity considerations tell us that three new unstable equilibria have been created on the edges. The situation with c=0 and with c=.6 are shown in figures 2 and 3, with filled circles representing stable equilibria and open circles representing unstable ones.

(figures 2 and 3 here)

What have all these dynamical fireworks done for the efficiency of the population? If we measure the results in terms of real payoff, without adding in the supposed satisfaction from conformism, the answer must be "Nothing." The average payoff at the original polymorphic equilibrium <1/3,1/3,1/3> is equal to one. A population at a stable monomorphism still gets the same payoff.

3.5 Rock-Scissors-Paper II [Zeeman (1980), Hofbauer and Sigmund (1998)]

	R	S	P
R	1-e	2	0
S	0	1-e	2
P	2	0	1-e

The original Rock-Scissors-Paper game (without conformist bias) was not structurally stable in the replicator dynamics. So the striking results of conformist bias may have been gotten rather cheaply. In Rock-Scissors-Paper II, with small e, we have a game which is structurally stable, and in which the <1/3,1/3,1/3> polymorphism *is* a stable attractor. Eigenvalues of the Jacobian have negative real part. Furthermore, the product of the population proportions increases along all interior orbits and reaches its maximum at <1/3,1/3,1/3>, so this state is a *global* attractor for all the interior of the space (where no type is extinct).. Orbits spiral in to the polymorphic equilibrium. The monomorphisms, All Rock, All Scissors, All Paper, are (unstable) saddles.

As we feed in conformist bias, the dynamics changes qualitatively at $b = e/(1+e)$. The equilibrium at <1/3,1/3,1/3> ceases to be an attractor, but remains stable. The eigenvalues of the Jacobian now have only imaginary parts. The quantity Pr(R)*Pr(S)*(1-Pr(R)-Pr(S) is again a constant of motion of the system, so that <1/3,1/3,1/3> is surrounded by closed orbits. We are in a situation qualitatively similar to the original Rock-Scissors-Paper without conformist bias.

As conformist bias increases, the system goes through all the changes noted in the discussion under Rock-Scissors-Paper I. At the end, the polymorphism has changed from an attractor to a repellor, the monomorphisms have changed from saddles to attractors, and three new (unstable) equilibria have been created. As before, conformist bias confers no collective benefit. A monomorphic population is no better off (in fact, slightly worse off) than a population at the polymorphic equilibrium.

4. **Conclusion.** In the simple setting considered in this paper, conformist bias can have dramatic effects on the dynamics of cultural evolution. These effects are sometimes positive, sometimes negative, and sometimes neutral with respect to the collective welfare of the group. Conformist bias is, therefore, not the same as group solidarity. Group solidarity would presumable work for a Pareto-efficient equilibrium - for mutual benefit, but we have seen (most clearly in 3.3) that conformist bias can work against it.

The story may, of course, be different when the matter is examined in different settings. [Compare Boyd and Richerson (1985) and Henrich and Boyd (1998)].

Interaction between multiple groups can quickly introduce additional complexity. Just moving from one population replicator dynamics to two population replicator dynamics makes a radical difference in Rock-Scissors-Paper games without conformist bias [Sato et a. (2002)]. Evaluation of the effects conformist and other types of bias in more complicated settings raises questions well worth pursuing.

References:

Boyd, R. and Richerson, P. J. (1985) *Culture and the Evolutionary Process*. University of Chicago Press: Chicago.

Hauert, C. , De Monte, S. , Hofbauer, J. and Sigmund, K. (2002) "Volunteering as Red Queen Mechanism for Cooperation in Public Goods Games." *Science* 296, 1129-1132.

Henrich, J. and Boyd, R. (1998) "The Evolution of Conformist Transmission and the Emergence of Between-Group Differences." *Evolution and Human Behavior,* 19: 215–242.

Hofbauer, J. and Sigmund, K.(1998) *Evolutionary Games and Population Dynamics.* Cambridge: N. Y.

Sato, Y. Akiyama, E., and Farmer, J. D. (2002) "Chaos in Learning a Simple Two-Person Game." *Proceedings of the National Academy of Sciences* 99, 4748-4751.

Schlag, K. H. (1997) "Why Imitate, and if so, How? A Boundedly Rational Approach to Multi-Armed Bandits." *Journal of Economic Theory* 78, 130-156.

Taylor, P. D. and Jonker, L. (1978) "Evolutionarily Stable Strategies and Game Dynamics." *Mathematical Biosciences* 40, 145-156.

Zeeman, E. C. (1980) "Population Dynamics From Game Theory" In *Global Theory of Dynamical Systems* Springer Lecture Notes on Mathematics 819.

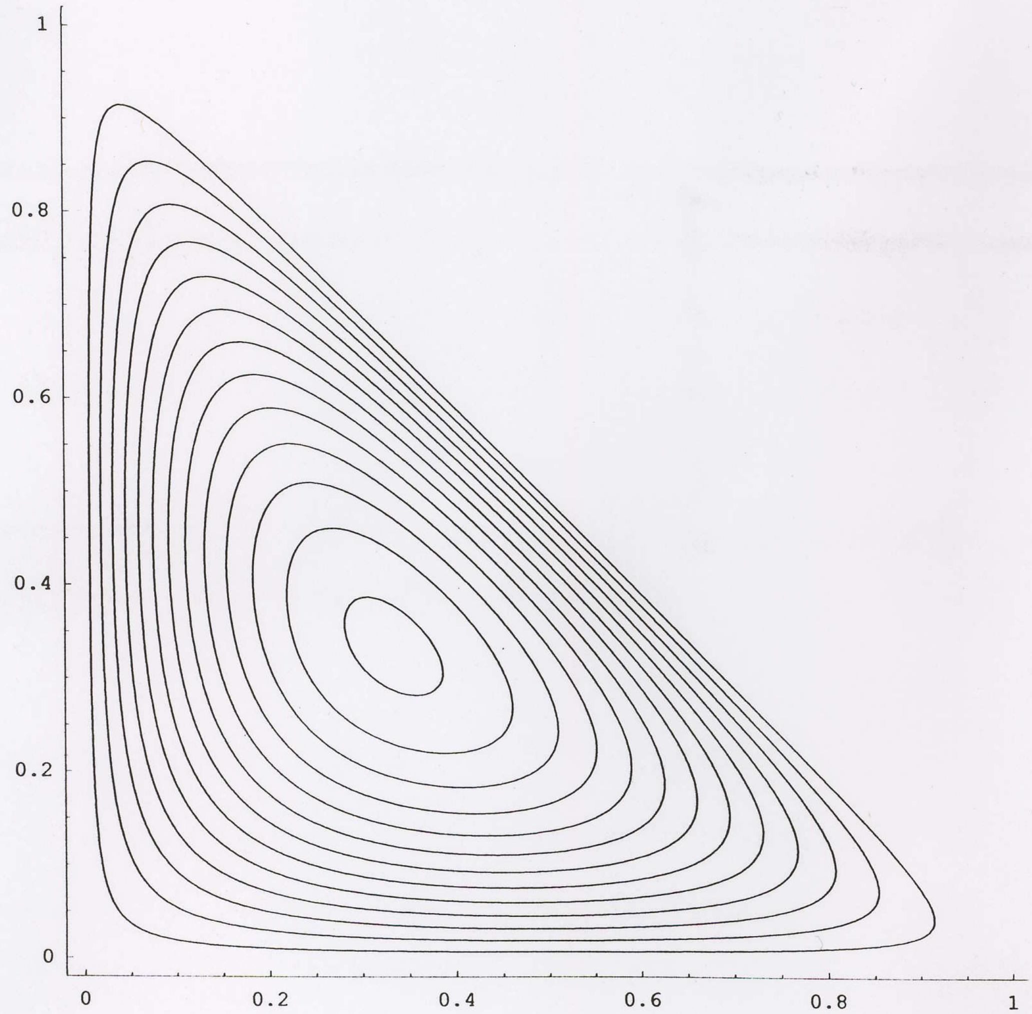

figure 1

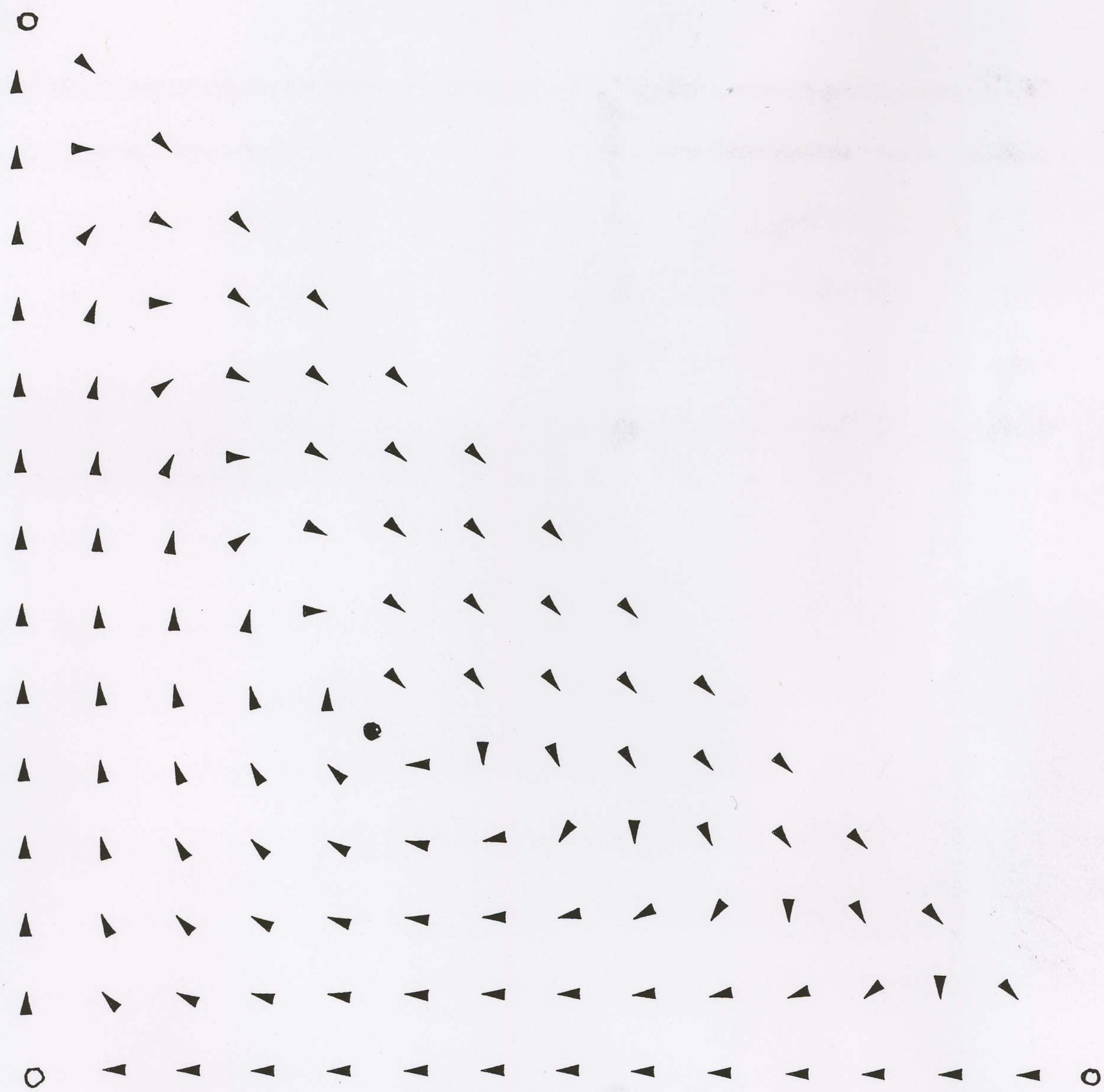

figure 2: $c = 0$

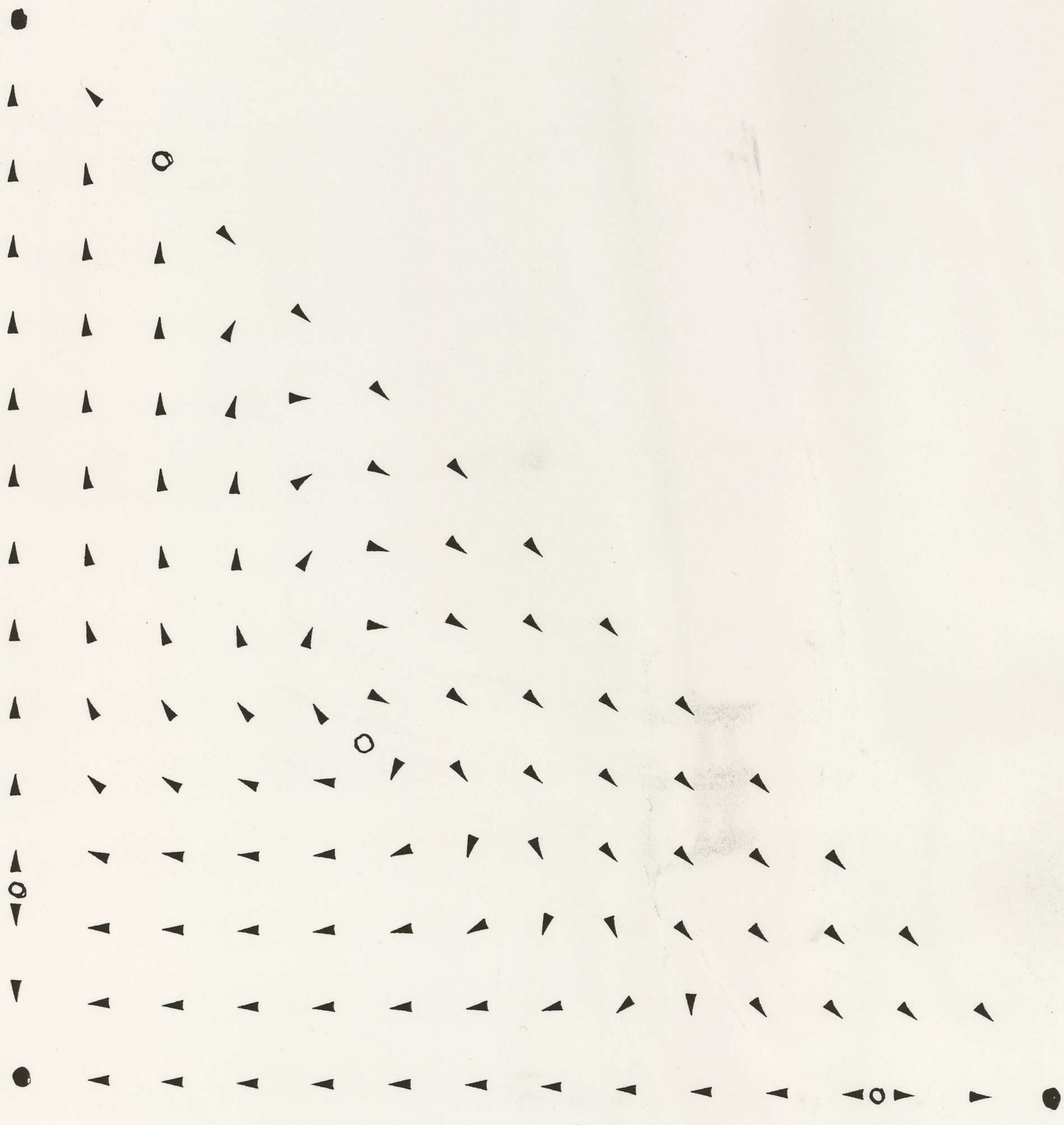

Figure 3: $c = .6$