

Opinion formation dynamics with contrarians and zealots

Kaitlyn Eekhoff *

Department of Mathematics and Statistics
Calvin College
Grand Rapids, MI 49546

Faculty advisor: Todd Kapitula †

Department of Mathematics and Statistics
Calvin College
Grand Rapids, MI 49546

January 14, 2019

Abstract. Mean-field type ODE models for opinion dynamics often assume that the entire population is comprised of congregators, who are agreeable. On the other hand, a contrarian opinion dynamics ODE model assumes the population has two personality types: congregators, and contrarians, who are disagreeable. In this paper we broadly study how contrarians influence the ability of the population to form a fixed and stable opinion. In particular, we re-examine the dynamics associated with the model introduced by Tanabe and Masuda [12] by looking at how the parameters effect the formation of stable periodic solutions (whose existence implies there is no fixed consensus opinion). Afterwards, we refine and analyze the model under two new hypotheses: (a) the contrarians bow to peer pressure and change their personality type to congregators if a large enough proportion of the entire population agrees on an opinion, and (b) there are zealots associated with one of the opinions. We conclude with a brief discussion on possible extensions of this work.

Keywords. Opinion formation, nonlinear ODEs, bifurcation analysis

CONTENTS

1. Introduction	18
2. Original contrarian model: parameter study	21
3. Modified contrarian model: change of personality type	23
4. Modified contrarian model: zealots for opinion 1	26
5. Conclusion	28

*E-mail: kmeekhoff@gmail.com

†E-mail: tmk5@calvin.edu

1. INTRODUCTION

People have different opinions on all kinds of issues, and when communicating with each other allow for their opinions to be changed or reinforced. Such phenomena have been studied by many mathematicians under the heading of opinion dynamics with the use of a variety of different models, including the voting (or election) model, as demonstrated in Yildiz et al. [14], Mellor et al. [8], Wang et al [13], and Bernardes et al [3], and the naming game model as demonstrated in Mistry et al [9]. More specifically, Yildiz et al. base their research off of the traditional voter model, which has a binary setup in which each individual holds one of two possible opinions. Upon communication with a neighbor, the individual changes opinions to that of the neighbor's opinion. In addition to this traditional model, Yildiz et al. adds stubborn agents (zealots), individuals who influence others but do not change their own opinions. In this sense, they have a fixed state, an effect which does not allow for a convergence to consensus. This work highly influenced the beginning of our research, leading to questions such as:

- (a) what would the addition of stubborn agents do in a non-binary system, and
- (b) are there other ways to model stubbornness or resistance?

The initial questions prompted by Yildiz et al. lead us into more studies on opinion modeling and stubborn agents. Some of the models found looked at the behavior of individual agents, such as in Ben-Naim et al. [1], whereas other models studied proportions of populations, such as in Marvel et al. [7]. Specifically inspirational was the work of Marvel et al. in their use of a simple, proportionally based model that could be used and added onto in so many ways to look at the concern of ideological conflict and radicalization. In a similar way to Yildiz et al., the model of Marvel et al. was foundational to the way that we approached modeling opinion dynamics. Marvel et al. offer a basic set of differential equations that act as stepping stones to examining intricate and complicated questions. While Marvel et al. were interested in finding out what factors were involved in deradicalizing, or moderating, two extreme opinions given the current climate of polarized opinions, we were interested in how resistance, an agent's willingness, or lack thereof, to change opinions, played a role in moderating opinions, given the appearance of not only extreme opinions but also dedicated and passionate opinion holders in the current political climate.

Other opinion dynamic models can operate both in discrete, see Stauffer and Sahimi [10] and continuous time, see Marvel et al. [7]. While we decided to use continuous time to work closer to Marvel's model, other models in discrete time were specifically interesting and offered areas of further research. Stauffer studies very similar topics to Marvel and Yildiz, which makes their work of great interest to us. However, what makes their work unique is their discrete model of opinion dynamics. While a continuous model has its benefits, a discrete model also has interesting impacts on the dynamics. Stauffer shows how a discrete model can capture the importance of spatial fluctuations in the different opinion groups. Stauffer's work leads us to question the next steps of our own research and modeling.

Specific areas of application and study for such models include riots, see Berestycki et al. [2], radicalization and de-radicalization, see Marvel et al. [7] and Chuang et al. [5], the influence of fanatics or zealots, see Stauffer and Sahimi [11], the art of compromises, see Ben-Naim et al. [1], and election outcomes, see Bernardes et al [3]. Like Marvel et al., Berestycki et al. look at a form of radicalization, but works on a more detailed level to explore riots specifically. With the goal of understanding what internal and external factors influence and play a role in riots, Berestycki et al. explore the effect of statistical norms in crime patterns as to predict or anticipate the patterns seen in riots. In addition to their work with discrete modeling, Stauffer et al. also deeply explore the influence of committed activists, as do Marvel et al. While a basic model has an agent changing opinions based on any interaction, a more complicated model includes agents who will only change their opinion after much convincing or may not change their opinion under any circumstances. This more complicated model mimics the very real extremists and dedicated believers that are a staple in modern society. Stauffer explores how these stubborn agents change other agents' opinions as well as how the threshold for convergence changes in correlation to the committed activists. With an emphasis on convergence and compromise, Ben-Naim et al. explore how only like-minded agents communicate with each other. This model captures another important phenomenon where not all agents are equally likely to communicate with each other. This was interesting because it would seem that stubborn agents, our particular interest, might not communicate with the same agents that a non-stubborn agent would communicate. Ben-Naim et al.'s

use of pairwise interactions inspired us to consider modeling interaction rates in different ways depending on the stubbornness of the agent. Throughout this paper we assume that an opinion on a particular issue is binary, and label it as 0 or 1 (not 0). Herein we focus on using the so-called mean-field model to study opinion formation in a given (large) group. In these models, the interest is not on the opinion of an individual within a population; instead, one looks at how proportions of the entire population which support 0 or 1 change over time. An underlying assumption for the validity of these mean-field equations is that there is a large number of individuals in the population, and that the interactions between individuals happens on a relatively fast time scale. Epitomized in Marvel et al. [7], the mean-field model is a deterministic and continuous model that can be used, e.g., to study the affect of zealots (those who will not change their opinion), and congregators (those who are agreeable, will change their opinion upon the influence of another opinion, and want to share the opinion of their neighbors). This model is deterministic because, with the knowledge of initial conditions, one can describe exactly what happens. For many mean-field models the final result is that one opinion eventually becomes the majority opinion. Such a result leads to the initial questions that sparked the current research. Can we produce a model such that we cannot know what opinion will become the majority opinion? Can we produce a model such that one opinion never wins and there is no consensus?

When modeling opinion formation, one must make an assumption about the personality types of the individuals within the entire population. It is often assumed that people are congregators. However, for large populations this assumption is suspect, as it will almost certainly be the case that within the entire population there will be some people who are contrarian, i.e., people who are disagreeable and wish to argue with their neighbors and have a differing opinion. The influence of stubbornness, or resilience, in opinion formation led to the contrarian model of Tanabe and Masuda [12]. Tanabe et al. assume the population is broken into two groups: congregators and contrarians. In other words, there are those who are agreeable and those who are not. With a non-linear opinion model, Tanabe et al. look at how the introduction of contrarians affects convergence to a dominant opinion. This work was specifically interesting because it builds off of Marvel et al.'s basic mean-field model but also adds the fascinating aspect of contrarians. Knowing that we wanted to look at the influence of stubbornness and resistance in opinion modeling, Tanabe et al.'s contrarians opened the door to a model that was both familiar and had the potential to be extended to model what was of interest to us. Like these authors, in this paper we assume that both types of people are present in the population. We will let $0 \leq X \leq 1$ represent the proportion of the total population who are congregators, and $Y = 1 - X$ represent the proportion that are contrarian. As we will see, the presence of contrarians, $Y > 0$, has a significant influence in forming consensus.

For our mathematical model we start with the contrarian opinion ODE model as given by Tanabe and Masuda [12],

$$\begin{aligned}\dot{x} &= r [(X - x)(x + y)^d - x(1 - x - y)^d] \\ \dot{y} &= (Y - y)(1 - x - y)^d - y(x + y)^d.\end{aligned}\tag{1.1}$$

The value of $x \in [0, X]$ corresponds to the proportion of the congregators who hold opinion 1 ($X - x$ of the congregators hold opinion 0), while the value of $y \in [0, Y]$ is the proportion of contrarians who hold opinion 1 ($Y - y$ of the contrarians hold opinion 0). The parameter r distinguishes between the reaction rates for congregators and contrarians, meaning that congregators do not change their opinions at the same rate that contrarians do. The differences in the agent types allows for different types of interactions at different rates; in particular, if $r > 1$, then congregators will change their opinion at a faster rate than contrarians. Noting that $x + y$ is the total proportion of those who hold opinion 1, and $1 - x - y$ is the total proportion who hold opinion 0, the parameter d can be interpreted as representing the number of interactions with like-minded people that must occur before an opinion is changed [12].

This particular model falls under the general class of compartment models with nonlinear reaction rates, see Figure 1. While there are four compartments, there needs to be only two equations to completely describe the dynamics. This is due to the fact that the opinion is assumed to be binary, e.g., knowing the proportion of those who hold opinion 1 in the congregators (x) automatically implies knowing the proportion who hold opinion 0 in the congregators ($X - x$).

In addition to the graphical representation of Figure 1, a critical point analysis allows us to mathematically distinguish between the two personality types. Suppose $X = 1$, i.e., the entire population is of congregator

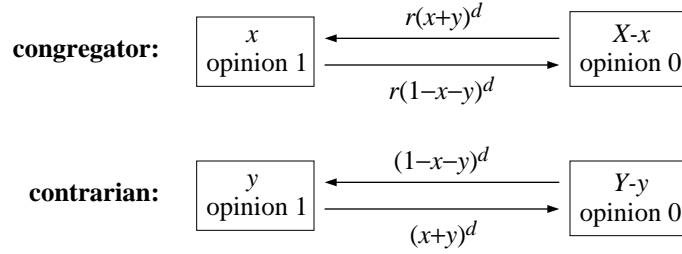


Figure 1: (color online) A graphical representation of the compartment model associated with the model (1.1). The reaction rates between the compartments are given just above or below the arrows pointing from one compartment to another. There is no direct interaction between the congregators and contrarians due to the assumption that personality type is invariant.

type. Equation (1.1) collapses to the scalar equation,

$$\dot{x} = rx(1-x) [x^{d-1} - (1-x)^{d-1}].$$

We need to assume $d > 1$ in order to have nontrivial dynamics and to ensure there are no singularities in the vector field; henceforth, this will be assumed. The three critical points are $x = 0, 1/2, 1$. The middle critical point is unstable, while the other two are stable. In the absence of contrarians, the congregators want to agree on either opinion 0 or opinion 1. Now suppose $X = 0$, i.e., the entire population is of contrarian type. The system collapses to the scalar equation,

$$\dot{y} = (1-y)^{d+1} - y^{d+1}.$$

The only critical point is $y = 1/2$, and it is stable. In the absence of congregators, the contrarians will have an equally split opinion; in particular, there is no consensus.

Now consider the full dynamics. It is easy to see that for any value of X there is the *egalitarian solution*, $(x, y) = (X/2, Y/2)$ (half the population holds opinion 0, while the other half holds opinion 1). When considering the full dynamics, we will see that, for fixed r and d , there is a range of X , say $X_- < X < X_+$, for which there is a stable periodic solution. While the periodic solution is not necessarily a global attractor, it is the case that if the initial condition is close to the egalitarian solution, then it is the attractor. For $X \notin [X_-, X_+]$, there are only critical points for the system, at least one of which is an attractor. As we will see, if $X < X_-$, the egalitarian solution is a global attractor. On the other hand, if $X > X_+$, there are two stable critical points. If the initial condition is close to the egalitarian solution, which is now a saddle point, then the asymptotic state depends on which side of the stable manifold is the initial condition. One stable state corresponds to a majority of the population holding opinion 0, while the other corresponds to a majority of the population holding opinion 1. When the periodic solution exists, the majority oscillates between the two opinions, and a stable majority opinion is not achieved.

The purpose of this paper is two-fold. First, we will examine the dependence of X_{\pm} on the parameters r and d . The determination of X_+ is primarily a numerical calculation using the MATLAB program Matcont (see [6]). We will then extend the model in two directions and examine the effect of such extensions. To begin, we will assume that there is a threshold associated with contrarian behavior. In particular, for a given $1/2 < A < 1$, we will assume that if $x + y > A$ (the proportion holding opinion 1 is greater than A) or $x + y < 1 - A$ (the proportion holding opinion 0 is greater than A), then contrarians become congregators. After that, we will assume there are zealots for opinion 1. As seen in the case of all congregators (see [4] and the references therein), it is known that there is a threshold such that if the proportion of zealots is above that threshold value, then the entire population will hold the same opinion as the zealots. We examine how the presence of contrarians effects this result, and to determine the manner in which they effect the size of the interval (X_-, X_+) . We conclude with some possible extensions for this work.

2. ORIGINAL CONTRARIAN MODEL: PARAMETER STUDY

The goal here is to understand the dynamics associated with the contrarian model (1.1). An analysis of this system was presented in [12, Section III]. Here we extend that analysis in order to better understand solution behavior as a function of the parameters; in particular, we look to see how the existence of periodic solutions depends on X, d, r . In all that follows, we assume $d > 1$.

We start by doing an analysis around the *egalitarian* critical point,

$$x_e = \frac{1}{2}X, \quad y_e = \frac{1}{2}Y \quad \rightsquigarrow \quad x_e + y_e = \frac{1}{2}.$$

The critical point describes the situation for which the opinion in the total population is equally split between opinion 0 and opinion 1. Based upon the analysis presented in Section 1, the expectation is that if X is sufficiently large, this point will be unstable; whereas, if X is sufficiently small, it will be stable. Consequently, it is possible for there to be interesting and nontrivial dynamics in the transition between having a small proportion of congregators in the total population and a large proportion of congregators.

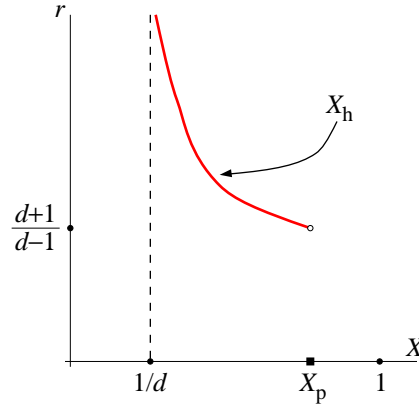


Figure 2: (color online) A cartoon plots of the Hopf bifurcation curve for a fixed $d > 1$ in the Xr -plane. The square represents the pitchfork bifurcation point, $X_p = (d + 1)/(2d)$. The bifurcation occurs for parameter values on the solid (red) curve.

In order to do the analysis we first linearize about the critical point. From [12, equations (7)-(8)], the relevant quantities associated with the linearization about this critical point are the trace, τ , and determinant, Δ ,

$$\tau = (1 + r)(dX - 1) - d, \quad \Delta = r[1 - d(2X - 1)].$$

The egalitarian solution is a saddle point if $\Delta < 0$. There is a bifurcation point which corresponds to a pitchfork bifurcation if $\Delta = 0$,

$$\Delta = 0 \quad \rightsquigarrow \quad X_p = \frac{1}{2} + \frac{1}{2d}.$$

The reason that the bifurcation is of pitchfork-type is that the equations are invariant under the transformation,

$$x \mapsto X - x, \quad y \mapsto Y - y. \quad (2.1)$$

Indeed, if (x^*, y^*) is a critical point, then so is $(X - x^*, Y - y^*)$. There is a Hopf bifurcation if $\Delta > 0$ and $\tau = 0$. The parameter values at which these bifurcations take place are,

$$\tau = 0 \quad \rightsquigarrow \quad X_h = \frac{1}{r + 1} + \frac{1}{d}.$$

Note that,

$$\Delta > 0 \quad \rightsquigarrow \quad X < X_p,$$

so a Hopf bifurcation must occur for a smaller value of X than does the pitchfork bifurcation. The Hopf bifurcation requires that the rate associated with the congregators is sufficiently large, $r > (d+1)/(d-1)$, and that the total proportion of congregators is in a certain range that depends upon d , $1/d < X < X_p$. Regarding the Hopf and pitchfork bifurcation points in (X, r) -space for fixed d , we have the cartoon as presented in Figure 2. This diagram is not given in [12].

Remark 2.1. If any type of bifurcation occurs at another critical point $(x^*, y^*) \neq (x_e, y_e)$, then the invariance (2.1) implies that the same bifurcation occurs at $(X - x^*, Y - y^*)$.

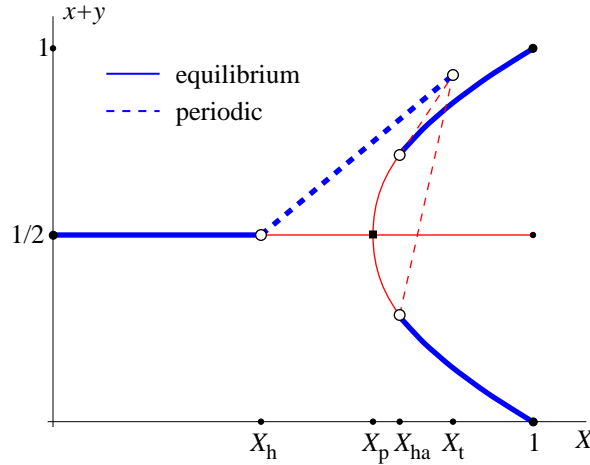


Figure 3: (color online) A cartoon plot of the bifurcation diagram for a fixed $d > 1$ and $r > (d+1)/(d-1)$ in the $X(x+y)$ -plane. The solid curves represent equilibrium solutions, and the dashed curves represent periodic solutions, each of which arises from a Hopf bifurcation. The solution is stable if it is represented by a thick (blue) curve, and unstable if it is given as a thin (red) curve.

Now fix $d > 1$ and $r > (d+1)/(d-1)$. A cartoon plot of the full bifurcation diagram is given in Figure 3. The solid curves represent equilibrium solutions, and the dashed curves represent periodic solutions. The solution is stable if it is represented by a thick curve and unstable if it is given as a thin curve. The point X_h is the Hopf bifurcation point from the egalitarian solution. The resulting periodic solution is stable. There is a pitchfork bifurcation for the egalitarian solution at X_p . At the point X_{ha} , there is another Hopf bifurcation from the equilibrium solutions arising from the pitchfork bifurcation. The resulting periodic solutions are unstable. Finally, the periodic solutions terminate at X_t . The termination is a consequence of the unstable interior periodic solutions colliding with the stable periodic solution as $X \rightarrow X_t^-$. A cartoon of the phase portrait in xy -space for $X_{ha} < X < X_t$ is given in Figure 4.

We now present plots for $X_- = X_h$ and $X_+ = X_t$ as a function of the parameters r and d . If $X_- < X < X_+$, then for initial data close to the egalitarian solution there is not an asymptotically stable dominant opinion; instead, the solution will approach the periodic solution, on which $x(t) + y(t)$ periodically crosses the line $x + y = 1/2$. If $X < X_-$, the egalitarian solution is globally stable, and if $X > X_+$, one of the two opinions will be dominant, with the final state depending upon the initial data and its location relative to the stable manifold of the unstable egalitarian solution. While the curve for X_- has been determined analytically, the curve for X_+ must be determined numerically. We use the MATLAB program Matcont to do so. In particular, for a fixed value of r and d , the point $X_+ = X_+(r, d)$ corresponds to that value for which a stable periodic solution collides with an unstable periodic solution, i.e., a saddle node bifurcation of periodic orbits. Matcont can numerically track these points.

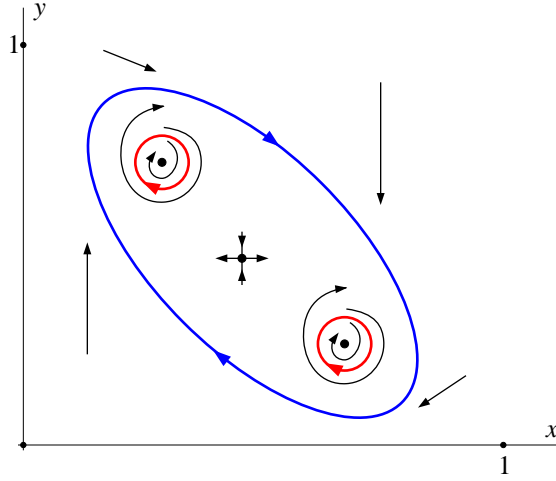


Figure 4: (color online) A cartoon phase portrait for $X_{\text{ha}} < X < X_{\text{t}}$ is given. There are two unstable periodic solutions, and one stable periodic solution. The egalitarian solution is a saddle point.

The results are plotted in Figure 5. The left panel plots the curves for $d = 2$ with r varying, while the right panel has these curves for $r = 4.5$ with d varying. Qualitatively, the results presented herein do not depend on the particular value of fixed r and d . The cusp point, where the two curves coincided, corresponds to a Takens-Bogdanov bifurcation ($\tau = \Delta = 0$) for the the egalitarian solution $((r, X) = (3, 3/4)$ in the left panel, and $(d, X) = (11/7, 9/11)$ in the right panel). Note that as r and d increase (with the other parameter fixed) the length of the interval (X_-, X_+) increases. Regarding the left panel, $X_+(r)$ is increasing and $X_-(r)$ is decreasing. The plots indicated that as the reaction rate difference between the two personality types increases, there is a larger range for proportions of contrarians in which there is no stable and fixed opinion. Regarding the right panel, both $X_{\pm}(d)$ are decreasing. As is the case for the left panel, increasing d increases the length of the interval.

3. MODIFIED CONTRARIAN MODEL: CHANGE OF PERSONALITY TYPE

We now assume that contrarians will succumb to social pressure and become congregators if a given threshold is reached. We will denote the threshold level by $A \in (1/2, 1]$. The contrarians will change their personality type if the proportion of 0's or 1's is too high,

$$x + y > A, \text{ or } (X - x) + (Y - y) > A \rightsquigarrow x + y < 1 - A.$$

Alternatively, the personality-type is unchanged if $1 - A < x + y < A$. The new model is a modification of (1.1),

$$\begin{aligned} \dot{x} &= r [(X - x)(x + y)^d - x(1 - x - y)^d] \\ \dot{y} &= (Y - y)g(1 - x - y) - yg(x + y), \end{aligned} \quad (3.1)$$

where

$$g(z) = \begin{cases} (1 - z)^d, & 0 \leq z < 1 - A \\ z^d, & 1 - A \leq z < A \\ (1 - z)^d, & A \leq z \leq 1. \end{cases}$$

Regarding this model,

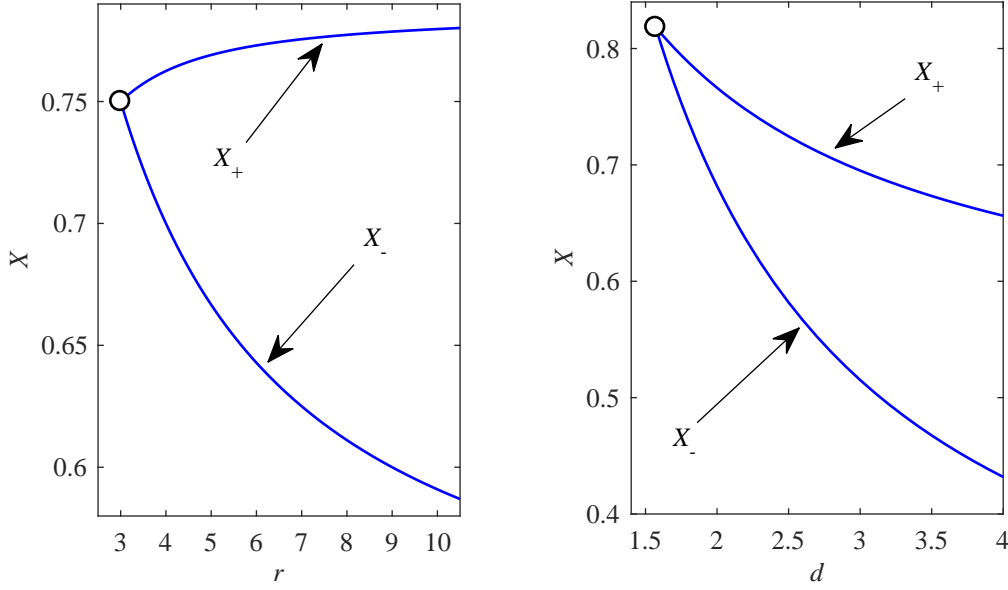


Figure 5: (color online) Plots of the curves X_{\pm} . The curve X_{-} is known analytically, whereas the curve X_{+} is found numerically using Matcont. The left panel plots $X_{\pm}(r)$ with $d = 2$, and the right panel plots $X_{\pm}(d)$ with $r = 4.5$. The point where the two curves meet corresponds to a Takens-Bogdanov bifurcation point.

- (a) the egalitarian solution exists, and has the same stability properties as for the original contrarian model
- (b) the equations continue to remain invariant under the transformation, $(x, y) \mapsto (X - x, Y - y)$, so for bifurcations and phase-plane analysis, there remains the symmetry with respect to the egalitarian solution.

On the other hand, assuming $A < 1$, the model also now allows for critical points associated with the congregator-only model,

$$(x, y) = (X, Y), \quad (x, y) = (0, 0),$$

which correspond to everyone agreeing on the same opinion. These solutions are stable for the ODE. Regarding the solution behavior, for $1 - A < x + y < A$, the model is the original contrarian model. The solution structure will remain the same, as will the bifurcations. For $x + y \notin (1 - A, A)$, the contrarians become congregators. The only fixed points are $(x, y) = (X, Y)$ and $(x, y) = (0, 0)$, and each of these are stable nodes.

When doing analysis and numerical simulations, we smooth the piecewise continuous function $g(z)$ and replace it with,

$$g(z) = (1 - z)^d + \frac{1}{2} [z^d - (1 - z)^d] \left[\tanh \left(\frac{z - (1 - A)}{\epsilon} \right) - \tanh \left(\frac{z - A}{\epsilon} \right) \right],$$

where $0 < \epsilon \ll 1$. We are now assuming that the change in personality type is not instantaneous, but instead, it occurs rapidly near the critical thresholds $x + y \in \{1 - A, A\}$. Upon making the reaction terms smooth, we introduce two new curves of critical points. While we will not provide the graphical analysis here, to leading order in a perturbation expansion these points lie on the transition curves, $x + y = 1 - A$ and $x + y = A$. Depending upon the value of A , we have the bifurcation picture as provided in [Figure 6](#). This diagram is found by taking the bifurcation diagram of [Figure 3](#), adding the lines $x + y = A$ and $x + y = 1 - A$ to it, and then removing those portions of the original curves for which $x + y < 1 - A$ and $x + y > A$.

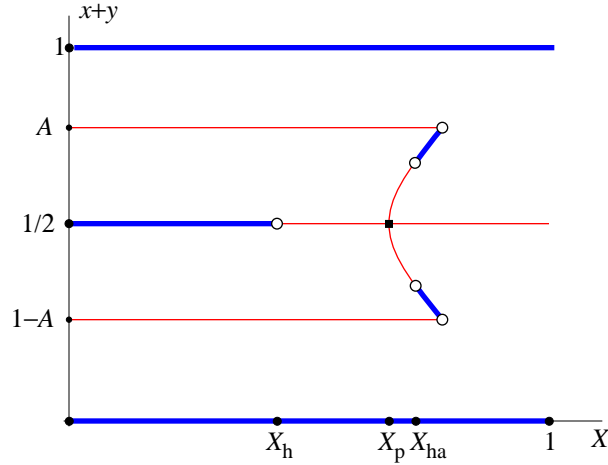


Figure 6: (color online) A cartoon plot of the bifurcation diagram for a fixed $d > 1$ and $r > (d+1)/(d-1)$ in the Xx -plane. The solid curves represent equilibrium solutions. The solution is stable if it is represented by a thick (blue) curve, and unstable if it is given as a thin (red) curve. There is a saddle-node bifurcation at the transition lines, $x + y = 1 - A$ and $x + y = A$. Hopf bifurcations take place at $X = X_h$ and $X = X_{ha}$.

Regarding the dynamics, we can construct a trapping region for the two critical points which represent total agreement on opinion, $x + y \in \{0, 1\}$. First, the curve $\dot{y} = 0$ intersects the line $x + y = A$ at the point (x_+, y_+) with,

$$x_+ = A - y_+, \quad y_+ = \frac{A^2 Y}{A^2 + (1 - A)^2},$$

and it intersects the line $x + y = 1 - A$ at the point (x_-, y_-) with,

$$x_- = A - y_-, \quad y_- = \frac{(1 - A)^2 Y}{A^2 + (1 - A)^2}.$$

Now, in the regions $x + y \notin (1 - A, A)$,

$$\begin{aligned} \dot{x} + \dot{y} &= \dot{x} + r\dot{y} + (1 - r)\dot{y} \\ &= 2r(x + y)(1 - x - y) \left(x + y - \frac{1}{2} \right) + (1 - r)\dot{y}. \end{aligned}$$

Consequently, if $x + y = A$ and $x < x_+$, then $\dot{x} + \dot{y} > 0$, while if $x + y = 1 - A$ and $x > x_-$, then $\dot{x} + \dot{y} < 0$. We also have $\dot{x} + r\dot{y} > 0$ on the line $x + ry = A + (r - 1)y_+$ (which intersects $x + y = A$ at (x_+, y_+) , and $\dot{x} + r\dot{y} < 0$ on the line $x + ry = (1 - A) + (r - 1)y_-$ (which intersects $x + y = A$ at (x_+, y_+)). The convex hull of the lines $x + y = A$ and $x + ry = A + (r - 1)y_+$ then forms a boundary for a trapping region for (X, Y) , whereas the convex hull of the lines $x + y = 1 - A$ and $x + ry = (1 - A) + (r - 1)y_-$ forms a boundary for a trapping region for $(0, 0)$. Any dynamics associated with the contrarian personality must then be contained in the complementary domain. Note that the saddle points are on the boundary of the convex hulls. The true boundary is given through the stable and unstable manifolds of these saddle points.

We now fix $r = 4.5$ and $d = 2$ and look to determine the curves $X_{\pm}(A)$. As was done previously, $X_-(A)$ can be determined analytically, whereas $X_+(A)$ must be computed numerically using Matcont. From our preceding analysis, we know there is only one stable periodic solution, and it arises from a Hopf bifurcation when $X = 15/22 \sim 0.6818$. Since the original dynamics always take place near the line $x + y = 1/2$, the value of $X_-(A)$ is fixed and equal to the bifurcation value, $X_- = 15/22$. On the other hand, because the trapping regions for $(0, 0)$ and (X, Y) depend upon A , the value of X_+ will depend upon A . The upper

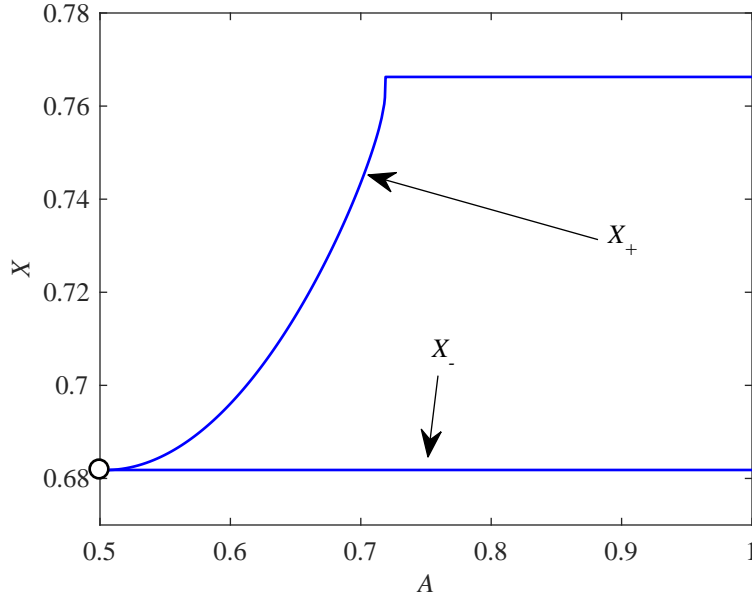


Figure 7: (color online) A plot of $X_{\pm}(A)$ when $r = 4.5$ and $d = 2$.

and lower bounds are illustrated in [Figure 7](#). For $A > A_c \sim 0.7190$, the upper limit is that associated with the original model, $X_+ \sim 0.7663$. Note that as $A \rightarrow 0.5^+$, i.e., the threshold at which the personality type changes converges to a simple majority, the size of the interval (X_-, X_+) decreases to zero.

4. MODIFIED CONTRARIAN MODEL: ZEALOTS FOR OPINION 1

A zealot is a person whose opinion is fixed; in particular, it does not change upon interaction with an individual who holds the opposing opinion. We let $0 < p < 1$ denote the total proportion of zealots who hold opinion 1, and we assume that they are equally distributed between the congregators and contrarians. In particular, there are pX zealots who are congregators and pY zealots who are contrarians. This assumption requires the bounds on x, y ,

$$0 \leq x \leq (1-p)X, \quad 0 \leq y \leq (1-p)Y \quad \rightsquigarrow \quad 0 \leq x + y \leq 1 - p.$$

The proportion of congregators who hold opinion 0 is $(1-p)X - x$, and the proportion of contrarians who hold opinion 0 is $(1-p)Y - y$. The contrarian model which allows for the presence of zealots is,

$$\begin{aligned} \dot{x} &= r \left[((1-p)X - x)(x + y + p)^d - x(1-p-x-y)^d \right] \\ \dot{y} &= [(1-p)Y - y](1-p-x-y)^d - y(x + y + p)^d. \end{aligned} \tag{4.1}$$

In all that follows we assume $d = 2$, which does not change the qualitative nature of the results.

In order to understand the effect of zealots on congregators, we first assume $X = 1$ (all congregators), so the system [\(4.1\)](#) collapses to a scalar equation,

$$\dot{x} = r \left[(1-p-x)(x+p)^2 - x(1-p-x)^2 \right].$$

The critical point $x = 1 - p$ is stable and represents the situation where the entire population agrees on opinion 1. The other two (potential) critical points are,

$$x_{\pm} = \frac{1}{4} \left(1 - 3p \pm \sqrt{1 - 6p + p^2} \right), \quad p \leq p^* = 3 - \sqrt{8} \sim 0.1716.$$

The point x_- is stable and represents the situation where the majority of the population holds opinion 0. The point x_+ is unstable. There is a saddle-node bifurcation at $p = p^*$. If $p > p^*$, then the only critical point is $x = 1 - p$, so in this case, the entire population will agree with the zealots no matter the initial proportion who hold opinion 0. Otherwise, it is possible for either opinion to be the majority opinion.

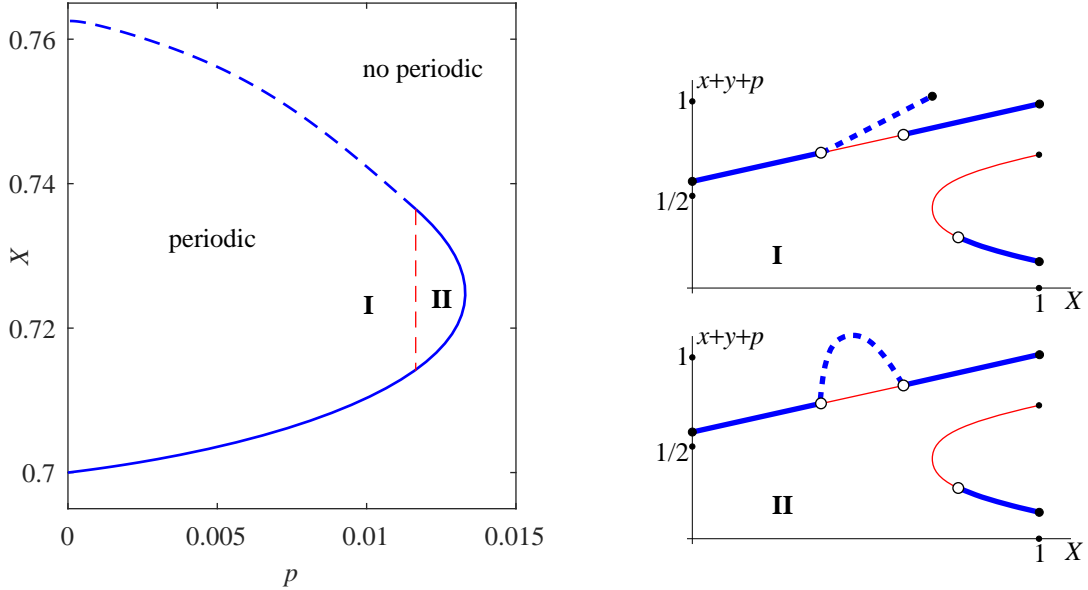


Figure 8: (color online) A plot of the region for which there exist stable periodic solutions when $r = 4$. The left figure contains the boundary for which there exist periodic solutions. The solid curve on the boundary corresponds to a Hopf bifurcation, and the dashed curve corresponds to the value for which a stable periodic solution ceases to exist. The right figures provide a cartoon of the bifurcation diagram in each of the two regions. The upper right figure corresponds to region I, and the lower right figure corresponds to region II. The thick solid curves correspond to stable critical points, and the thin solid curves are unstable critical points. The open circles correspond to Hopf bifurcation points. The thick dashed curves correspond to the periodic solution arising from the Hopf bifurcation. Note that in region II the periodic solutions connect the two Hopf bifurcation points.

We now consider the full problem, $X < 1$. The bifurcation diagram associated with $p = 0$ is that in Figure 3. First, regarding the pitchfork bifurcation at $X = X_p = 3/4$, the fact that p is a symmetry-breaking term in the equations implies that $p > 0$ and sufficiently small there will now be a saddle-node bifurcation. The saddle-node bifurcation point will limit to the pitchfork bifurcation point as $p \rightarrow 0^+$ (for a cartoon description see the right panel of Figure 8, and for a numerically generated picture see the right panel of Figure 9). Regarding the Hopf bifurcation points, they will vary smoothly as a function of (for at least small) p .

As for the curves X_{\pm} , they are depicted in the left panel of Figure 8 for $r = 4$ using Matcont (again, the results are qualitatively the same for any value of $r > 3$). Note that there is a saddle-node bifurcation for these curves at $p = p_{mp} \sim 0.0133$. While $X_-(p)$ always corresponds to a Hopf bifurcation, the curve $X_+(p)$ corresponds to either a collision of two periodic orbits (as in the $p = 0$ case), region I, or another Hopf bifurcation, region II. An important consequence is that a small proportion of zealots forces the entire population to form a majority opinion, i.e., for $p > p_{mp}$ the only stable solutions permitted by the model correspond to states at which the opinion is constant. Note also that even if $p < p_{mp}$, the range of X values for which there is a stable periodic solution is relatively small (the largest range, when $p = 0$, is $0.7 < X < 0.7625$).

Regarding the bifurcation diagram, henceforth assume $p > p_{mp}$, so there are no Hopf bifurcations. There

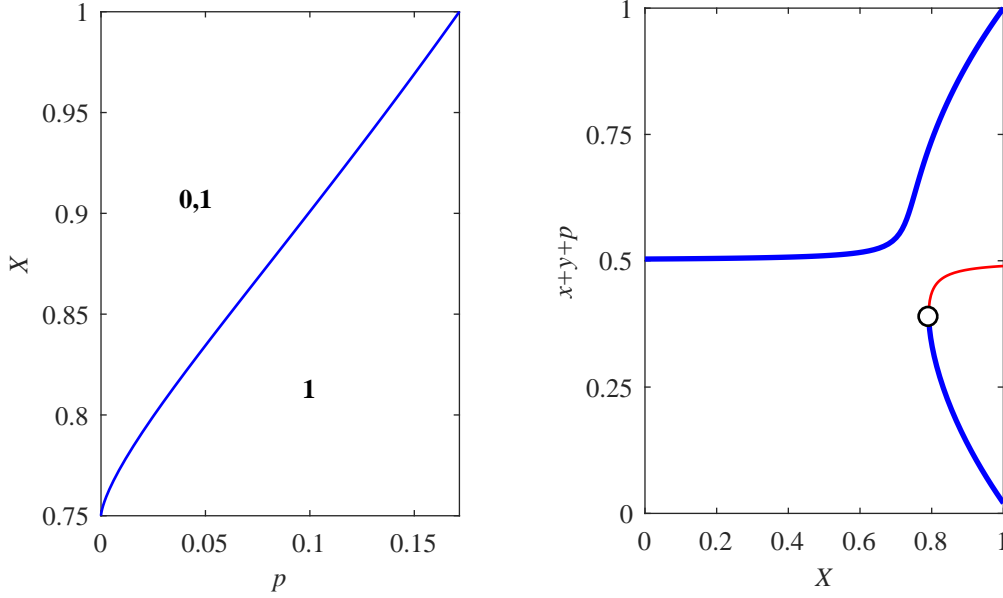


Figure 9: (color online) The left figure contains the transition curve in (p, X) -space between there being two possible opinions, or only one possible opinion. Above the curve both opinions are possible, and below the curve only opinion 1 is possible. The right figure gives the numerically generated bifurcation diagram for $p = 0.02 > p_{\text{mp}}$. The thick curves correspond to stable solutions, whereas the thin curves correspond to unstable solutions. Note the saddle node bifurcation, which occurs as X decreases across the transition curve.

are now only critical points, each of which corresponds to there being a majority opinion. We have three distinct regions in (p, X) -space (see the left panel of Figure 9). If $p > p^* \sim 0.1716$, then the only stable critical point, for any value of X , corresponds to a state where opinion 1 is the dominant opinion (see the right panel in Figure 9, and imagine that the lower curve has been removed). Now suppose $p < p_c$. There is a transition curve between either opinion being possibly dominant, versus only opinion 1 being dominant (see the left panel in Figure 9). If X is above the transition curve, i.e., the proportion of contrarians is not too large, it is possible for the dominant opinion to be either 0 or 1. If there are too many contrarians, i.e., X is below the curve, then only opinion 1 is the final end state. Finally, note that the transition curve is decreasing. Consequently, the more contrarians there are in the population, the smaller the proportion of zealots that is needed in order for the zealot opinion to prevail as the majority opinion.

5. CONCLUSION

In this paper, we analyzed a simple opinion dynamics model first introduced in [12], which allows for both congregators and contrarians. We did a parameter study associated with the stable periodic solution, which had not been done previously. We also looked at two modifications of the model. The first allowed for the possibility of change of personality; in particular, we assumed that contrarians would become congregators if the proportion of people holding an opinion reached a certain threshold (analogously, the contrarians bow to peer pressure at some point). The primary finding is that the range of contrarian proportions for which there can be periodic behavior shrinks as the threshold decreases. Regarding application, this finding suggests that the convergence of beliefs or opinions is more likely when the contrarian proportions cannot produce periodic behavior. This means that the contrarians, those who want to be disagreeable, have the ability to prevent convergence, but certain thresholds can also put boundaries on this ability of the contrarians.

The second modification assumes there is a proportion of the population who are zealots, i.e., people whose opinion remains unchanged after interactions. There we find that a small proportion of zealots can remove the possibility of there being periodic behavior. Moreover, if no periodic behavior is possible, the inclusion of zealots also guarantees that there will be a majority opinion, no matter the proportion of contrarians. The majority opinion will be the opinion of the zealots. Finally, we find that as the proportion of zealots increases, the proportion of contrarians needed in order needed to ensure that the majority opinion is also the zealot opinion decreases.

There are several possible future avenues of research. One is to modify the model of [Section 3](#) and allow only a subpopulation of the contrarians to change personality type. Another is to assume that there is a moderate subpopulation, as is the case in Marvel et al. [7] (where it is implicitly assumed that the entire population is of congregator-type). Moderates have neither opinion 0 or 1; in other words, they cannot make up their mind. When moderates are present there exists the possibility that there is not a majority opinion, even if there are no contrarians. Regarding the modeling, instead of two ODEs describing the opinion dynamics as in [Section 2](#), there will be a set of four nonlinear ODEs. Preliminary analysis indicates that the results associated with the more complicated model are qualitatively the same as the model discussed in [Section 2](#). Finally, we could explore several interacting population groups, each of which is called a city. Each city is assumed to have internal opinion dynamics as described by the original system (1.1), and then there would be some coupling between the cities. This problem of interacting cities, each of which is composed only of congregators who follow the Marvel et al. [7] model, was investigated in Bujalski et al. [4]. One of the interesting findings in that paper is that increasing interaction strength between individual cities leads to a clustering phenomena. Mathematically, the clustering was a consequence of a saddle-node bifurcation. The expectation is that adding contrarians to the model will increase the richness and complexity of the underlying dynamics, and perhaps lead to new and surprising conclusions.

REFERENCES

- [1] E. Ben-Naim, P. Krapivsky, and S. Redner. Bifurcations and patterns in compromise processes. *Physica D*, 183:2003.
- [2] H. Berestycki, J. Nadal, and N. Rodriguez. A model of riots dynamics: shocks, diffusion and thresholds. *Networks and Heterogeneous Media*, 10.3:2015.
- [3] A. Bernardes, D. Stauffer, and J. Kertesz. Election results and the Sznajd model on Barabasi Network. *Eur. Phys.*, 25:2002.
- [4] J. Bujalski, G. Dwyer, T. Kapitula, Q.-N. Le, H. Malvai, J. Rosental-Kay, and J. Ruiter. Consensus and clustering in opinion formation on networks. *Phil. Trans. R. Soc. A*, 376:20170186, 2018.
- [5] Y. Chuang, M. D’Orsogna, and T. Chou. A Bistable Belief Dynamics Model for Radicalization within Sectarian Conflict. *Quarterly of Applied Mathematics*, 75:2016.
- [6] A. Dhooge, W. Govaerts, and Y. Kuznetsov. Matcont: a MATLAB package for numerical bifurcation analysis of ODEs. *ACM TOMS*, 29:141–164, 2003.
- [7] S. Marvel, H. Hong, A. Papush, and S. Strogatz. Encouraging moderation: clues from a simple model of ideological conflict. *Phys. Rev. Lett.*, 109:118702, 2012.
- [8] A. Mellor, M. Mobilia, and R. Zia. Characterization of the Nonequilibrium Steady State of a Heterogeneous Nonlinear q-Voter Model with Zealotry. *Physics.soc-ph*:2016.
- [9] D. Mistry, Q. Zhang, N. Perra, and A. Baronchelli. Committed activists and the reshaping of the status-quo social consensus. *Physical Review*, 92:2015.
- [10] D. Stauffer and M. Sahimi. Discrete simulation of the dynamics of spread of extreme opinions in a society. *Physica A*, 364:90089, 2006.
- [11] D. Stauffer and M. Sahimi. Can a few fanatics influence the opinion of a large segment of society? *The European Physical Journal B*, 57:2007.

- [12] S. Tanabe and N. Masuda. Complex dynamics of a nonlinear voter model with contrarian agents. *Chaos*, 23:043136, 2013.
- [13] S. Wang, L. Rong, and J. Wu. Bistability and Multistability in Opinion Dynamics Models. *Applied Mathematics and Computation*, 289:2016.
- [14] E. Yildiz, A. Ozdaglar, D. Acemoglu, A. Saberi, and A. Scaglione. Binary Opinion Dynamics with Stubborn Agents. *ACM Transactions on Economics and Computations*, 1.4:2013.