

Opinion Dynamics with Slowly Evolving Zealot Populations

Ashlyn DeGroot *

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

Emma Schmidt †

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

Project advisor: Todd Kapitula ‡

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

Abstract. We introduce and analyze a model for opinion dynamics comprised of nonlinear ODEs. The variables are the proportion of moderates in the population who hold opinion A, the proportion of zealots who hold opinion A, and the proportion of zealots who hold opinion B (not A). The zealots are willing to change their opinion at a much slower rate than the moderates. Our model takes into account such things as the inherent attractiveness of one opinion over the other, the indoctrination of moderates by the zealots, and deradicalization of the zealots by the moderates. A combination of theoretical and numerical analysis shows there are many different types of asymptotic configurations of the population. Many of these correspond to critical points of the system. The most intriguing finding is that if both A and B are roughly equally attractive, and the rate of indoctrination is roughly equal to the rate of deradicalization, then there will be a stable periodic orbit. The dynamics of this orbit show that a precursor to an opinion being dominant is that the proportion of zealots for the opinion must first grow to some critical value. Moreover, when the periodic orbit exists, there are no other solutions which allow for coexistence between the two opinions.

Keywords. opinion dynamics, nonlinear ODEs, bifurcation analysis

*E-mail: degrootashlyn@gmail.com

†E-mail: schmidtem2020@gmail.com

‡E-mail: tmk5@calvin.edu

CONTENTS

1. Introduction	2
2. Zealots constant ($\epsilon = 0$)	3
2.1. $a \geq 0, b = 0$	4
2.2. $a \geq 0, b = 0.1$	6
2.3. $a = 0, b \geq 0$	8
2.4. $a = 0.1, b \geq 0$	9
3. Reduced dynamics for $\epsilon > 0$	10
3.1. $\dot{b} \approx 0$	12
3.2. $\dot{a} \approx 0$	14
4. Dynamics for the full system	15
4.1. Critical points	16
4.2. Jumping between the attracting surfaces	18
4.3. Full dynamics and periodic solutions	20
5. Conclusion	23

1. INTRODUCTION

We consider a closed population in which every member believes opinion A or opinion B (not A). We are interested in how the proportions change over time, and not the individual. Each individual is either a moderate or a zealot for the opinion. Moderates do not have a strong attachment to their opinion, and can be relatively easily convinced to change their mind. Zealots, on the other hand, hold their belief very strongly, and it is very difficult to get a zealot to change their mind. Thus, the total population is made up of four distinct subgroups:

- those who believe A moderately (labeled, x)
- those who are zealots for A (labeled, a)
- those who are zealots for B (labeled, b)
- those who believe B moderately (labeled, y).

Since the variables correspond to proportions, $a + x + b + y = 1$. Consequently, we can remove y as a variable, and instead describe moderate B's through the variable $1 - a - b - x$. We summarize using the below table:

a	proportion who are zealots for A
x	proportion who are moderate for A
b	proportion who are zealots for B
$1 - a - b - x$	proportion who are moderate for B

To mathematically differentiate between the two types of groups, we assume the beliefs of zealots change on a much slower time scale than those for moderates. This is illustrated in [Figure 1.1](#). We see there is slow movement (single arrows) between moderates and zealots of the same opinion, and fast movement (double arrows) between the moderates of the two opinions.

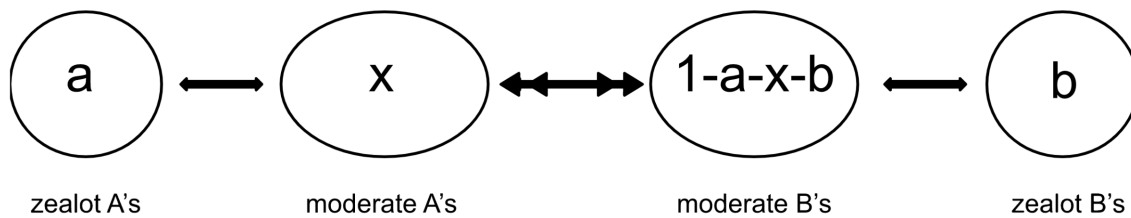


Figure 1.1: Movement between groups. The arrows between the groups represent the ways members can move between groups. The single arrows represent a slow rate of change while the double arrows represent a fast rate of change.

In the context of opinion formation, this listing of subgroups is perhaps too simplistic, as it assumes there are no undecided people. However, in the context of, e.g., religious affiliation, it is more natural, as A refers to those who have some type of religious affiliation, and B refers to those who do not [2].

The model implicitly assumes everyone in the population has the personality type of being agreeable, i.e., they all wish to hold the same opinion if possible. Eekhoff [5] (also see the references therein) looked at the problem when the population also included contrarians, i.e., those who are disposed to have a different opinion simply because they are disagreeable.

We will explore the ways in which these population subgroups interact with each other, and how the proportions within each group changes as a consequence of the interactions.

The model that we will use to represent this movement will have the general form,

$$\begin{aligned}\dot{x} &= f(x, a, b) + \epsilon g(x, a, b) \\ \dot{a} &= \epsilon h(x, a, b) \\ \dot{b} &= \epsilon m(x, a, b),\end{aligned}\tag{1.1}$$

where $0 < \epsilon \ll 1$. Of course, the various functions will also depend upon parameters. Inspired by the work of Abrams and Strogatz [1] on modeling language competition and decline (also see Kapitula and Keeverkidis [9] and the references therein), the function f will be given by,

$$f(x, a, b) = s(a + x)^2(1 - a - b - x) - (1 - s)(1 - a - x)^2x.\tag{1.2}$$

Here $0 \leq s \leq 1$ is a reaction parameter which measures the underlying receptivity of the entire population towards opinion A. In linguistic terms, s represents the prestige associated with a particular language. Generally speaking, this model allows for a competition between all A's trying to convince moderate B's to hold opinion A versus all B's trying to convince moderate A's to hold opinion B.

Previous papers, e.g., see Bujalsk et al. [3] and Marvel et al. [10] and the references therein, assumed the proportion of zealots must stay constant. However, Short et al. [12] relaxed this condition and assume the proportions of zealots ("sects" in their language) evolve over time, i.e., they too allow $\epsilon > 0$. Like them, we use our functions g , h , m to model such effects as indoctrination, deradicalization, and spontaneous radicalization. Unlike them, we allow the total proportion of those who hold opinion A and B to vary over time. We find that relaxing this constraint allows for more interesting, and perhaps physically relevant, dynamics.

The paper is organized as follows: In Section 2, we consider in great detail the case when the proportions of zealots are fixed. We find this analysis crucial in understanding the $\epsilon > 0$ problem. In Section 3, we look at various subproblems associated with the full problem. We find that a proper perspective allows us to understand the full dynamics through various concatenations of the subproblems. In Section 4, we consider the dynamics of the full problem for small ϵ . The most interesting conclusion we demonstrate is that in a large region of parameter space there exist stable periodic solutions. These solutions correspond to A and B alternating between being the dominant opinions. Moreover, we find the transition from one opinion being dominant to the other being dominant happens very quickly, but is preceded by a slow increase in the number of zealots. Finally, in Section 5, we conclude and provide some possible future directions for research.

Remark 1.1. In all that follows, we say an opinion "wins" if the total proportion of the population which eventually believes that opinion is greater than 0.5. In other words, A wins if $a + x > 0.5$, and B wins if $a + x < 0.5$.

2. ZEALOTS CONSTANT ($\epsilon = 0$)

Before diving into the three-dimensional system as broadly defined in equation (1.1), we will first begin our study with the simpler model which arises when setting $\epsilon = 0$. In this case, both a and b are constant, and so act as parameters. This reduction allows us to identify normally hyperbolic manifolds for the full system, i.e., two-dimensional surfaces which are invariant for the full flow for small ϵ . The manifolds are approximately the x -nullclines, $f(x, a, b) = 0$. These surfaces will be labeled, x_T , x_M , x_B , and will have the property that when they all exist they can be ordered,

$$0 \leq x_B \leq x_M \leq x_T \leq 1.$$

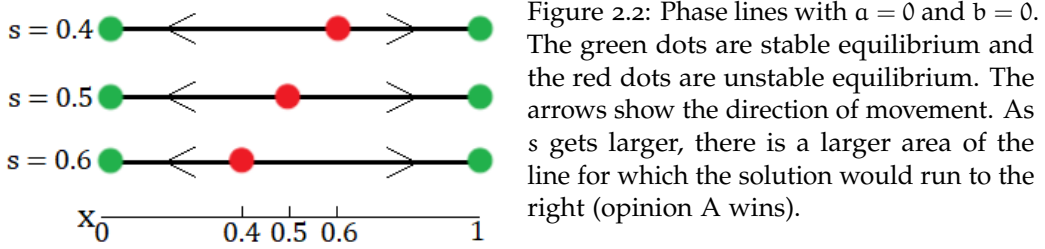
The surfaces x_B , x_T will be exponentially attracting, whereas x_M will be exponentially repelling.

Using (1.2) the equation to be originally studied is,

$$\dot{x} = s(a + x)^2(1 - a - b - x) - (1 - s)(1 - a - x)^2x.\tag{2.1}$$

The first term represents all the A's, $a + x$, convincing the moderate B's, $1 - a - b - x$, to become moderate A's, which increases x . The second term represents all the B's, $1 - a - x$, convincing the moderate A's, x , to become moderate B's, which decreases x . The rate parameter $0 \leq s \leq 1$ tells us which opinion is more favored, i.e., which opinion is intrinsically more appealing.

To see what this means mathematically, suppose $a = b = 0$, and consider the initial condition $x(0) = 0.5$ (both opinions are equally represented). If $s > 0.5$, we have $x(t) \rightarrow 1$ as $t \rightarrow +\infty$, so we say opinion A is more favored by the population. On the other hand, if $s < 0.5$, then $x(t) \rightarrow 0$ as $t \rightarrow +\infty$, so opinion B is more favored. This phenomena is illustrated using phase lines with different values for s in Figure 2.2.



Remark 2.1. We could generalize the reaction terms in the model (2.1) and write,

$$(a + x)^2 \mapsto (a + x)^p, \quad (1 - a - x)^2 \mapsto (1 - a - x)^p.$$

As long as $p > 1$, however, the results will not qualitatively change.

The $\epsilon = 0$ problem is well-understood when $a = b = 0$. We now need to look at what happens for nonzero proportions for the zealots. We will first do an analysis for the edge cases, $a = 0$ or $b = 0$. These analytic expressions will help us to have a better qualitative understanding of the dynamics on the x -nullcline surfaces; in particular, when the proportions of zealots for each opinion are small. We will do numerics for the nullclines when $a, b > 0$.

2.1. $a \geq 0, b = 0$

We start by setting $b = 0$. In this case, the null surfaces can be explicitly determined.

Lemma 2.2. *The x -nullcline surfaces when $b = 0$ are the curves,*

$$\begin{aligned} x_T(a, 0) &= 1 - a \\ x_M(a, 0) &= \frac{1}{2} \left[(1 - s)(1 + a) + \sqrt{(1 - s)^2(1 + a)^2 - 4(1 - s)a} \right] - a \\ x_B(a, 0) &= \frac{1}{2} \left[(1 - s)(1 + a) - \sqrt{(1 - s)^2(1 + a)^2 - 4(1 - s)a} \right] - a. \end{aligned}$$

Moreover, for each s there is a unique saddle-node bifurcation, $x_B(a, 0) = x_M(a, 0)$, when $a = a_{SN}(s)$ with,

$$a_{SN}(s) = \frac{1 - \sqrt{s}}{1 + \sqrt{s}}.$$

Proof: When $b = 0$ we can factor and write,

$$f(x, a, 0) = (1 - a - x) \left[s(x + a)^2 - (1 - s)(1 - a - x)x \right].$$

The equation for x_T follows immediately. Upon rewriting the term in square brackets,

$$s(x + a)^2 - (1 - s)x(1 - a - x) = (x + a)^2 - (1 - s)(1 + a)(x + a) + (1 - s)a,$$

the other two expressions follow via an application of the quadratic formula.

The saddle-node bifurcation occurs when the term inside the radical vanishes,

$$(1-s)^2(1+a)^2 - 4(1-s)a = 0 \quad \rightsquigarrow \quad a^2 - 2\frac{1+s}{1-s}a + 1 = 0.$$

Using the quadratic formula gives the physically relevant, i.e., $0 \leq a \leq 1$, solution,

$$a = B - \sqrt{B^2 - 1}, \quad B = \frac{1+s}{1-s}.$$

Simplify the term under the radical gives,

$$a = \frac{1+s-2\sqrt{s}}{1-s}.$$

The final expression arises upon writing $s = (\sqrt{s})^2$ and simplifying. □

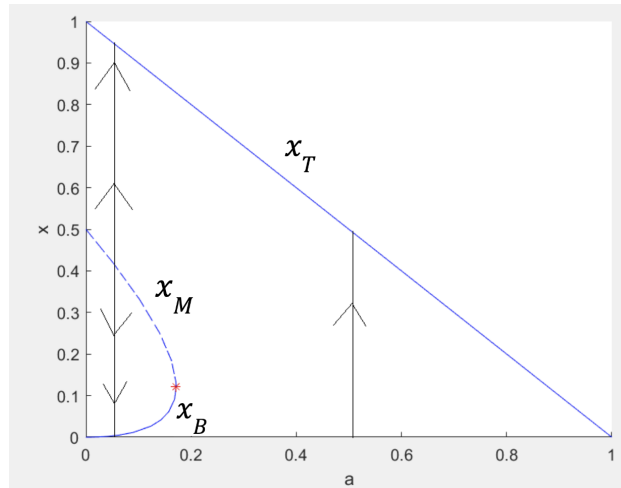


Figure 2.3: Bifurcation plot when $s = 0.5$ and $b = 0$. The blue curves are the equilibrium curves. The solid curves are stable solutions, and the dashed curve is an unstable solution. The top line is x_T , the dashed middle curve is x_M , and the solid bottom curve is x_B . The black lines are phase lines with arrows in the direction of motion.

Since we know the nullclines are where $\dot{x} = 0$, every other region will have either $\dot{x} > 0$ or $\dot{x} < 0$. The arrows on the representative phase lines indicate the direction of motion. When the arrows point towards the nullcline, you have an exponentially stable equilibrium (the solid curves in Figure 2.3). On the other hand, when two arrows point away a nullcline, you have an unstable equilibrium (the dashed curve in Figure 2.3), and the solution curve will flow away from that equilibrium towards the stable equilibrium.

For any equilibrium on x_T , $a + x > 0.5$, so if the solution went towards this equilibrium, A would win. Likewise, for any equilibrium on x_B , $a + x < 0.5$, so B would win. We can see from Figure 2.3 that when $a < a_{SN}$, A or B can win depending on the initial condition of x . When $a > a_{SN}$, however, A will win regardless of the initial condition, as all trajectories will go to x_T . The saddle-node bifurcation point is marked with a red star in Figure 2.3. As s changes, the bifurcation diagram looks qualitatively similar. When s increases, x_M and x_B will get smaller and close in on the origin. As s decreases, x_M and x_B will expand towards the upper right. This means that as an opinion becomes more favored, it needs fewer and fewer zealots to ensure that it will win.

Figure 2.4 provides the a coordinate of the saddle-node bifurcation points when $b = 0$ (blue curve). It also gives us a visual for which opinion wins for which values of a and s . The red region indicates A will win no matter the initial condition, and the purple region indicates the winner depends upon the initial

condition. For $s \geq 0.25$, the boundary is given by the saddle-node bifurcation curve. Since the proportion of zealots on the saddle-node bifurcation curve increases as s decreases, for $s < 0.25$ it is possible for some critical points on x_B to be associated with A winning. This is the underlying reason why the boundary between the two regions no longer follows the saddle-node bifurcation curve.

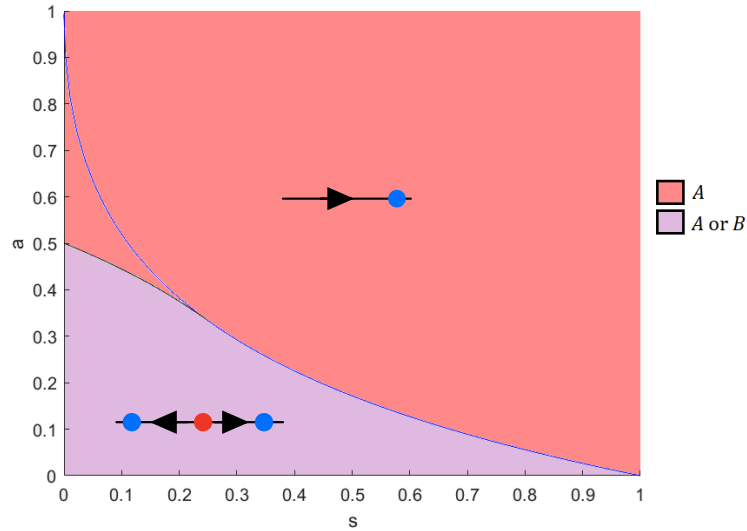


Figure 2.4: A curve, $a = a(s)$, of the saddle-node bifurcations is shown as a blue curve for the case, $b = 0$. When $0.25 \leq s \leq 1$ it is the boundary between the red and purple regions. At $s = 0.25$ it is no longer the boundary, and instead continues into the red region. In the red region opinion A will win no matter the initial conditions (see the associated phase line), while in the purple region opinion A or B could win depending on initial conditions (see the associated phase line).

2.2. $a \geq 0, b = 0.1$

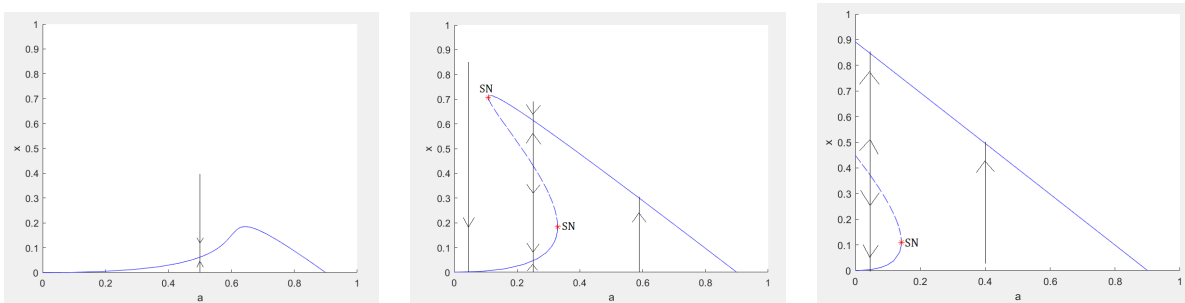


Figure 2.5: Bifurcation diagrams for $b = 0.1$. The blue curves are the equilibrium curves; the solid ones are stable and the dashed are unstable. The black lines are phase lines with arrow in the direction of motion. SN marks the saddle-node bifurcations. In the left figure $s = 0.1$, in the middle figure $s = 0.3$, and in the right figure $s = 0.6$. The choices for s are dictated by Figure 2.6.

We now look at the case when $b = 0.1$. We are no longer able to easily do analytics, so we turn to MATLAB-based numerical bifurcation program Matcont [4] to generate the bifurcation curves. We look at bifurcation diagrams for multiple values of s as they have more significant differences. The bifurcation diagrams with $s = 0.1$ (left plot), $s = 0.3$ (center plot), $s = 0.6$ (right plot) are shown in Figure 2.5.

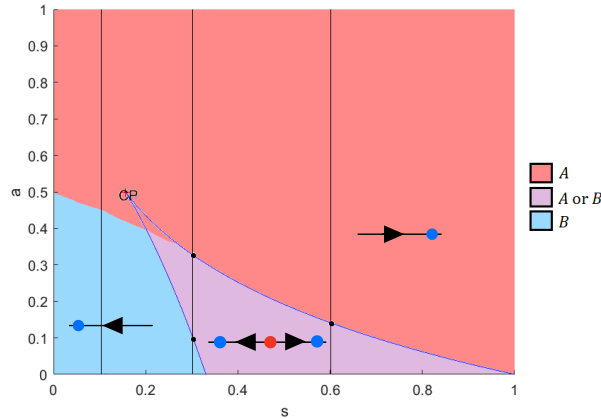


Figure 2.6: The curves, $a = a(s)$, of the saddle-node bifurcation points when $b = 0.1$. Each vertical black line represents a bifurcation diagram from Figure 2.5. For a given s value, the bifurcation diagram can have 0, 1, or 2 saddle-node bifurcation points. In the red region opinion A will win no matter the initial condition (see the associated phase line), while in the blue region opinion B will win (see the associated phase line). In the purple region either of opinion A or B can win (see the associated phase line).

For $s = 0.6$, when $a \lesssim 0.15$, opinion A or opinion B can win depending on the initial condition, but when $a \gtrsim 0.15$, A wins regardless of the initial condition. Similar phenomena happen for $s = 0.3$ and $s = 0.1$. For $s = 0.3$, when $a \lesssim 0.1$, B wins regardless of the initial condition. When $0.15 \lesssim a \lesssim 0.35$, A or B can win depending on the initial condition, and when $a \gtrsim 0.35$, A wins regardless of the initial condition. For $s = 0.1$, when $a \lesssim 0.45$, B wins regardless of the initial condition, and when $a \gtrsim 0.45$, A wins regardless of the initial condition. The key feature of these bifurcation diagrams that determines which side will win is the number and location of saddle-node bifurcations (marked on the diagrams as SN).

Figure 2.6 shows the a coordinate of the saddle-node bifurcation points as a function of s when $b = 0.1$. The label CP marks the cusp point, which is the point at which the individual saddle-node bifurcation curves intersect. The different colors on the graph represent which opinion will be in the majority for given values of a and s . In the red and blue sections, opinion A and B respectively will win for any initial condition. In the purple section, there are two stable equilibrium, so A or B can win depending on the initial condition. The red section goes below the cusp point because when a is approximately 0.5, an equilibrium with even a small number for x will make $x + a > 0.5$, so A will win. This graph means that as A becomes more highly favored by the general population, fewer zealots are needed to ensure that A wins.

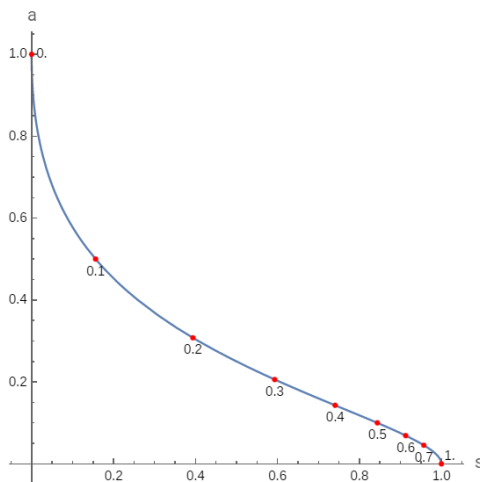


Figure 2.7: A parametric trace of the cusp point as b changes. The corresponding values of b are given below the points marked in red.

If b is changed to a different value, Figure 2.6 remains qualitatively the same. We can see the graph of the cusp point as a parametric function of b in Figure 2.7. The cusp point moves down and to the right as b increases. The graph coming out of the cusp point is qualitatively similar to the graph in Figure 2.6. Physically speaking, we can conclude that when the proportion of B zealots increases, the size of the domain in (s, a) -space for which B wins also increases, and that for A decreases.

2.3. $a = 0, b \geq 0$

Much of the above analysis for $b = 0$ can also be done when $a = 0$. We start with expressions for the x -nullclines (the analogue of Lemma 2.2).

Lemma 2.3. *The x -nullcline surfaces when $a = 0$ are the curves,*

$$\begin{aligned} x_T(0, b) &= \frac{1}{2} \left[2 - s - sb + \sqrt{s^2(1+b)^2 - 4sb} \right] \\ x_M(0, b) &= \frac{1}{2} \left[2 - s - sb - \sqrt{s^2(1+b)^2 - 4sb} \right] \\ x_B(0, b) &= 0. \end{aligned}$$

Moreover, for each s there is a unique saddle-node bifurcation, $x_T(0, b) = x_M(0, b)$, when $b = b_{SN}(s)$ with,

$$b_{SN}(s) = \frac{2 - s - 2\sqrt{1-s}}{s}$$

(see the left figure in Figure 2.8).

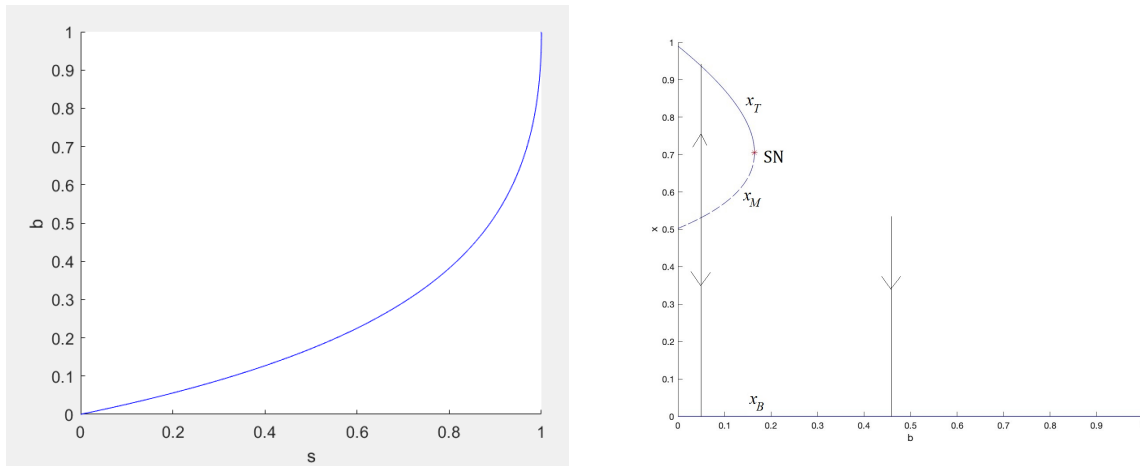


Figure 2.8: The left figure gives the b -value of the saddle-node bifurcation point. The right figure is the bifurcation plot when $a = 0$ and $s = 0.5$. The saddle-node bifurcation point is marked with the SN label. The solid curves, x_T and x_B , are both stable, while the dashed x_M is unstable. Representative phase lines are given with the arrows marking the direction of motion.

Proof: When $a = 0$ we can factor and write,

$$f(x, 0, b) = x \left[sx(1 - b - x) - (1 - s)(1 - x)^2 \right].$$

The equation for x_B follows immediately. The other two expressions follow via an application of the quadratic formula on the term in the square bracket.

The saddle-node bifurcation occurs when the term inside the radical associated with x_M and x_T vanishes,

$$s^2(1+b)^2 - 4sb = 0 \rightsquigarrow b^2 - 2\frac{2-s}{s}b + 1 = 0.$$

Using the quadratic formula gives the physically relevant solution,

$$b = B - \sqrt{B^2 - 1}, \quad B = \frac{2-s}{s}.$$

Routine algebra gives the final result. □

A typical bifurcation diagram is provided in the right figure of Figure 2.8. For $b < b_{SN}$, both x_T and x_B correspond to stable critical points, while x_M is an unstable critical point. The winner depends upon the initial condition. For $b > b_{SN}$, the only critical point is on x_B , so B will win no matter the initial condition.

2.4. $\alpha = 0.1, b \geq 0$

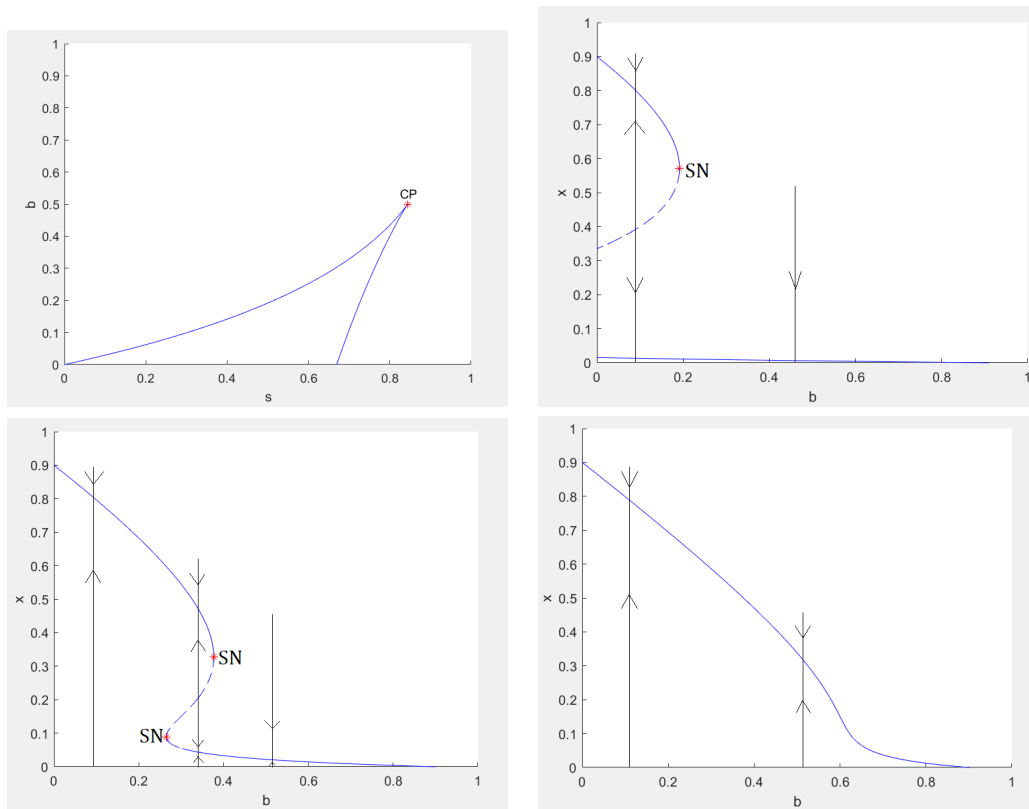


Figure 2.9: The top left figure shows the saddle-node bifurcation points when $\alpha = 0.1$. Bifurcation diagrams are shown in the other three figures. In the top right figure $s = 0.5$, in the bottom left figure $s = 0.75$, and in the bottom right figure $s = 0.9$. The blue curves are the equilibrium curves; the solid ones are stable and the dashed are unstable. The black lines are phase lines with arrows pointing in the direction of motion. SN marks the saddle-node bifurcations.

We now look at the case when $\alpha = 0.1$ using Matcont [4]. The top left figure in Figure 2.9 shows the b -coordinate of the saddle-node bifurcation. The bifurcation curves for $s = 0.5$ (top right figure), $s = 0.75$ (bottom left figure), and $s = 0.9$ (bottom right figure) are also shown in Figure 2.9. The choices for s are dictated by the top left figure. For $s = 0.5$, when $b \lesssim 0.2$, opinion A or opinion B can win depending on the

initial condition, but when $b \gtrsim 0.2$, B wins regardless of the initial condition. For $s = 0.75$, when $b \lesssim 0.27$ A wins regardless of the initial condition. When $0.27 \lesssim b \lesssim 0.4$, A or B can win depending on the initial condition, and when $b \gtrsim 0.4$ B wins regardless of the initial condition. For $s = 0.9$, when $b \lesssim 0.5$, A wins regardless of the initial condition, and when $b \gtrsim 0.5$, B wins regardless of the initial condition.

3. REDUCED DYNAMICS FOR $\epsilon > 0$

We are now ready to allow the proportion of zealots for A and B to evolve over time. For our model (1.1) we specifically set,

$$\begin{aligned} h(x, a, b) &= dx^3 + wa^2x - (1-w)x^2a + rb^2x \\ m(x, a, b) &= d(1-a-b-x)^3 + wb^2(1-a-b-x) - (1-w)(1-a-b-x)^2b + ra^2(1-a-b-x) \\ g(x, a, b) &= -h(x, a, b). \end{aligned}$$

It is straightforward to check that with this choice of functions the physical constraints are met (i.e. all variables are non-negative). The parameters have the following interpretation:

- $d > 0$ controls the rate at which moderates become zealots for no apparent reason (spontaneous radicalization)
- $w \in [0, 1]$ corresponds to the rate at which the zealots convert like-opinioned moderates to become zealots (indoctrination)
- $1 - w$ is the rate at which moderates convince like-opinioned zealots to become moderates (deradicalization)
- $r > 0$ is the rate at which moderates become like-minded zealots as a reaction to the opposing zealots (repulsive radicalization).

Remark 3.1. We will assume in this paper that the rate of radicalization is small. When doing numerical simulations we will set $d = 0.05$ and $r = 0.01$.

Remark 3.2. Again, the exact power is not qualitatively important. Replacing pure quadratic reaction terms with a more generic power, e.g., $a^2x \mapsto a^p x$, does not effect the results as long as $p > 1$.

In order to justify the argument that we can legitimately consider various submodels of the full model to understand the dynamics, we first need a better understanding of the nullcline surfaces. In the statements below, the notation, “ $+\dots$ ”, means there are higher-order terms in the Taylor expansion.

Lemma 3.3. *Assume a, b, d are small, and assume $0 < \epsilon \ll 1$. The attracting x -nullcline surfaces have the Taylor expansions,*

$$\begin{aligned} x_B &= \mathcal{O}\left((a+b)^2\right) \\ x_T &= 1 - a - b + \mathcal{O}\left((a+b)^2\right). \end{aligned}$$

The higher-order terms also depend upon ϵ . The a -nullcline surfaces are,

$$x = 0, \quad x = \frac{w}{1-w}a + \mathcal{O}(a^2) \text{ with } b = 0, \quad x = \frac{1-w}{d}a + \mathcal{O}(a^2) \text{ with } b = 0.$$

The b -nullcline surfaces are,

$$x = 1 - a - b, \quad x = 1 - \frac{1}{1-w}b + \mathcal{O}(b^2) \text{ with } a = 0, \quad x = 1 - \frac{1-w}{d}b + \mathcal{O}(b^2) \text{ with } a = 0.$$

The a - and b -nullcline surfaces do not depend upon ϵ .

Proof: First consider x_B . Write the approximation as $x = Aa + Bb + \mathcal{O}((a+b)^2)$. Plugging into the equation and collecting terms yields,

$$(1-s)(Aa + Bb) + \mathcal{O}((a+b)^2) = 0.$$

At the linear level we need,

$$(1-s)(Aa + Bb) = 0 \quad \rightsquigarrow \quad A = B = 0.$$

The result regarding x_B is proven.

Now consider x_T . Write the approximation as, $x = 1 + Aa + Bb + \mathcal{O}((a+b)^2)$. Plugging into the equation and collecting terms yields, Upon expanding we see at the linear level,

$$-s((A+1)a + (B+1)b) + \mathcal{O}((a+b)^2) = 0.$$

At the linear level we need,

$$-s((A+1)a + (B+1)b) = 0 \quad \rightsquigarrow \quad A = B = -1.$$

The result regarding x_T is proven.

The a -nullclines are found by solving $h(x, a, b) = 0$. After factoring out $x = 0$ this means,

$$dx^2 + wa^2 - (1-w)xa + rb^2 = 0.$$

Write the linear approximation for x , $x = Aa + Bb + \mathcal{O}((a+b)^2)$. Substitution into the above yields,

$$d(A^2a^2 + 2ABab + B^2b^2) + wa^2 - (1-w)(Aa + Bb)a + rb^2 + \mathcal{O}((a+b)^3) = 0.$$

After keeping only the quadratic terms,

$$(dA^2 - (1-w)A + w)a^2 + (2dAB - (1-w)B)ab + (dB^2 + r)b^2 = 0.$$

For the expansion to make sense, each coefficient must be zero. Setting the first coefficient to zero gives,

$$A = \frac{1}{2d} \left(1 - w \pm \sqrt{(1-w)^2 - 4d} \right).$$

A Taylor expansion gives for small d ,

$$A = \frac{w}{1-w} + \mathcal{O}(d), \quad A = \frac{1-w}{d} + \mathcal{O}(1).$$

The last two coefficients being zero requires, $b = 0$. For small a, b the leading equation is,

$$wa^2 - (1-w)xa = 0 \quad \rightsquigarrow \quad a[wa - (1-w)x] = 0.$$

The result is now proven.

The b -nullclines are found by solving $m(x, a, b) = 0$. Upon setting $y = 1 - a - b - x + \mathcal{O}((a+b)^2)$, this is mathematically equivalent to solving for the a -nullclines. Substituting y in for x in the a -nullcline expressions, and then solving for x , gives the result. \square

We are now ready to look at the evolution of the zealot variables on the attracting x -nullclines. First suppose $x = x_B$. The expansion provides,

$$\dot{a} = \epsilon \mathcal{O}((a+b)^4), \quad \dot{b} = \epsilon [d(1-3a-3b) - (1-w)b + \dots].$$

For a and b small this means $\dot{b} \gg \dot{a}$ as long as the trajectory is not close to the b -nullcline. Since d is assumed small, this further implies that $\dot{b} < 0$, i.e., $b(t)$ is decreasing, until the trajectory gets close to the b -nullcline. In conclusion, until the b -nullcline is reached, we have the approximation $a(t) \equiv \text{const}$ (see

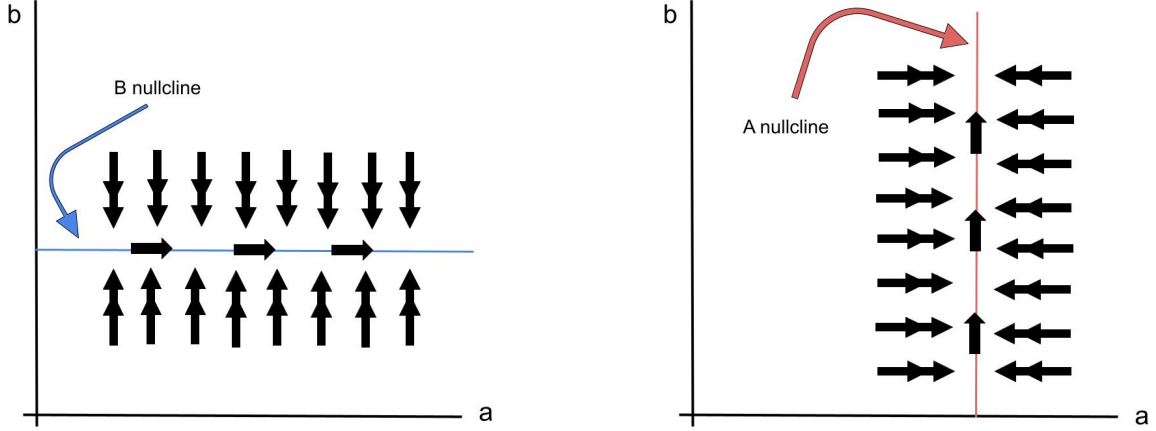


Figure 3.10: In the left figure, a cartoon of the dynamics on x_B is given for small a, b . The rate of motion towards the b -nullcline is much larger than the rate of change for a . In the right figure, a cartoon of the dynamics on x_T is given for small a, b . The rate of motion towards the a -nullcline is much larger than the rate of change for b .

the left figure in Figure 3.10). Now suppose $x = x_T$. The expansion for x_T implies this attracting surface is $\mathcal{O}((a+b)^2)$ -close to the b -nullcline. On the other hand,

$$\dot{a} = \epsilon [d(1 - 3a - 3b) - (1 - w)a + \dots],$$

which to leading order is the ODE for b on x_B . We conclude $\dot{a} \gg \dot{b}$ as long as the trajectory is not near an a -nullcline. Since d is small, we also have $a(t)$ is decreasing until the trajectory nears the a -nullcline. In conclusion, until the a -nullcline is reached, we have the approximation $b(t) \equiv \text{const}$ (see the right figure in Figure 3.10). We now see that in order to understand the full dynamics, we must first look at the subproblems where a is constant, $\dot{a} \approx 0$, and b is constant, $\dot{b} \approx 0$.

3.1. $\dot{b} \approx 0$

Let us first assume $\dot{b} \approx 0$. The governing equations are,

$$\begin{aligned} \dot{x} &= s(a+x)^2(1-a-x-b) - (1-s)(1-a-x)^2x - \epsilon(dx^3 + wa^2x - (1-w)x^2a - rb^2) \\ \dot{a} &= \epsilon(dx^3 + wa^2x - (1-w)x^2a + rb^2x) \end{aligned} \quad (3.1)$$

For this subproblem, we treat b as a parameter. Since b is implicitly assumed to be small, for ease we will initially assume $b = 0$. This choice allows us to derive analytic expressions (e.g., see Lemma 2.2).

For ϵ small, the x -nullclines, $x_B(a, 0)$ and $x_T(a, 0)$, are attracting curves in the (a, x) -phase plane. On the other hand, $x_M(a, 0)$ is a repelling curve. The critical points for the system are found through the intersection of the a -nullclines with the x -nullclines. For small a and b , expressions for the nullclines are given in Lemma 3.3. Supposing s is fixed, we have two possibilities depending on the value for w . If w is sufficiently small, the a -nullclines will intersect x_B and x_M , which allows for a stable and unstable interior critical point (see the left figure in Figure 3.11). On the other hand, if w is sufficiently large the a -nullclines will intersect x_M only, and both interior critical points will be unstable (see the right figure in Figure 3.11).

Figure 3.11 shows a plot of the a and x -nullclines with $s = 0.3$ for two different w values. The blue curves are the x -nullclines, and the red lines are the a -nullclines. The stable equilibrium are marked with green dots while the unstable are marked with red dots. Because $\epsilon \ll 1$, the trajectories quickly move vertically to an x -nullcline, and then move slowly along the nullcline. The arrows along the nullclines

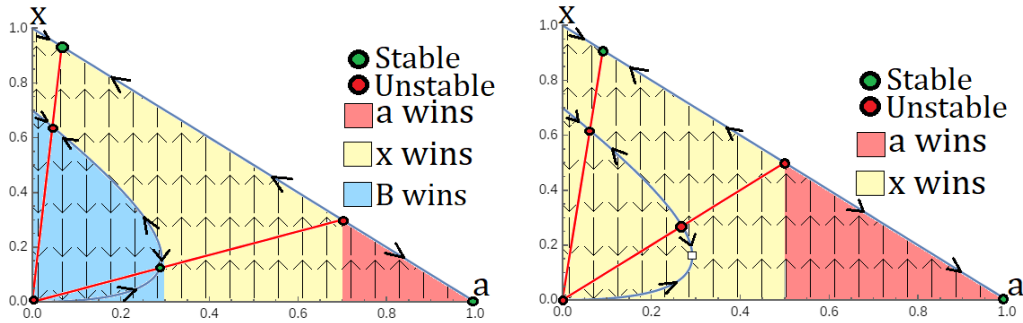


Figure 3.11: Phase planes with nullclines with $s = 0.3$. The blue curves are the x -nullclines and the red lines are the a -nullclines. The stable equilibrium are marked with green dots while the unstable are marked with red dots. In the left figure $w = 0.3$, and in the right figure $w = 0.5$. Note that the critical point moves from x_B (stable) to x_M (unstable) as w changes. The shaded regions indicate which equilibrium the solution goes to for a given initial condition. The white square in the figure on the right indicates the jump point (i.e., the saddle-node bifurcation point). A trajectory starting near that point moves towards it before “jumping off” and quickly moving to x_T .

indicate where the solution curves will go once they hit the x -nullcline. When a solution gets to the white square (saddle-node bifurcation point), it leaves the curve and quickly jumps to x_T . The shaded regions indicate the equilibrium towards which the solution goes for a given initial condition for small ϵ . An initial condition in the yellow region will go to the equilibrium on x_T with small a , so moderate A wins. Similarly, an initial condition in the red region will go to the equilibrium on x_T with $x = 0$ and $a = 1$, hence zealot A wins. Finally, an initial condition in the blue region will go to the stable equilibrium on x_B with small x and small a , so moderate B wins. The left figure in Figure 3.11, where $w = 0.3$, shows three stable equilibria, so B can win for some initial conditions. The right figure in Figure 3.11, where $w = 0.5$, shows that B can never win.

This bifurcation takes place when the (almost linear) a -nullcline with the $\mathcal{O}(1)$ slope intersects the saddle-node bifurcation point on the x -nullcline. The transition value for w can be approximated using the analytic approximations for the nullclines. The below result is illustrated in Figure 3.12.

Lemma 3.4. *When $b = 0$ the a -nullcline intersects the saddle-node bifurcation points when,*

$$w_T = \frac{\sqrt{s}}{1 + \sqrt{s}}.$$

Proof: The a -value of the saddle-node bifurcation point is given in Lemma 2.2,

$$a = a^* = \frac{1 - \sqrt{s}}{1 + \sqrt{s}}.$$

Plugging this value into the expression for $x_B(a, 0)$ yields the x -coordinate,

$$x = x^* = \frac{\sqrt{s} - s}{1 + \sqrt{s}}.$$

The transition occurs when the saddle-node bifurcation point intersects the a -nullcline which has the approximation,

$$x \approx \frac{w}{1 - w} a \rightsquigarrow \frac{x^*}{a^*} = \frac{w}{1 - w} \rightsquigarrow w = \frac{x^*}{a^* + x^*}.$$

Simplifying gives the desired result, which is illustrated in Figure 3.12. □

The analysis above used the expressions available for $b = 0$. If $b > 0$ is small, then the picture will be qualitatively the same. There will be a threshold, $w = w_T$, such that either opinion can win if $w < w_T$,

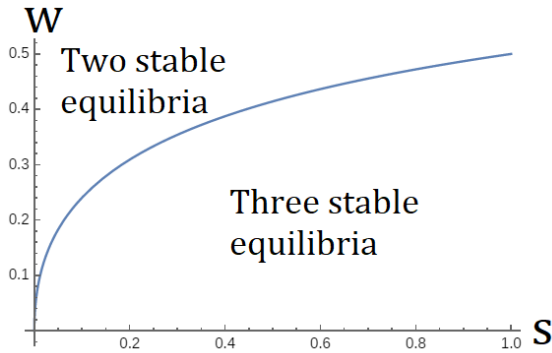


Figure 3.12: The dividing curve, $w = w_T$, is given in Lemma 3.4. For $w > w_T$, both stable equilibria are on x_T , so opinion A always wins (see the right figure in Figure 3.11). For $w < w_T$, there is an additional stable critical point on x_B , so it is possible for opinion A or B to win depending on initial conditions (see the left figure in Figure 3.11).

and A will always win if $w > w_T$. The x -nullclines will look as in Figure 2.5, and the a -nullclines will intersect them at certain points. There will be at most three stable equilibria, and B can win only if one of the equilibria resides on x_B .

3.2. $\dot{a} \approx 0$

Now assume $\dot{a} \approx 0$. In this case the governing equations are,

$$\begin{aligned} \dot{x} &= s(a+x)^2(1-a-x-b) - (1-s)(1-a-x)^2x - \epsilon(dx^3 + wa^2x - (1-w)x^2a - rb^2) \\ \dot{b} &= \epsilon \left[d(1-a-b-x)^3 + wb^2(1-a-b-x) - (1-w)(1-a-b-x)^2b + ra^2(1-a-b-x) \right]. \end{aligned} \quad (3.2)$$

For this subproblem, we treat a as a parameter. Since a is implicitly assumed to be small, for ease we will initially assume $a = 0$. This choice allows us to derive analytic expressions (e.g., see Lemma 2.3).

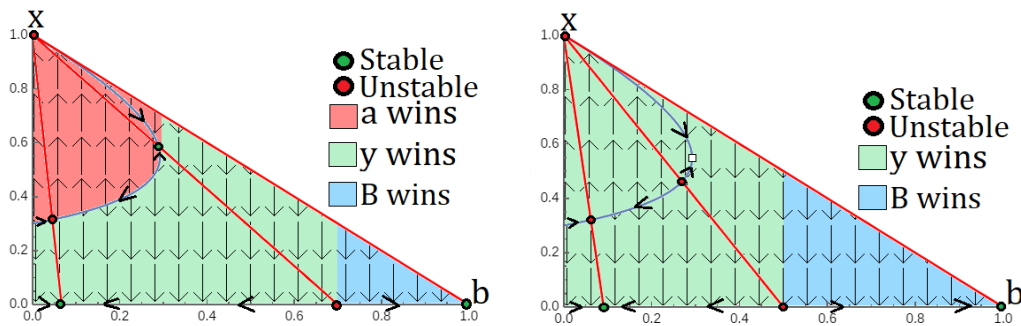


Figure 3.13: Phase planes with nullclines with $s = 0.7$. The blue curves are the x -nullclines and the red lines are the b -nullclines. The stable equilibria are marked with green dots while the unstable are marked with red dots. The left figure has $w = 0.3$, and the right figure has $w = 0.5$. Note the critical point moves from x_T (stable) to x_M (unstable) as w changes. The shaded regions indicate which equilibrium the solution goes to for a given initial condition. The white square in the figure on the right indicates the jump point (i.e., the saddle-node bifurcation point). A trajectory starting near that point moves towards it before “jumping off” and quickly moving to x_B .

For ϵ small the x -nullclines, $x_B(0, b)$ and $x_T(0, b)$, are attracting curves in the (b, x) -phase plane. On the other hand, $x_M(0, b)$ is a repelling curve. The critical points for the system are found through the intersection of the b -nullclines with the x -nullclines (see Figure 3.13). Supposing s is fixed, we have two possibilities depending on the value for w . If w is sufficiently small, the b -nullclines will intersect all three x -nullclines, which allows for a stable and unstable interior critical point (see the left graph in Figure 3.13). On the other hand, if w is sufficiently large the b -nullclines will intersect x_M and x_B only, and both interior

critical points will be unstable (see the right graph in [Figure 3.13](#)). When $a = 0$, the transition value for w can be explicitly computed.

Lemma 3.5. *When $a = 0$ the b -nullcline intersects the saddle-node bifurcation points when,*

$$w_T = \frac{2s - 2 + (2 - s)\sqrt{1 - s}}{s[1 - \sqrt{1 - s}]}$$

(see [Figure 3.14](#)).

Proof: The b -value of the saddle-node bifurcation point is given in [Lemma 2.3](#),

$$b = b^* = \frac{2 - s - 2\sqrt{1 - s}}{s}.$$

Plugging this value into the expression for $x_T(0, b)$ yields the x -coordinate,

$$x = x^* = \frac{1}{2}(2 - s - sb^*).$$

The transition occurs when the saddle-node bifurcation point intersects the b -nullcline which has the approximation,

$$x \approx 1 - \frac{1}{1 - w}b \rightsquigarrow \frac{1}{2}(2 - s - sb^*) = 1 - \frac{1}{1 - w}b^* \rightsquigarrow w = 1 - \frac{2b^*}{s + sb^*}.$$

Simplifying gives the desired result. □

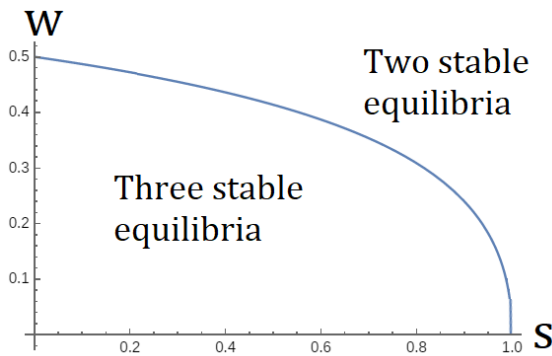


Figure 3.14: The dividing curve $w = w_T$ given in [Lemma 3.5](#). For $w > w_T$, there are only two stable equilibria, both of which are on x_B , so opinion B always wins (see the right figure in [Figure 3.13](#)). For $w < w_T$, there is an additional stable equilibrium on x_T , so it is possible for A or B to win (see the left figure in [Figure 3.13](#)).

The analysis above used the expressions available for $a = 0$. If $a > 0$ is small, then the picture will be qualitatively the same. There will be a threshold, $w = w_T$, such that either opinion can win if $w < w_T$, and B will always win if $w > w_T$. The x -nullclines will look as in [Figure 2.9](#), and the b -nullclines will intersect them at certain points. There will be at most three stable equilibria, and A can win only if one of the equilibria resides on x_T .

4. DYNAMICS FOR THE FULL SYSTEM

We are now ready to consider the dynamics for the full system,

$$\begin{aligned} \dot{x} &= s(a + x)^2(1 - a - b - x) - (1 - s)(1 - a - x)^2x - \epsilon(dx^3 + wa^2x - (1 - w)x^2a - rb^2x) \\ \dot{a} &= \epsilon(dx^3 + wa^2x - (1 - w)x^2a + rb^2x) \\ \dot{b} &= \epsilon \left(d(1 - a - b - x)^3 + wb^2(1 - a - b - x) - (1 - w)(1 - a - b - x)^2b + ra^2(1 - a - b - x) \right) \end{aligned} \tag{4.1}$$

We will primarily be looking at how the dynamics change as a function of s and w . The results of the previous section are crucial in all that follows. We continue to assume $0 < \epsilon \ll 1$. When doing numerical simulations, we will fix $\epsilon = 0.1$. In all that follows we will use colloquial language like “hitting a surface”, or “being on a surface”. What is meant with this language is that a solution curve gets to, or is very close to, the surface, and so the dynamics are governed by the surface dynamics. Going forward we will use the following notation:

- a_{SN} : the a -value for the saddle-node bifurcation point when b is frozen
- a_{u} : the a -value for the unstable point on x_{T} when b is frozen
- b_{SN} : the b -value for the saddle-node bifurcation point when a is frozen
- b_{u} : the a -value for the unstable point on x_{B} when a is frozen.

4.1. Critical points

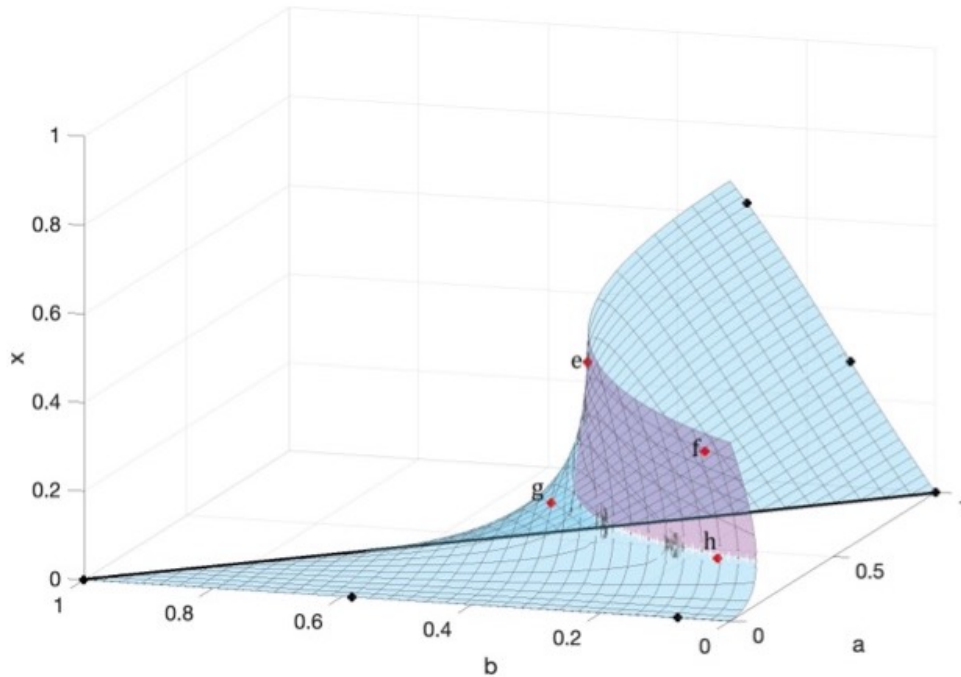


Figure 4.15: x -nullcline and equilibria when $s = 0.6$ and $w = 0.4$. The blue parts are x_{B} and x_{T} , while the purple part is x_{M} . The labeled red dots are the equilibria in the interior. The black dots are equilibria on the edges of the surface, $a + x = 0$ (all B) and $a + x = 1$ (all A). The black line, $a + b = 1$ with $x = 0$, is an attracting continuum of critical points where the entire population is filled with zealots.

The above discussion showed that the dynamics crucially depended on the existence of interior stable critical points. Consequently, we expect the solution behavior for the full system to be governed by the interior stable critical points on x_{B} and/or x_{T} . We will need to see where they are located as a function of s and w . For an example of what is possible, suppose $s = 0.6$ and $w = 0.4$. In Figure 4.15 we have a graph of the x -nullcline, as well as all the critical points. The parts shaded blue represent those parts of the surface which are stable for small ϵ , and that shaded purple is the unstable part, x_{M} . The dots represent the equilibria, and can be categorized as follows:

- There are three equilibria on the line $a + x = 1$, where the entire population supports opinion A. The middle point is unstable, and the other two are stable.
- There are three equilibria on the line $a + x = 0$, where the entire population supports opinion B. The middle point is unstable, and the other two are stable.
- The line $a + b = 1$ with $x = 0$ is comprised entirely of critical points, and is stable. In this case, the entire population is filled with zealots.
- The red dots in the interior of the surface can be stable or unstable depending on the parameter values. Equilibria f and g are always unstable. The point f always remains on the unstable surface, x_M , and so will always be unstable in the x direction. Additionally, it will have a stable and unstable direction on the x -nullcline. The point g can be stable or unstable in the x direction, depending whether it is on x_B or x_M , but will always have a stable and unstable direction on the x -nullcline. Equilibria e and h can be stable or unstable. They will always be stable along the x -nullcline, so the overall stability of the point depends on whether or not it is on a stable x -nullcline. Depending on the value of s and w , these points can reside on any of the three parts of the x -nullcline.

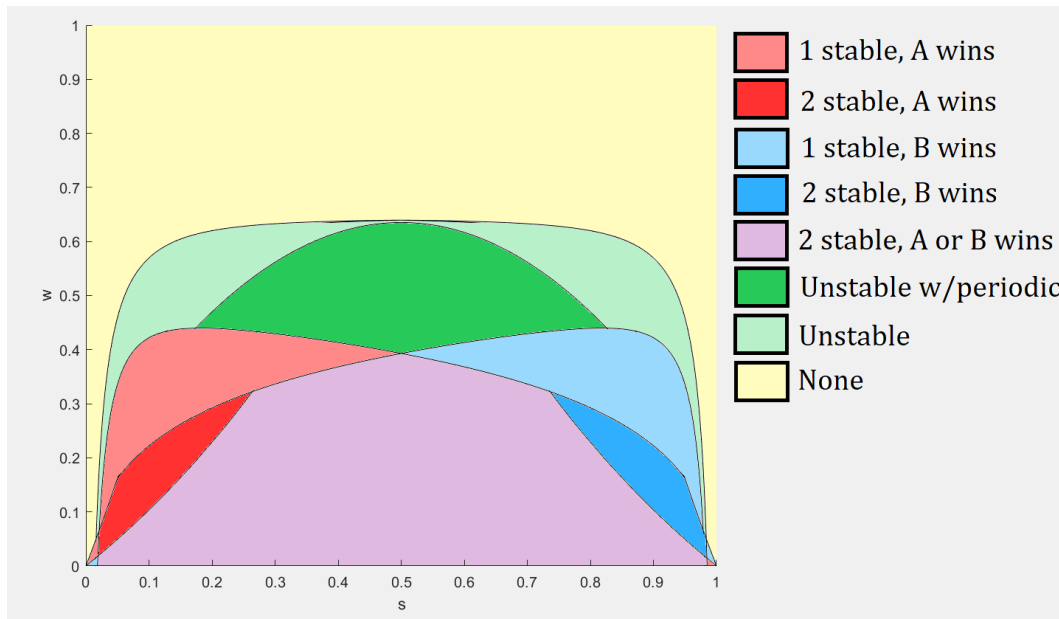


Figure 4.16: A graph of the stability of interior equilibria, and showing which opinion wins. The legend indicates how many stable equilibria there are, and which opinion wins at those equilibria.

The interior equilibria, shown with red dots in [Figure 4.15](#), are the most interesting of the equilibria because they move around and can change stability as s and w change. For certain values of s and w , the points e and f , and the points g and h , can disappear via a saddle-node bifurcation. Consider [Figure 4.16](#). The legend dictates how many stable critical points there are, and computes which opinion wins according to the stable critical point. In the purple, dark red, and dark blue regions, there are two stable critical points. The transition curve between two stable and one stable critical point follows from either a saddle-node bifurcation (extreme s values) or Hopf bifurcation (moderate s values). This curve was computed using Matcont [4]. The next transition curve follows from a Hopf bifurcation for the stable critical point. Again, this curve was generated using Matcont. It was determined numerically using Matcont that the size of the domain for which there is a stable periodic solution arising from the Hopf bifurcation is very small; hence, it is not shown here. The light green and dark green regions are where there are interior critical points, but they are all unstable. The upper boundary of the light green region was computed using Matcont. The

upper boundary of the dark green region is where $b_{SN} = b_u$, or $a_{SN} = a_u$. The dark green region is where there are no stable interior critical points, and where $b_{SN} < b_u$ and $a_{SN} < a_u$.

More specifically, consider what happens when $s = 0.2$. For small w there are two stable equilibria in the interior, one of which represents opinion A winning and one of which represents B winning. As w increases there are still two stable critical points, but now both are associated with opinion A winning. Once we move into the pink region one of the stable critical points has become unstable via a Hopf bifurcation. The remaining stable equilibrium is associated with A winning. As we further increase w the stable critical point destabilizes via a Hopf bifurcation, say at w^* . The bifurcating solution is stable, but it exists for only a very small range, i.e., $0 < w - w^* < 10^{-3}$. The point where this periodic solution ceases to exist is not marked. As w further increases there are no longer any stable critical points in the interior, but $b_{SN} < b_u$ and $a_{SN} < a_u$. We will discuss this in more detail shortly. When we next pass into the light green area, we have no stable interior equilibria, and $b_{SN} > b_u$ and/or $a_{SN} > a_u$. Finally, when we enter the yellow region there are no longer any interior equilibria. Something similar happens for $s = 0.8$, except A winning is replaced by B winning.

4.2. Jumping between the attracting surfaces

We have already talked about the dynamics on x_B and x_T . We now discuss the dynamics associated with jumping from x_B to x_T or vice-versa. Assume for the initial condition that $a(0)$, $b(0)$ are small. This assumption ensures the discussion surrounding [Figure 3.10](#) is applicable. Without loss of generality, assume $x(0)$ is sufficiently small so that the initial condition lies below the unstable surface, x_M .

Because ϵ is small the solution will first quickly move to the stable surface, x_B , and do so in such a manner that the a - and b -values do not significantly change. Once on x_B , the a -value will remain unchanged, while the b -value will change monotonically until the b -nullcline surface is hit. Once the b -nullcline is hit, the value for b is frozen, and we have the dynamics associated with the subproblem [\(3.1\)](#) (see the top left figure in [Figure 4.17](#)).

There are now two possibilities:

- there is a stable critical point on x_B (top right plot in [Figure 4.17](#)), or
- there is not a stable critical point on x_B (bottom two plots in [Figure 4.17](#)).

In the former case, the trajectory asymptotically approaches the stable critical point. In the latter case, the trajectory moves along x_B until $a = a_{SN}$, which is the saddle-node bifurcation point for the subproblem associated with b being frozen. The solution now jumps up to the stable surface, x_T . Since the jump is vertical we still have $a = a_{SN}$. If at the jump $a < a_u$, the solution moves to the left (bottom left plot in [Figure 4.17](#)). If at the jump $a > a_u$, the solution moves to the right and eventually goes to the line of zealots, $a + b = 1$ (bottom right plot in [Figure 4.17](#)). The value of w for which $a = a_u$ when $b = 0$ is given in [Lemma 3.5](#).

Suppose the solution is now on the surface x_T . When the trajectory first hits the surface, the a -value will be approximately the value associated with the saddle-node bifurcation point, a_{SN} , and the b -value will be that associated with the intersection of the b -nullcline and x_B . If $a_{SN} < a_u$, the b -value will remain unchanged, while a will decrease until the attracting a -nullcline is hit (see the bottom left plot in [Figure 4.17](#)). On the other hand, if $a_{SN} > a_u$, the b -value will remain unchanged, while a will increase until the attracting line of critical points representing all zealots, $a + b = 1$, is hit (see the bottom right plot in [Figure 4.17](#)).

If $a_{SN} < a_u$ and the a -nullcline is hit, the value for a is frozen, and we have the dynamics associated with the subproblem [\(3.2\)](#) (see [Figure 4.18](#)). There are now two possibilities:

- there is a stable critical point on x_T (top right plot of [Figure 4.18](#)), or
- there is not a stable critical point on x_T (bottom two plots of [Figure 4.18](#)).

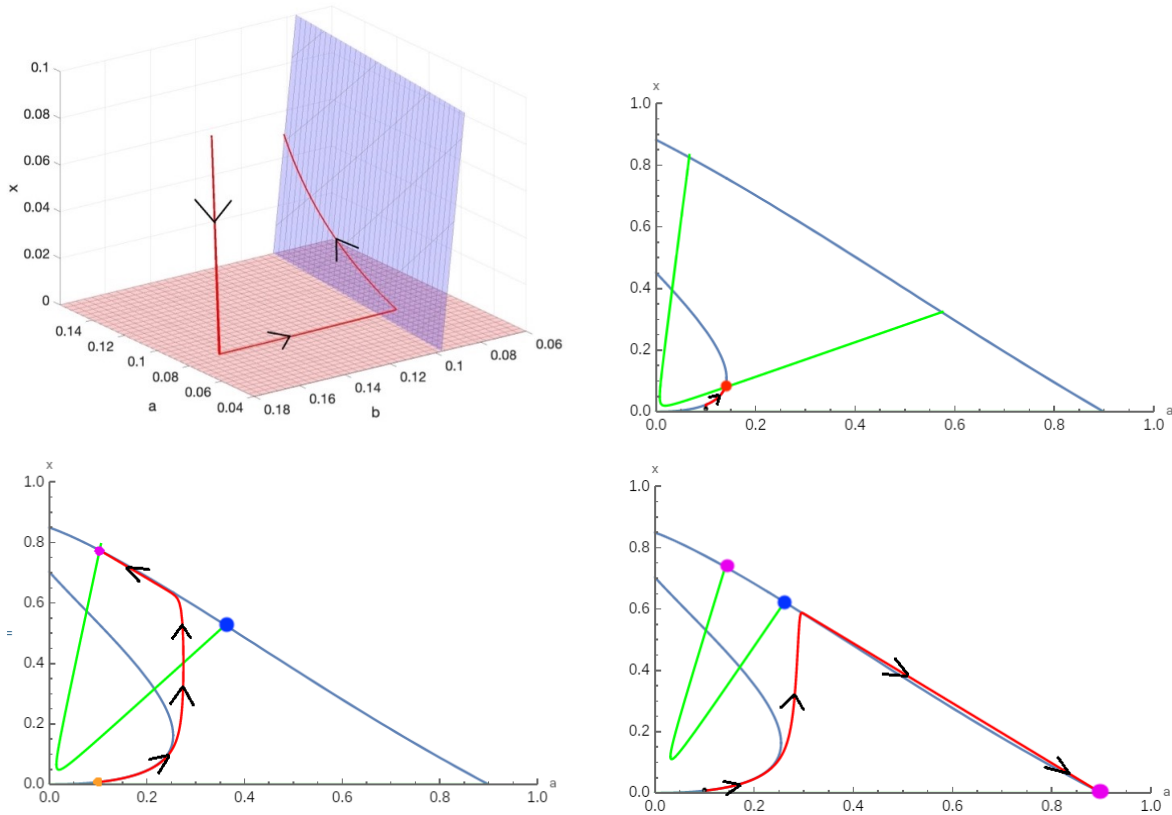


Figure 4.17: A sample trajectory when $s = 0.5$ and $w = 0.5$ is given in the top left figure. The a -nullcline surface is shown in light red, and the b -nullcline surface is shown in blue. Once the solution has fallen onto x_B , it slowly evolves with approximately constant a until it hits the b -nullcline. Once on the b -nullcline surface, the trajectory is governed by the subproblem (3.1). The next three figures show possible trajectories on the attracting b -nullcline surface for $b \approx 0.1$. The red curves are the solutions with arrows pointing in the direction of motion. The blue curves are the x -nullclines, and the green lines are the a -nullclines. In the top right plot, $s = 0.6$ and $w = 0.35$. There is a stable critical point on x_B (marked with a red dot), so the solution does not jump up to x_T . In the bottom left plot, $s = 0.4$ and $w = 0.55$, and there is no stable critical point on x_B . The solution jumps up to x_T and moves to the left as $a_{SN} < a_u$. In the bottom right plot, $s = 0.4$ and $w = 0.62$. The solution jumps up to x_T and moves to the right as $a_{SN} > a_u$.

In the former case, the trajectory asymptotically approaches the stable critical point. In the latter case, the trajectory moves along x_T until $b = b_{SN}$. Since ϵ is small, the solution drops vertically to the stable surface, x_B . If $b < b_u$, the solution moves to the left (bottom left figure in Figure 4.18), and if $b > b_u$, the solution moves to the right and eventually goes to the line of zealots, $a + b = 1$ (bottom right plot in Figure 4.18). The value of w for which $b = b_u$ when $a = 0$ is given in Lemma 3.4.

Suppose the solution has fallen back down to x_B . The b -value, b_{SN} , will be that associated with the saddle-node bifurcation point for the subproblem associated with a being frozen. There are now two possibilities. If $b_{SN} < b_u$, then the solution will head back towards the plane, $b = 0$, with a remaining constant (see the bottom left plot in Figure 4.17). Eventually, b will be small enough that the a -nullcline is hit, and we start the cycle over. On the other hand, if $b_{SN} > b_u$, the solution will move towards the boundary, $a + b = 1$ with $x = 0$, and asymptotically approach a critical point (see the bottom right plot in Figure 4.17).

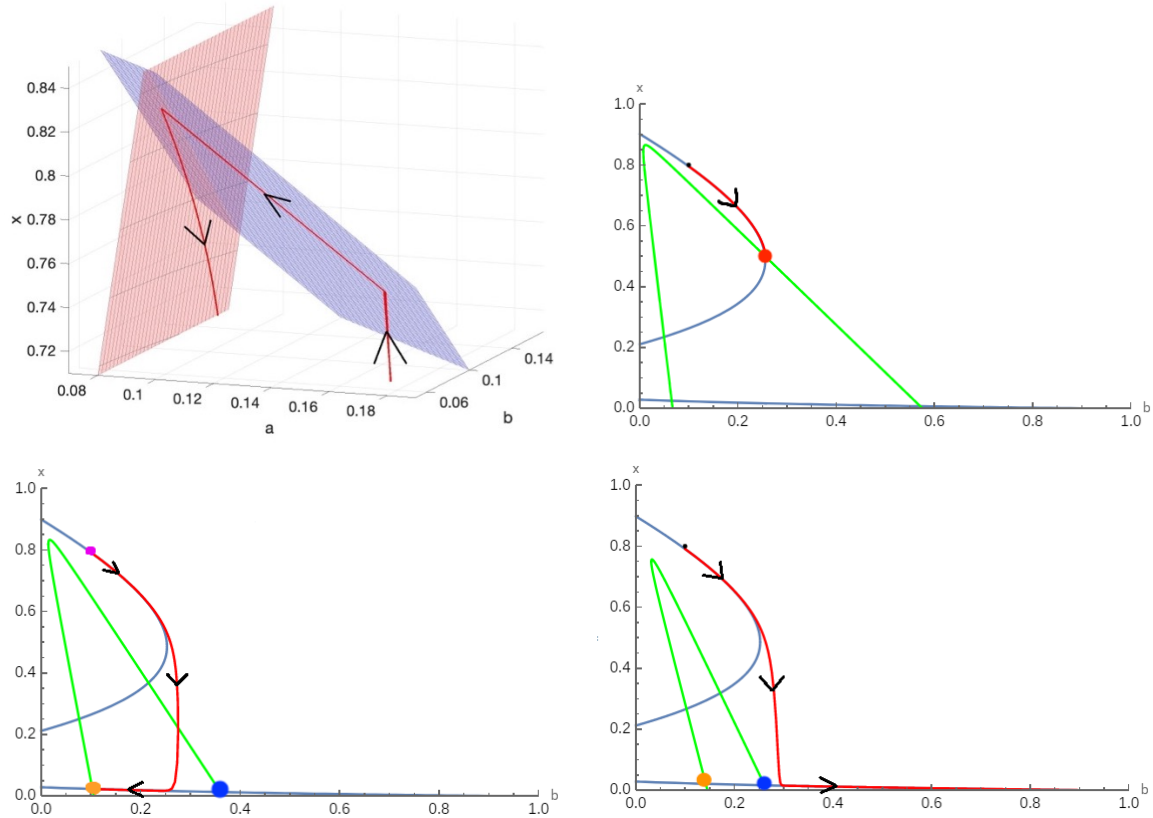


Figure 4.18: A sample trajectory when $s = 0.5$ and $w = 0.5$ is given in the top left figure. The a -nullcline surface is shown in light red, and the b -nullcline surface is shown in blue. Once the solution has jumped to x_T , it slowly evolves with approximately constant b until it hits the a -nullcline. Once on the a -nullcline surface, the trajectory is governed by the subproblem (3.2). The next three figures show possible trajectories for $s = 0.6$ on the attracting a -nullcline surface for $a \approx 0.1$. The red curves are the solutions with arrows pointing in the direction of motion. The blue curves are the x -nullclines, and the green lines are the b -nullclines. In the top right plot, $w = 0.35$. There is a stable critical point on x_T (marked in red), so the solution does not fall down to x_B . In the bottom left plot $w = 0.55$, and there is no stable critical point on x_T . In this case, the solution does fall down to x_B and moves to the left ($b_{SN} < b_u$). In the bottom right plot, $w = 0.62$, and the solution falls down to x_B , and moves to the right ($b_{SN} > b_u$).

4.3. Full dynamics and periodic solutions

We are now ready to discuss the dynamics for the full problem. The stability diagram of Figure 4.16 will be our guide for choosing representative values of s and w . Figure 4.19 provides a variety of solution curves against the x -nullcline to study in further detail.

First, suppose s and w are chosen in the pink region. There is one stable critical point in the interior, and it lies on x_T . A solution starting below x_M will drop to x_B , and from there, it will slowly move to the b -nullcline. Once there, it will follow the trajectory outlined in the bottom left plot of Figure 4.17, hit the b -nullcline, then will follow the trajectory outlined in the top right plot of Figure 4.18 and approach a stable critical point on x_T . A typical solution trajectory with $s = 0.2$ and $w = 0.35$ is plotted in the top left of Figure 4.19.

Now suppose s and w are chosen in the light blue region. There is one stable critical point in the interior, and it lies on x_B . A solution starting above x_M will rise to x_T , and from there, it will slowly move to the a -nullcline. Once there, it will follow the trajectory outlined in the bottom left plot of Figure 4.18, hit the

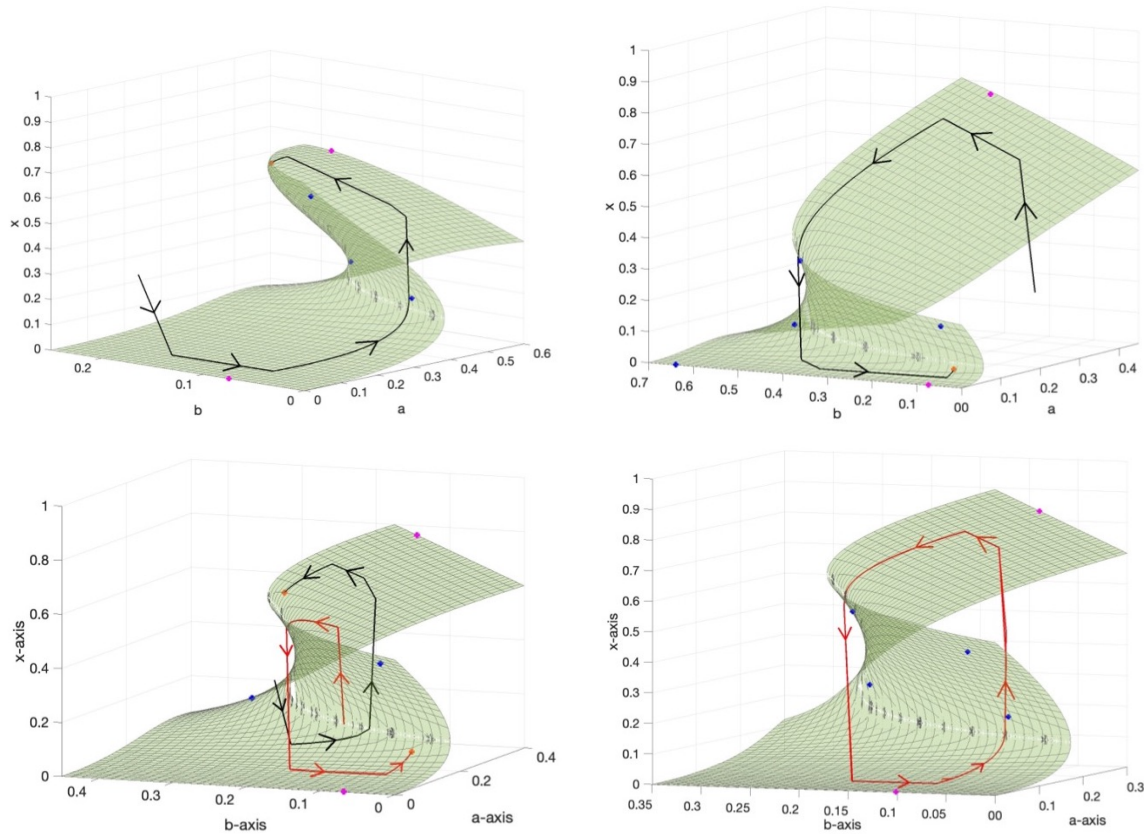


Figure 4.19: A variety of solution curves against the x -nullcline. The blue dots are unstable equilibria, the pink dots are degenerate equilibria, and the orange dots are stable equilibria. In the top left $s = 0.2$ and $w = 0.35$, in the top right $s = 0.8$ and $w = 0.35$, in the bottom left $s = 0.5$ and $w = 0.3$, and in the bottom right $s = 0.5$ and $w = 0.5$.

a -nullcline, then follow the trajectory outlined in the top right plot of Figure 4.17 and approach a stable critical point on x_B . A typical solution trajectory with $s = 0.8$ and $w = 0.35$ is plotted in the top right of Figure 4.19.

Next, suppose s and w are chosen in the purple region. There are two stable critical points in the interior, and one lies on x_B , while the other lies on x_T . Depending on the initial condition, one of the two above scenarios will be followed. A typical set of solution trajectories with $s = 0.5$ and $w = 0.3$ is plotted in the bottom left of Figure 4.19.

Finally, suppose s and w are chosen in the dark green region. Here, the trajectory on the b -nullcline is illustrated in the bottom left figure in Figure 4.17, and the trajectory on the a -nullcline is illustrated in the bottom left figure in Figure 4.18. Since there are no stable critical points, the solution keeps on jumping between the two attracting x -nullclines. The solution always falls for $b = b_{SN}$, and after falling, the a -value for the trajectory stays constant until the a -nullcline is hit (see the left plot in Figure 3.10). The solution jumps for $a = a_{SN}$, and after jumping, the b -value stays constant until the b -nullcline is hit (see the right plot in Figure 3.10). Thus, it is plausible that there may be an attracting periodic solution. This is precisely what is seen in the bottom right figure in Figure 4.19.

A periodic solution is plotted in more detail in Figure 4.20. Here, we started with an arbitrary initial condition, and did not plot the transients associated with the solution. The left figure is the solution in the phase space. The upper right figure is the proportion of zealots for opinion A, a , and the total proportion of those who hold opinion A, $a + x$, plotted as a function of time. Note that A wins only after the proportion of zealots increases to a critical level, which is approximately the value of a for the appropriate saddle-node

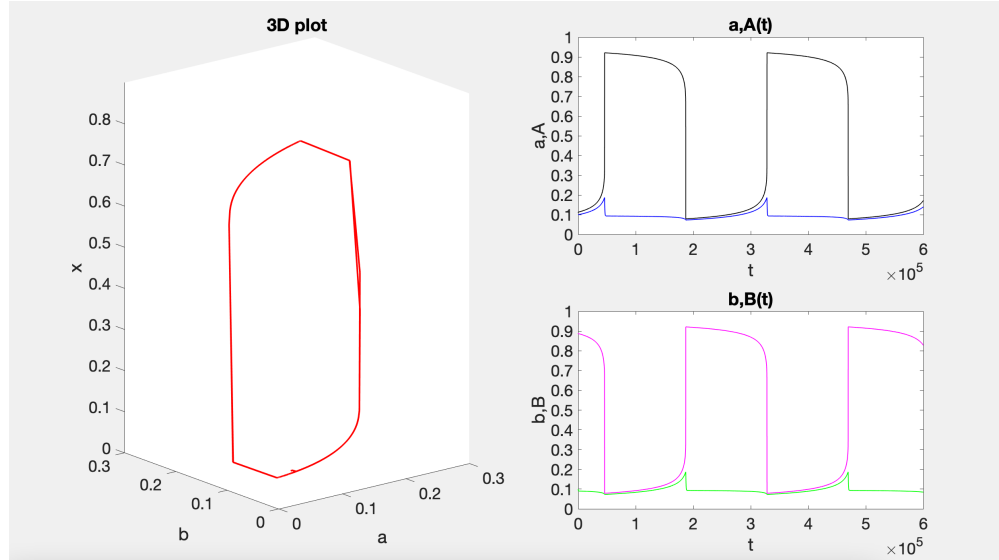


Figure 4.20: The left plot shows a graph of the periodic solution in the phase space for $s, w = 0.5$, and $\epsilon = 0.1$. From the shown perspective, the solution traverses the curve in a counterclockwise fashion. Note how long it takes the solution to traverse x_B and x_T . The right upper plot shows a graph of a (blue) and all opinion A (black) over time. The right lower plot shows b (green) and all opinion B (pink) over time. Note that a precursor for an opinion winning is that the proportion of zealots must rise to a critical level.

bifurcation. The lower right figure has the same features as the upper right, except now we are looking at the evolution for opinion B. In both plots, note the time scales. When $\epsilon = 0.1$, the zealots evolve over a time scale of 1.2×10^5 , whereas a significant change in the total proportion for an opinion changes over $\mathcal{O}(1)$ time scales. The disparity is fueled not only by the singular nature of the ODE, but also the fact that the attracting x -nullclines are also close to the a - and b -nullclines.

The periodic solution requires that when x_B is hit, a decreases, and when x_T is hit, b decreases. In other words, the requirement is that once an opinion becomes dominant, the proportion of zealots for that opinion begins to decrease. The bottom right figures in [Figure 4.17](#) and [Figure 4.18](#) present the scenario when the requirement no longer holds, i.e., $a_{SN} > a_u$ and/or $b_{SN} > b_u$. In this case, the solution will go towards the edge of the x -nullcline,

- $a + b = 1$, all zealots, or
- $a + x = 1$, all A, or
- $a + x = 0$, all B.

Let us more closely consider the physical implications of the behavior of the periodic solution. Without loss of generality, suppose the initial condition is on x_B . The solution will remain on x_B with constant a until it hits the b -nullcline (see the left plot in [Figure 3.10](#)). Along this portion of the curve, opinion B is winning, but the proportion of zealots for B is decreasing. Once the b -nullcline is hit, the proportion of zealots for B is fixed, but the proportion of zealots for A begins to increase. When $a = a_{SN}$, there is a quick transition to x_T . In other words, once a critical proportion of the population become zealots for A, that opinion quickly becomes the dominant opinion. Once the solution is on x_T , the proportion of zealots for A slowly decreases, while the proportion of zealots for B remains fixed (see the right plot in [Figure 3.10](#)). However, opinion A is still winning. Once the a -nullcline is hit, the proportion of zealots for A is fixed, but those for B begin to increase. When $b = b_{SN}$, there is a quick drop to x_B . In other words, once a critical proportion of the population become zealots for B, that again becomes the dominant opinion. Upon dropping to x_B , the cycle begins anew.

5. CONCLUSION

We have created and analyzed a nonlinear ODE model to explore how proportions of a population who hold an opinion with a binary choice change over time. A relatively novel feature of the model is that we allow both moderates and zealots to evolve, but with the allowance that the rate of change for moderates is much greater than that for zealots. Depending on the affinity the total population has towards an opinion, and the rate of indoctrination and deradicalization, there are several possible final configurations:

- one opinion wins, but the losing opinion still has adherents
- the entire population holds one opinion only
- the population becomes filled with zealots (even though there are initially very few zealots)
- the two opinions coexist, but the winning opinion periodically changes.

The existence of a periodic solution is especially intriguing, as, to our knowledge, such a thing has not been seen in previous studies on opinion dynamics where the population is assumed to be filled with only agreeable people.

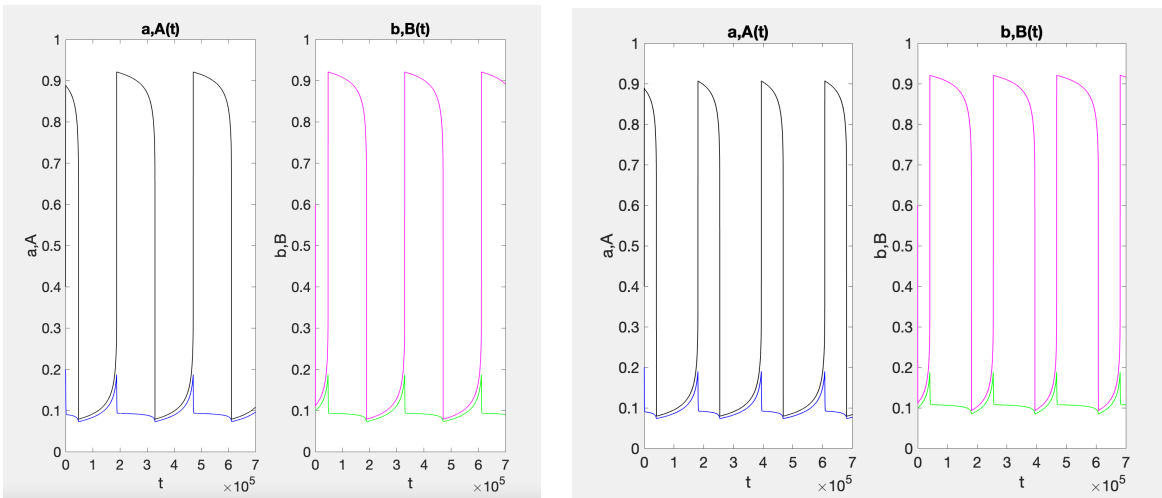


Figure 5.21: Graphs showing a periodic solution for differing values of w for opinion B. In both plots, $w = 0.5$ for opinion A. The left plot has $w = 0.5$ for B, and the right plot has $w = 0.55$ for B.

There is good evidence to show that there are real world scenarios that could easily fit this model, particularly the periodic solution curve where we have oscillatory behavior between opinions. For example, Hayward's [8] paper on church growth models revivals like The Alpha Course or The Toronto Blessing, during which the number of converts grew very quickly due to a group of zealous enthusiasts promoting their cause. However, these enthusiasts only retain their conversion potential for a fixed amount of time, and so eventually the amount of believers will decline again rather sharply, reflective of the periodic solution we discovered in this paper. Similar phenomena occur for diseases or spread of beliefs other than religion [7], though data has not been collected to explicitly prove this.

The work here can be extended in several possible directions. Some of these are:

- extend the model by allowing for a subgroup that does not hold either opinion (see [3, 10] and the references therein)
- extend the model to a network (see [9] and the references therein) to see how zealots in one node effect the entire network, and how the network influences the dynamics of the zealots

- allow for more than two choices for an opinion (see [6])
- allow zealots to be either solely stubborn, or also activists (see [11])
- allow ϵ to take on a larger value
- allow for the indoctrination strategies of the zealots to have different effectiveness.

Concerning the penultimate bullet point, there is preliminary evidence that the periodic solution persists at least for ϵ up to 0.65.

Regarding the implications of the last bullet point, consider Figure 5.21. Here, two periodic solutions are plotted. In the left plot, the indoctrination rates are the same for both opinions, $s = w = 0.5$. In the right plot, $s = 0.5$, and $w = 0.5$ for opinion A, but $w = 0.55$ for opinion B. In other words, the indoctrination rate for opinion B is slightly larger than that for opinion A. The solution is still periodic, but now B wins for longer stretches of time than A. As the rate constant between the two strategies widens further, we expect the periodic solution to be destroyed, and B to always win.

Acknowledgments: This work was supported by the Jack and Lois Kuipers Applied Mathematics Endowment.

REFERENCES

- [1] D. Abrams and S. Strogatz. Modelling the dynamics of language death. *Nature*, 424:900, 2003.
- [2] D. Abrams, H. Yapel, and R. Wiener. Dynamics of social group competition: modeling the decline of religious affiliation. *Phys. Rev. Lett.*, 107:088701, 2011.
- [3] J. Bujalski, T. Kapitula, G. Dwyer, Q-N. Le, H. Malvai, J. Rosenthal-Kay, and J. Ruiter. Consensus and clustering in opinion formation on networks. *Phil. Trans. R. Soc. A.*, 376(20170186):1–21, 2017. URL <http://dx.doi.org/10.1098/rsta.2017.0186>.
- [4] A. Dhooge, W. Govaerts, and Y. Kuznetsov. Matcont: a MATLAB package for numerical bifurcation analysis of odes. *ACM Trans. Math. Software*, 29(779362):141–164, 2003. doi: doi:10.1145/779359.
- [5] K. Eekhoff. Opinion formation dynamics with contrarians and zealots. *SIAM J. Undergraduate Research Online*, 12, 2019.
- [6] R. Fujie, K. Aihara, and N. Masuda. A model of competition among more than two languages. *J. Stat. Phys.*, 151: 289–303, 2013.
- [7] J. Hayward. Mathematical modelling of church growth. *The Journal of Mathematical Sociology*, 23:255–292, 1995.
- [8] J. Hayward. A dynamic model of church growth and its application to contemporary revivals. *Review of Religious Research*, 43:3:218–241, 2002.
- [9] T. Kapitula and P. Kevrekidis. Language competition on lattices. *Studies in Appl. Math.*, 148(1):219–247, 2022.
- [10] S. Marvel, H. Hong, A. Papush, and S. Strogatz. Encouraging moderation: Clues from a simple model of ideological conflict. *American Physical Society*, 109(118702):1–4, 2012. doi: 10.1103/PhysRevLett.109.118702.
- [11] D. Mistry, Q. Zhang, N. Perra, and A. Baronchelli. Committed activists and the reshaping of status-quo social consensus. *Physical Review*, 2015.
- [12] M. Short, S. McCalla, and M. D’Orsogna. Modelling radicalization: how small violent fringe sects develop into large indoctrinated societies. *Royal Society Open Science*, 4:1–21, 2017. URL <http://dx.doi.org/10.1098/rsos.170678>.