Results on the Hegselmann-Krause Model in Opinion Dynamics

Edvin Wedin
Department of Mathematical Sciences
Gothenburg University
SE-412 96 Gothenburg, Sweden

Abstract

This master thesis consists of an introductory text to the so called Hegselmann-Krause bounded confidence model, a well known simple model within the field of opinion dynamics, along with three original papers. The model emulates a discussion among a group of idealised people, or agents, who are assumed to be willing to compromise, but only with those whose opinions are sufficiently close to their own. As it turns out, the interactions between the agents give rise to very complex behaviours on a population level, and the purpose of our work has been to understand these behaviours mathematically. The introduction presents the model and some of its properties, discusses some open problems and then summarises and discusses the papers.

In the first paper we prove a conjecture by Hendrickx et al., that consensus is not guaranteed even for a continuum of agents and a regular opinion function. The continuous agent model is meant to capture the behaviour of the traditional model as the number of agents grows to infinity.

In the second paper we investigate what happens when agents hold one dimensional opinions which are equidistant on an interval. In particular, we exactly determine the freezing time of a system where the initial distances between the agents equal the radius of confidence. The behaviour of such configurations was previously observed in simulations by Krause, and presumably many others, but no formal proof was known.

In the third paper we present a set of configurations on the line with a freezing time of $\Theta(n^2)$, where n is the number of agents. This is the first known example of freezing times that exceed $O(n)$.

Keywords: Hegselmann-Krause Model, Opinion dynamics, Multi-Agent Systems

Contents

List of appended papers

The following three papers are included:

Paper I: **Edvin Wedin** and Peter Hegarty, "The Hegselmann-Krause dynamics for continuous agents and a regular opinion function do not always lead to consensus". Submitted to IEEE Transactions on Automatic Control. Preprint at `http://arxiv.org/abs/1402.7184`

Paper II: Peter Hegarty and **Edvin Wedin**, "The Hegselmann-Krause dynamics for equally spaced agents". Preprint at `http://arxiv.org/abs/1406.0819`

Paper III: **Edvin Wedin** and Peter Hegarty, "A quadratic lower bound for the convergence rate in the one-dimensional Hegselmann-Krause bounded confidence dynamics". To appear in Discrete and Computational Geometry. Preprint at `http://arxiv.org/abs/1406.0769`

Preprints of all the appended papers are also available at `http://publications.lib.chalmers.se/rweb/?personID=220822`.

Acknowledgements

First and foremost, my deepest gratitude goes to my advisor, Peter Hegarty. Day or night, any day of the year, he has been a constant support throughout the work on this thesis. Along the way he has also played an enormous part in my general development towards becoming a mathematician. He has taught me about journal submissions, what to do when some German seems close to getting that same idea that you have been working on, how to address a professor in a professional correspondence, and many other things, large and small. Most importantly, he has taught me something about real mathematical rigour, a lesson for which I am forever grateful.

I would like to thank the students on all levels here at the institution for many fun and helpful discussions. A special thanks goes to Tom for all the language support. I also want to thank my family for always being there for me and encouraging my choices.

Finally, I wish to thank Sandra for her help and support; her love and friendship; and her patience. You are the reason I came to Göteborg and the reason I've stayed. Thank you.

Chapter 1

Introduction

The subject of this thesis is what has become known as the *Hegselmann-Krause bounded confidence model* (or HK-model for short)[1] , also known as the *Hegselmann-Krause dynamics*. It is named after philosopher Rainer Hegselmann and mathematician Ulrich Krause, and was developed to model the opinions of people debating a given topic. It and was first published in 1997 in [7], and popularised in 2002 in [6]. The basic idea is this: Let real numbers represent opinions and assume that people, whom we will refer to as *agents*, find an opinion reasonable if and only if it is close to their own. Then assume that people are willing to compromise with others if and only if they find their opinions reasonable. This is implemented by the model supposing that there is some radius r, commonly taken to be 1, such that everyone is willing to replace their opinion with the average of the opinions within this radius. It further supposes that all the agents update their opinions in this fashion, and that they do so simultaneously in discrete time steps.

1.1.1 Definitions

To put this formally, for n agents let the vector $x_t \in \mathbb{R}^n$ represent their respective opinions at time t. The set of possible opinions, in this case the real numbers, will be referred to as the *opinion space* and the individual vectors will often be referred to as *configurations*. By convention, the agents are labelled 1, 2, ... n and the opinions assumed to be sorted so that they lie in increasing order.[2] For an opinion a in the opinion space and a time t, an agent i such that $x_t(i) = a$ is said to be a *follower* of that opinion. The agents 1 and n are referred to as *extremists* along with all agents that share either of their opinions. For each agent i we define a set $\mathcal{N}_t(i) = \{j : |x_t(j) - x_t(i)| \leq 1\}$ which is referred to as the *neighbours*[3] of i at time t. Note that $j \in \mathcal{N}_t(i) \Leftrightarrow i \in \mathcal{N}_t(j)$. If i and j are neighbours, we often say that they can *see* each other. From an initial vector x_0 of opinions, the system then evolves according to the rule

$$x_{t+1}(i) = \frac{1}{|\mathcal{N}_t(i)|} \sum_{j \in \mathcal{N}_t(i)} x_t(j). \tag{1.1.1}$$

Under this rule we see that any two agents that share the same opinion at time t must also be of the same opinion at time $t+1$ and, by induction, at all subsequent time steps. A maximal set of agents with the same opinion is refereed to as a *cluster*, and the number of agents in a

[1]The abbreviation *HK* for Hegselmann-Krause is also used in other constructions

[2]It is a good exercise to show that the order is preserved under the dynamics defined in (1.1.1).

[3]This choice of nomenclature is admittedly not in line with our view that the numbers $x_t(i)$ are opinions, but still standard. It originates in the idea that, for every t, one may construct a graph where the nodes represent agents, and where the nodes i and j are connected iff $i \in \mathcal{N}_t(j)$.

cluster is referred to as its *size*. Should all agents belong to a single cluster the system is said to have reached *consensus*, and otherwise to result in $fragmentation$. If for some time step $T < \infty$ we get $x_{T+1} = x_T$ we obviously have $x_{T+t} = x_T$ for all $t \in \mathbb{N}$, and we say that the system is then *frozen*. The smallest such T is called the *freezing time* of the system. One of the most basic results concerning the HK-model is that every initial configuration of any finite number of agents must freeze in a finite number of steps, and the freezing time can thus be defined for any configuration of agents. There are various proofs of this, some of which give explicit, albeit maybe not tight, bounds on how large the freezing time can be for a configuration of n agents. For a discussion on the best bounds known to date, see section 1.2.

For some arguments it is useful to consider stationary agents with varying opinions. Other arguments are better conducted by considering agents that move around in the opinion space. For this reason we will often abuse the terminology by referring to agents when we in fact mean their opinions. It should be clear from the context which view is employed.

Over the years numerous variations of this model have been presented. It has been modified to allow agents to have individual thresholds for which opinions they consider, and even asymmetrical thresholds. The model has also been generalised to opinion spaces of higher dimensions, with each agent holding an opinion in $\mathbb{R}^d$ for some $d \geq 1$. Even more exotic opinion spaces like the circle have been studied, see e.g [5]. The work for this thesis has mainly been focused on the classic one dimensional model described above, although Paper I concerns a generalisation to an uncountably infinite number of agents.

1.1.2 Motivation

From a mathematical point of view, the HK-model describes a dynamical system with an evolution rule defined by (1.1.1). If each agent retains the same set of neighbours the evolution is linear and can be described in terms of a matrix A acting on x_t. The challenge in understanding the dynamics of the system lies in the fact that the the sets $\mathcal{N}_t(i)$ depend on x_t, in effect making the evolution rule *piecewise linear*. As the neighbourhoods change discretely it is not, however, continuous. By finding the freezing time T, calculating a matrix A_t for each $t < T$ and multiplying all the matrices $A_0, A_1, \ldots, A_T$ together,[4] one can, in theory, obtain the correct frozen state for any initial condition by multiplying it by a single carefully chosen matrix. Actually finding this matrix for a general initial state has proven very difficult without explicitly calculating all the intermediate matrices A_t one at a time, at each step using the product of the previous matrices to find which neighbourhoods each agent should have. In the resulting product the discontinuities from the individual matrices A_t all contribute to making the system potentially chaotic.

The very simple rule give rise to an extraordinary amount of unexplained phenomena. Many of these arise as one considers parametrised sets of initial configurations and varies the parameters. Though one may observe tendencies, the evolution of the system often depends on the parameters in quite subtle ways. It is, for instance, still an open problem to classify the initial configurations of opinions that will ultimately lead to consensus, and equally open is the question of determining the number of clusters that will emerge when consensus fails. Simulations suggest that if opinions are generated independently at random from any reasonable probability distribution, the number of final clusters, and even their location and relative sizes, is relatively independent of the number of agents, but from a mathematical point of view these are still wide open questions. Another example comes from agents deterministically positioned in periodic configurations: As it seems, this often results in almost, but not entirely, periodic frozen

[4]One could check that if the system is frozen at time T we will have that $A_{T+t} = A_T$ for all $t > 0$, so an infinite product would work just as well.

configurations. There is not much theory to explains this either. Unexpected regularities and irregularities seem to appear for almost any construction on scales ranging from the very local to the very global.

With all this beautiful complexity arising from such a simple definition, the HK-model readily lends itself to mathematical study; it has been analysed by means of functional analysis, Lyapunov functions, linear programing, Markov theory and several other mathematical tools. New ideas are required to further explore the nature of this system, and we hope the interest for its many problems will continue to grow.

There is also a more worldly justification for the study of systems like the HK-model. Despite its origin, and despite us thinking about the values in x as opinions, what makes this model interesting and important to so many non-mathematicians is not necessarily applications to the study of actual opinions of actual people, though this is indeed one motivation. The HK model is a very simple example of a *multi-agent system* with *local rules*. The meaning of this is that even though the new position for an agent does depend on the opinions of other agents, only those that are in some sense near enough at that particular time will count. This is a broad class of problems with many applications in fields such as biology, economics, robotics and the study of social interactions. The agents in these models may be birds in a large flock who only watch their closest neighbours, sick humans who infect those who come too close or autonomous robots exploring unknown territory with limited vision. The rules these agents follow are comparatively complicated, and though they may be easy to simulate to get an idea of their behaviour, knowing anything at all about the dynamics in a mathematical sense is often difficult. Working with a model as simple as the HK-model allows us to go deeper into the theory and better understand the dynamics involved. This follows the well known principle of understanding the basics well in order to help with the intuition of more complicated cases. Even so, very little is known about the HK-model, but we are just beginning to uncover its secrets.

1.1.3 Basic examples

To get some intuition for what the model can do, consider the following examples:

Example 1.1.1. For a very basic first example, consider $x_0 = (0, 1, 2)$. For this vector we have $\mathcal{N}_0(1) = \{1, 2\}$, $\mathcal{N}_0(2) = \{1, 2, 3\}$ and $\mathcal{N}_0(3) = \{2, 3\}$. We calculate the average opinion for each set of neighbours using (1.1.1) and obtain $x_1 = (\frac{1}{2}, 1, \frac{3}{2})$. We now see that $\mathcal{N}_1(1) = \mathcal{N}_1(2) = \mathcal{N}_1(2) = \{1, 2, 3\}$ and make the observation that agents sharing the same set of neighbours must belong to the same cluster in the next time step. We thus have $x_2 = (1, 1, 1)$ and see that the the system is frozen with freezing time 2, and that we have reached consensus.

Example 1.1.2. For a slightly more complicated system, consider $x_0 = (0, \frac{1}{2}, 1, 2, 3)$. We first note that agents 2 and 4 each have two neighbours apart from themselves, and that they both lie exactly centred in between these neighbours. Their opinions will thus remain unchanged until the next step. Agent 1 share the same set of neighbours as agent 2, and they will thus adopt the same opinion. Agent 5 can only see 4, and will thus move to $\frac{5}{2}$. As for agent 3, we have $\mathcal{N}_0(3) = \{1, 2, 3, 4\}$ which yields $\frac{7}{8}$, using (1.1.1). Thus we have $x_1 = (\frac{1}{2}, \frac{1}{2}, \frac{7}{8}, 2, \frac{5}{2})$. We now note that $2 - \frac{7}{8} > 1$, which implies that the system has fragmented and now consists of two sub-systems of agents, each with their respective set of neighbours. The two sub-systems will, according to the observation made in the previous example collapse into clusters in the next step, and we have $x_2 = (\frac{5}{8}, \frac{5}{8}, \frac{5}{8}, \frac{9}{4}, \frac{9}{4})$. Again, the freezing time is 2, but we do not reach consensus.

These examples show, among other things, that adding more agents to a system that reaches consensus might actually cause fragmentation. A fact that might be surprising is that one could then again force the agents in Example 1.1.2 into a consensus by spreading their opinions further apart:

Example 1.1.3. Consider $x_0 = (0, 1, 2, 3, 4)$. Because of symmetry, agent 3 will necessarily retain her opinion for all times t. Note also that, because of this symmetry, it would suffice to consider the lower half of the configuration. Using our experience from the previous examples, we immediately see that only the extremists will move in the first step, and so $x_1 = (\frac{1}{2}, 1, 2, 3, \frac{7}{2})$. In the subsequent steps, agents 2 and 4 will try to stay in equilibrium with their respective extremist and agent 3, and the extremists will chaise after them. Not until $t = 4$ will the extremist see agent 3, to form a cluster with their penultimate agents at $t = 5$. The two clusters will now be a distance $\frac{1250}{1296}$ apart, close enough for the system to reach consensus at $t = 6$.

1.2 What is known?

Very little is known about the nature of the HK-dynamics. For particular constructions it is easy to do the calculations and see where they lead, but for an arbitrary configuration of agents or even a class of configurations there are only a handful of known rigorous results. Arguably the most important of these results is the following:

Theorem 1.2.1. *Let $f_1(n)$ denote the maximal number of steps needed for a configuration of n agents to freeze. Then $f_1(n) = O(n^3)$.*

This theorem was proven independently in both [9] and [1]. The proofs are similar in essence: They both show that if the extremists on one side do not disconnect or increase in number at time t, they must move at least a distance proportional to $\frac{1}{n^2}$ towards the centre, at least every other time step. In [1] this methodology is pronounced, while it is only implicitly used in [9] where the authors instead prove a stronger statement about the total motion of the extremists and an additional penultimate agent. They both continue to argue that if the extremists on either side disconnect from the rest of the agents and thereby freeze, the same lower bound on the movement can be applied to the most extreme non-frozen agent or agents. We can assume that no two opinions differ by more than n, so with the lower bound on the movement of the extremists we see that they can move no more than $O(n^3)$ times. Would they initially differ by more than n some pair of consecutive agents would have opinions differing by more than 1, and the system could then be divided into two subsystems that could be analysed separately.

Though the proofs share similarities they are conducted in markedly different manners and each provides insights about the problem. The partial results used in [9] are slightly stronger, and the coefficient of n^3 is slightly lower than in [1]: 3 instead of 4. The proof in [1] on the other hand is much shorter, half a page as compared to almost four pages.

The problem of finding upper and lower bounds on $f_1(n)$ is further discussed in section 1.4.3.

1.3 Concerning simulations

Simulations play a key part in the study of the HK-model. Most importantly, they have allowed us to discover a wide range of patterns and phenomena that would hardly have been found by hand without considerable effort. Also, the ease with which one could implement a simple simulation and play around with different parameters and initial configurations has doubtlessly contributed to its increased popularity. Although the model was introduced in 1997, it went rather unnoticed until 2002 when [6] was published. This article use computer simulations of the model to demonstrate interesting phenomena and to visualise these in clear and elegant pictures. When this was written [6] had 938 citations on Google Scholar, which is notably more than earlier papers on this topic. Simulations are a also valuable tool for testing hypotheses; before trying to prove a statement, it is very convenient to be able to implement simulations to see whether the statement is plausible or not.

For most of our simulations we have used variations on the following Matlab code:

```
N=50;                              % Choosing the number of agents N.
L=10;                              % Choosing opinion space.
v=sort(L*rand(1,N));               % Generating x0, here uniformly at random.
M=v;                               % Opinions are stored in a matrix M.

t=0;                               % t denotes the current time step.

while 1                            % The loop will run until terminated.
    u=zeros(1,N);                  % Resetting the vector that will become.
                                   % the next v
    for i=1:N
        w=v;         ↓             % A temporary vector w of potential neighbours.
        w(abs(v-v(i))>(1+1e-13))=[];% Finding the neighbours of i by removing
                                   % opinions that differ by more than
                                   % some r>1, see discussion below.
        u(1,i)=mean(w);            % Agent i chooses the mean of neighbouring
                                   % opinions.
    end
    v=u;                           % Updating the vector.
    M(t+1,:)=u;                    % Updating M.
    t=t+1;                         % Moving on to the next time step.

    q=abs(diff(v));                % These lines detects if the system is
    q(q==0)=2;                     % frosen and break the main loop if it is.
    if min(q) > 1
        break
    end
end
plot(0:t,M,'.-')                   % See Figure 1.1
```

All the phenomena discussed in this thesis may be observed by initialising this code with appropriate opinion vectors, or by iterating the code to gather quantitative information. Though not very sophisticated, it manages to simulate the evolution of single random configurations consisting of up to a few tens of thousands of agents in a few minutes. At each time step it goes through all the agents, for each making a copy of the current configuration and then filtering unwanted opinions. It calculates the mean of what is left and stores this in a temporary vector, which then replaces the original vector before the new configuration is stored and the loop repeats. To determine when the configuration has frozen, in which case the simulation should terminate, the code tests whether any two distinct opinions in the configuration differ by more than 1. One can easily check that this is true if and only if the configuration is frozen.

The inequality marked by an arrow (↓) deserves special attention: We want the agents to only consider opinions that differ at most 1 from their own, so the term 10^{-13} seems somewhat out of place. Still, it is absolutely necessary in order to get the results we want, and the reason for this is that the calculations are carried out using *floating point numbers*. This means that the computer, which internally works in binary, needs to round all numbers to the closest binary representation that uses a fixed number of digits.[5] For small integers, halves, quarters etc. this causes no problems, but for any other numbers there will be errors. Any error at all, however minor, could make the computer mistake a neighbour for a non-neighbour, and as if this were

[5]Our code use 60 binary digits.

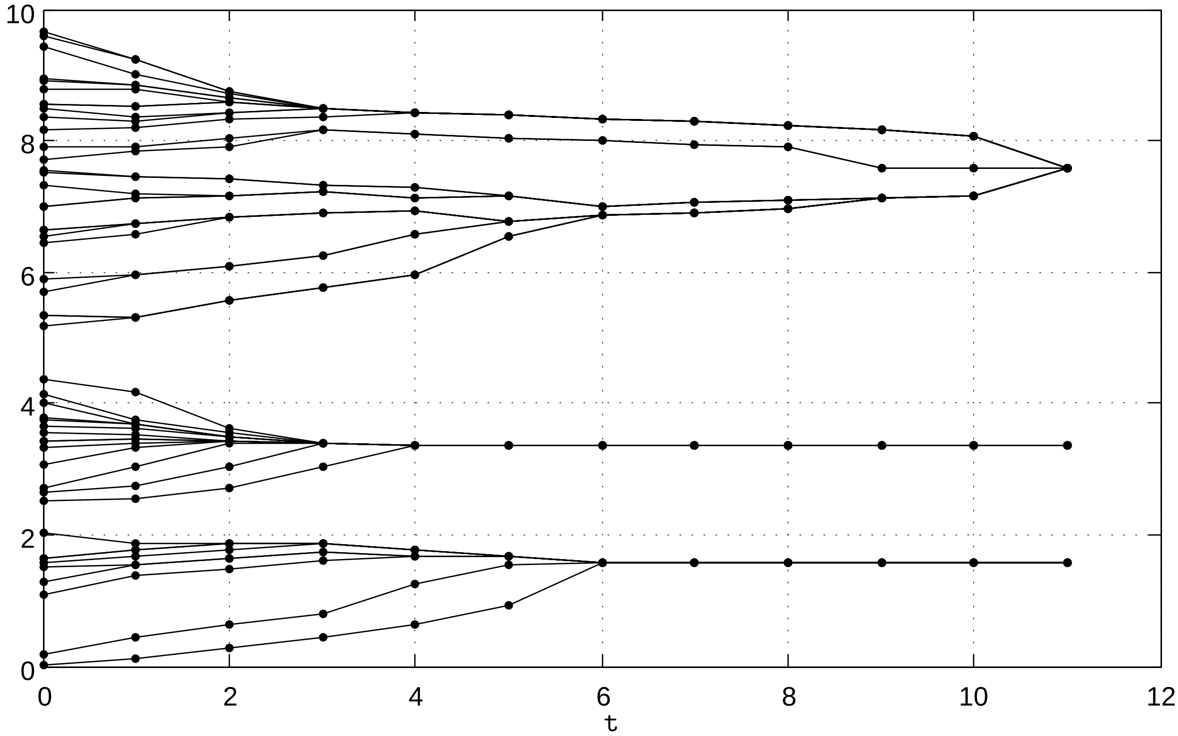

Figure 1.1: A plot generated from the Matlab-code with opinions on the vertical axis and time on the horizontal axis.

not enough these errors may accumulate.

Consider, for example, an initial configuration made up entirely of integer multiples of $\frac{1}{5}$: in the program we set `v=0.2*[0,1,2,3,4,5, ..., k]` for some `k` and observe what happens.[6] Already for `k=6` the first mistake will occur, as the second agent will not recognise the sixth as its neighbour. These errors are initially small and build up slowly, and do not become larger than 10^{-13} until `k=3072`. The addition of the extra term thus protect us from some errors that could potentially drastically alter the evolution some configurations.

A legitimate question would now be if the extra 10^{-13}-term itself could cause mistakes of the other kind, and count a non-neighbour as a neighbour. The answer is of course yes, but this has turned out to be a less severe problem in the cases we have considered.

For deterministic configurations, this is due to the fact that if opinions are assumed to be rational numbers, and we denote the biggest denominator at time t by D_t, we easily have $D_{t+1} \leq \frac{D_t^2}{\max_i \mathcal{N}_t(i)|}$ for all times t. Since numbers whose difference from one is less than 10^{-13} necessarily have large denominators, for as long as $D_t \leq 10^{13}$ we are thus protected from errors of this second kind. This protection lasts only for a very small number of steps, but one could argue that, intuitively, this should be enough in most cases: After a few time steps, the opinions have affected each other in complicated ways and are difficult to guess without calculating them, and we thus think of them as random variables. The "probability" of an error occurring now depends on how densely the opinions are packed. If the density is not small enough for the probability of errors to be low, the sets of neighbours must also be large, and adding or subtracting one agent will not drastically change the evolution. As the updates are linear, one could, in theory, reason in a similar way for subsequent time steps as well, although this leads to messy calculations.

As for random configurations, one could of course apply the same argument as above.

[6] The correct Matlab notation for this is `v=0.2*(0:k)`.

To summarise, we believe our program to be rather robust against errors due to the rounding of the the floating point numbers. Since our research has mostly been focused on proving results using exact arithmetic, with simulations as a mere supplement, we have not put much effort into supporting these beliefs with more formal results. The constant 10^{-13} was chosen more or less arbitrarily to keep both of the errors discussed rare enough for our needs.

1.4 Some open questions

Since very little is known about the HK-dynamics, as has been mentioned earlier, this section could be made quite extensive. Among the mysteries of this model, however, there are a few well known problems that have remained, despite many efforts. In this section, we will discuss a few of them.

In simulations, it is common to construct an initial configuration by assigning independent random opinions to the agents from some probability distribution. This makes for interesting simulations, as quite different behaviours may be observed from what may at a first look like similar sets of agents. On the one hand, this can give a good idea of what "typically" happens to random configurations generated from a certain distribution. After a few hundred simulations one might *feel* rather confident about their behaviour, although in a rather informal, intuitive way. On the other hand, one gets very little or no information about what *may* happen. There might be rare classes of configurations which behave very differently from what is normally observed. Thus, the frustrating truth is that these simulations show phenomena that we often cannot explain, while they withhold the most interesting special cases.

In this section we will discuss three such open questions. The first two of these are strongly related and concern explaining observed phenomena, while the third one concerns the existence of special cases.

1.4.1 Cut-off for consensus with uniform distribution

As we have seen, the rules of the HK-dynamics do *not* guarantee consensus. In fact, if we want opinions to be spread over an arbitrary interval it is non-trivial to construct any configuration of agents that reach consensus at all.[7] Simply choosing independent uniformly random opinions on an interval typically results in fragmentation unless the interval in question is very small: With opinions contained in the interval [0,1] one step will give immediate consensus, and one can easily see that a large number of opinions spread evenly over [0,2] will coincide after 2 or maybe 3 iterations. On the other hand, simulations suggest that opinions drawn from a uniform distribution over [0,10] will practically never reach consensus. But what about the interval [0,5]? If small intervals behave in one way, and bigger ones in another there needs to be some bifurcation.

Assuming all "short" intervals behave in one way, and all "long" intervals in another, which is suggested by simulations, this reasoning can be made formal as follows:

Conjecture 1.4.1. *Let n agents have opinions drawn independently and uniformly at random from the interval $[0,L]$, and let $p_{L,n}$ be the probability that the system converges to consensus. Then there exists a number L_c such that*

$$p_{L,n} \to \begin{cases} 1 & \text{if } L < L_c \\ 0 & \text{if } L > L_c \end{cases} \tag{1.4.1}$$

as $n \to \infty$.

[7] Hint: the only way we have found requires an extremely uneven distribution of an extreme number of agents.

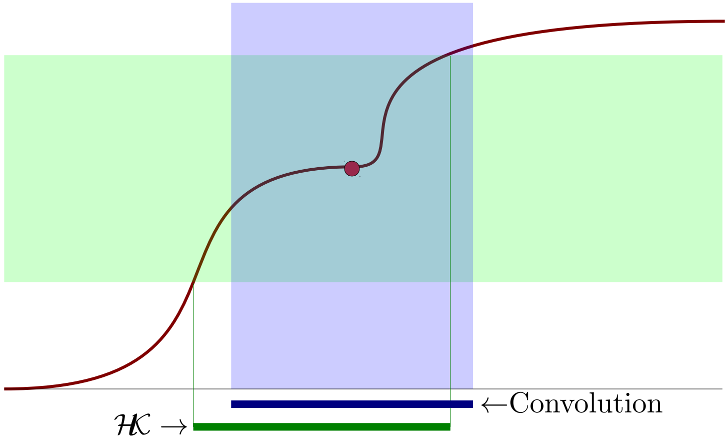

Figure 1.2: The different intervals of integration for the standard convolution with a box and the $\mathcal{HK}$-operator.

This conjecture could conceivably be approached in several different ways. One such way would be to assume that the agents are so densely packed that they can be approximated by an agent continuum, as is done in [2]. In this generalised model, for a given time step t a configuration can be represented by a function x_t from the real interval $[0, 1]$ that indexes the agents, and into the opinion space. We will not get into too much detail here, but if considered as a limit of an increasing finite number of agents generated independently from some probability distribution, the initial function x_0 turns out to be precisely the inverse density function of that distribution.[8] In the case of uniform distribution, this translates to x_0 being a linear function from agents to opinions.

As in the discrete agent model, we obtain the next state by letting the agents replace their own opinion with the average of those nearby. Since we are now dealing with ordinary real valued functions in one variable, one might be tempted to erroneously interpret this as simply convolving x_t with a box function. If this interpretation would be correct any function would become gradually flattened out and in the end approach a constant function, which corresponds to consensus in the continuous model. However, in order to then update the opinion of an agent α using normal convolution, one would have to integrate the function x_t over the interval $[\alpha - 1, \alpha + 1]$. In contrast, the correct way to make an update analogous to (1.1.1) is instead to evaluate x_t at α and then integrate over the inverse image $x_t^{-1}([x_t(\alpha) - 1, x_t(\alpha) + 1])$. The result must then be divided by the length of the interval that constitutes this inverse image. The rule for updating the function can thus be considered a form of "skewed convolution" with a box function. Figure 1.2 tries to illustrate this. If one could better understand this operator, call it $\mathcal{HK}$, and find some manageable way to use it in calculations, one would simply have to apply it repeatedly to a linear function. If this has slope k, say, all that then remains to finish the proof is to vary k to see what happens near $k = 5$ for $\lim_{t\to\infty} \mathcal{HK}^t$.

However, no one has thus far succeeded in this and the problem remains open.

1.4.2 The $2r$-conjecture

In the previous section we mentioned the idea of assigning opinions in some interval independently and uniformly at random to a set of agents. This is a natural first class of configurations

[8] If α is a random variable denoting the opinion, this is defined as the inverse of the function $F(\xi) = \mathbb{P}(\alpha \le \xi)$.

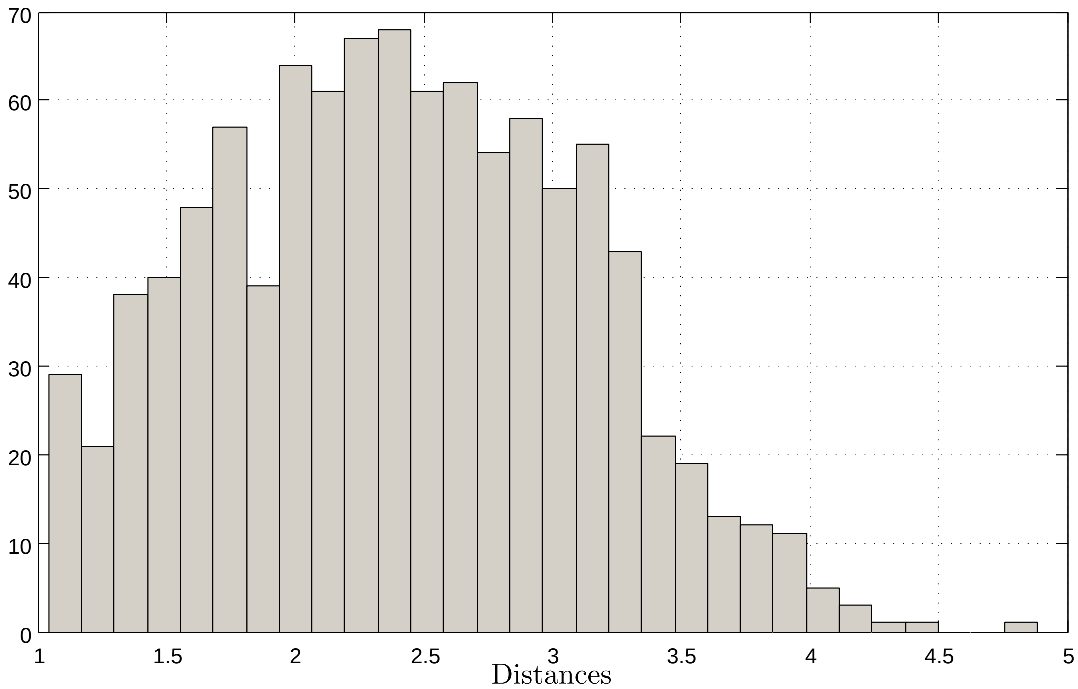

Figure 1.3: A histogram showing the distribution of distances between clusters obtained from 25 simulations, each of 5000 agents equally distributed on the interval $[0, 100]$. The total number of distances is 1003, the sample mean is 2.4 and the sample variance is 0.46.

to try when implementing the HK-model, and it is discussed in countless papers. As was mentioned in the last section, intervals above some length tend to result in opinions ending up fragmented into multiple clusters, and this appears to happen in a somewhat regular manner: It seems like the distances between the clusters tend to stay above $2r$, where r is defined as in Section 1.1.1; hence the name of the phenomenon. It should be noted that this is not one formal conjecture, but rather an unexplained phenomenon about which several possible conjectures could be posited. For this reason we will refer to it as the *$2r$-phenomenon*. It should also be noted that this tendency is merely a tendency, and that distances well below two are frequent even for large numbers of agents. Twenty-five simulations with 10^5 agents equally distributed on $[1, 100]$ give a sample mean of 2.4 and a sample variance of 0.46, see Figure 1.3.

This behaviour is still far from understood. The first step in understanding the system would be to pose an explicit conjecture, and this has partially been done.

In [2] the authors introduce the concept of *stability*. Basically, a frozen configuration of a large number of agents is stable if it cannot be altered radically by the introduction of an additional agent.[9] For example, a system where two clusters that consist of a hundred agents each, holding opinions 0 and $\frac{3}{2}$ respectively, is certainly frozen, but introducing a single agent of opinion $\frac{3}{4}$ will force the system to start evolving again and eventually reach consensus. The system is thus unstable. If we instead let one of the clusters consist of 300 agents a new perturbing agent will be so attracted to the bigger cluster that it immediately moves too far away from the smaller cluster to have any influence over it. In the next step the system will once again freeze. The new frozen configuration will not be much different from the initial one and the initial system is thus considered stable. For a more rigorous description, see [2] or [3].

[9] This is a simplification of the original notion of stability, which uses a weighted version of (1.1.1). It is easy to check, however, that Conjecture 1.4.2 as it is stated here an the original conjecture more or less directly imply each other.

The authors show that if $\#A > \#B$ denotes the sizes of two clusters A and B, a configuration is stable if and only if the distance between any two clusters A and B is greater than $1 + \frac{\#B}{\#A} + O(\frac{1}{\#A})$.[10] They then make the following conjecture:

Conjecture 1.4.2. *(Conjecture 1 in [2]) Fix an interval from which you draw opinions independently according to a fixed continuous probability distribution with connected support. Then the probability of the system evolving to a stable frozen configuration approaches* 1 *as the number of agents goes to infinity.*

Intuitively, if the opinions are drawn uniformly and independently, the final clusters should consist of approximately the same number of agents. If this intuition is true, Conjecture 1.4.2 would imply that the distances between clusters should stay above 2, and thus this would partially explain the $2r$-phenomenon.

1.4.3 The average and maximal freezing times

One of the hottest, most productive, questions concerning the HK-model is the maximal number of steps $f_1(n)$ during which a configuration of n agents may evolve before it freezes. The example provided in Paper III shows that $f_1(n) = \Omega(n^2)$, which still leaves a gap to the best known upper bound $f_1(n) = O(n^3)$.

Both of these bounds are far higher than one might expect from observing random simulations. Let $\bar{f}_1(n,L)$ denote the expected value of the freezing time for n agents with opinions uniformly and independently distributed on $[0, L]$.[11] For up to a few thousand agents $\bar{f}_1$ seems to grow very slowly as the number of agents increases. Determining precisely how slowly is complicated by the fact that the width L of the interval seems to affect the growth rate in a non-linear fashion. This is most easily seen for very small values of L: It is straight forward to see that a system freezes after a single step if $L \leq 1$ and after two steps if $L \in (1, 2)$ asymptotically almost surely as $n \to \infty$. With a little work one can convince oneself that four steps will asymptotically almost surely suffice for any $L \leq 3$, but for larger values of L finding how $\bar{f}_1$ depends on L and n appears highly non-trivial. Figure 1.4 shows simulations that suggest that $\bar{f}_1$ might not even be monotone in L.

A first clue that convergence might require many time steps if n is large comes in the form of *semi-stable states*. By this we mean a configuration containing two large clusters of approximately the same size and with opinions too different for them to see each other, and a small cluster in the middle that can see both of the big ones. These structures appear relatively frequently in random simulations and the freezing time is linear in the number of agents they contain. For example, place the opinions of m agents each on 1 and -1 and let one agent have opinion 0. With only one agent affecting them the extremists will move a distance $\frac{1}{m}$ towards the centre, reducing their difference by a factor $(1 - \frac{2}{m})$. For as long as the distance remains larger than 1 it will reduce by the same factor. Once the extremists see each other the system will reach consensus in a single step. If we denote the freezing time by T we thus have that

$$1 \approx 2(1 - \frac{2}{m})^T \Leftrightarrow 0 \approx \log 2 + T \log(1 - \frac{2}{m}) \approx \log 2 - T\frac{2}{m} \Leftrightarrow T \approx \frac{m \log 2}{2}$$

and we see that the time T depends linearly on m.

The example in the previous paragraph is important, as it introduces the idea that opinions might change very slowly until some time step when suddenly new interactions with appear and the system collapses. We believe that it plays a vital role in explaining the observations in

[10] With the original definition of stability the $O(\frac{1}{\#A})$ term is omitted.

[11] To our knowledge this function has not been previously studied.

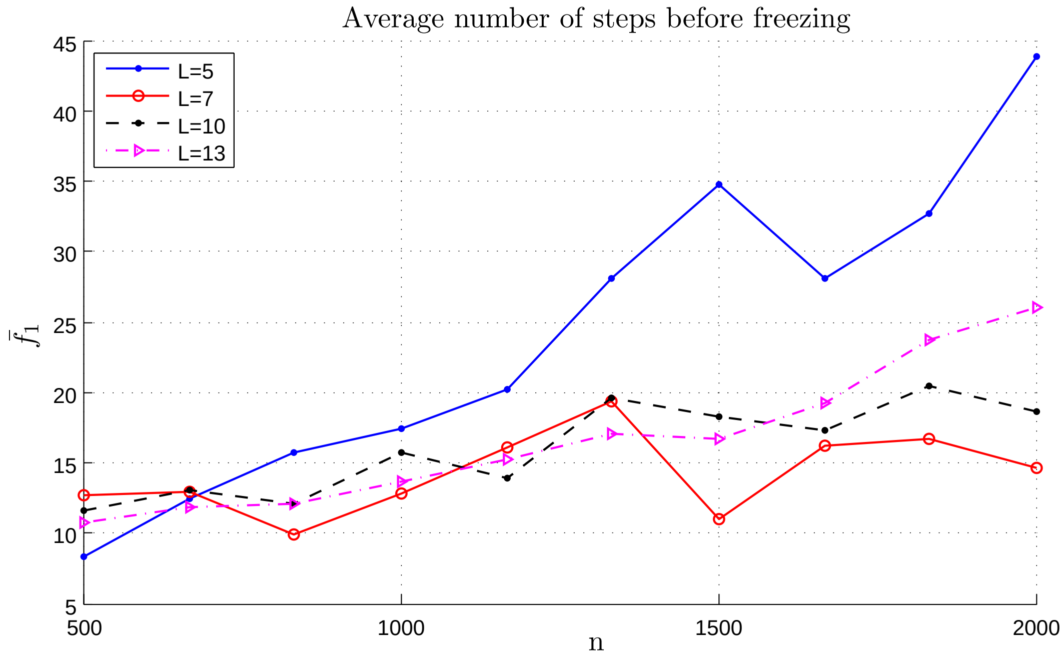

Figure 1.4: Empirical approximations of $\bar{f}_1$ for some values of L and n. Each point is computed from the average of 100 simulations with opinions uniformly distributed on $[0, L]$. Note that for $L = 5$, $\bar{f}_1$ seems to depend nearly linearly on n, while it hardly seems to depend on n at all for $L = 7$. Increasing L further appears to once again increase the correlation between L and $\bar{f}_1$. Note also that the graphs seem to intersect near $n = 650$; passing this point even seems to revert their order, even though this might be due to random noise. There is currently no rigorous theory to explain any of these observations.

Figure 1.4. If enough agents form a semi-stable state it might not collapse for tens or hundreds of steps.[12] Different values of L seem to be more or less prone to give rise to these semi-stable states, and this affects $\bar{f}_n$. The idea is also used in the proofs of both the best known upper and lower bounds on $f_1(n)$ known to date, where slow-moving extremists play a crucial part. In the calculation of these estimates the speed of the extremists is, however, reduced even further to $\Theta(\frac{1}{n^2})$ compared to the just presented $\Theta(\frac{1}{n})$.

The discrepancy between the bounds originates from our lack of understanding of how and why splits occur. While the extremists in the example of Paper III steadily move at a speed of about $\frac{1}{n^2}$ per time step, the distance they travel is bounded by some constant. After moving this distance the extremists absorb all their neighbours, disconnect and form clusters. The most extreme agents in the remaining central component are too few to move that slowly, and their structure will comparatively quickly collapse. If there would have been more agents occupying the opinions close to the original extremists, they could not have sustained their slow movement for long enough, and the time before the disconnections would shorten. Someone might suggest that somehow assigning opinions slightly outside the neighbourhoods of the extremists to more agents could remedy this. One must keep in mind, however, that the freezing time is measured in relation to the number of agents: If we want to arrange any number n of agents so that the extremists retain minimum speed, the number of extremists cannot go below Cn for some absolute constant C. As we want the freezing time to be asymptotically bigger than n^2, the diameter L of the configuration must also grow with n. With only n agents to distribute, this puts a limit on the overall "density" of agents in the opinion space, as it can not grow like $\Theta(n)$.

[12]Indeed, the time is unbounded as n grows.

From these observations we deduce that only some bounded number of opinions, that is the same for all n, can have as many followers as the extremists. For all other opinions the density of followers must asymptotically be much lower, and with these large differences in opinion density splits are hard to avoid. Finding a larger lower bound would require a very clever distribution of the opinions, possibly with extremists moving faster than this slowest possible speed which is used in the current proof, and we doubt it can be done.

As for the upper bound, it does not presuppose any disconnections, and only uses the trivial upper bound $O(n)$ for the distance travelled by an extremist.[13] We believe these conditions can never hold, and that an extremist can either move at close to minimum speed or close to maximal distance, but never both; it seems as though minimum speed of the extremists somehow implies that they will disconnect before moving a distance longer than some universal constant. To prove this, one would supposedly need to formalise the observations from the previous paragraph of why raising the lower bound seems so difficult.

1.5 Summary and discussion of appended papers

1.5.1 Paper I

This paper proves a conjecture by Hendrickx et al. about a continuous version of the HK-model where agents are not indexed by a finite set, but by a real interval. The definition of their model is as follows: Let $x_0 : I \to \mathbb{R}$ be a real-valued monotone function on some interval I, and for every time step t and each $\alpha \in I$ let $\mathcal{N}_0(\alpha) = \{\beta : |x_t(\alpha) - x_t(\beta)| \leq 1\}$ be the set of neighbours of α at time t. Using the notation $|\cdot|$ to denote the length of an interval, we then define x_t for $t > 0$ recursively by the rule

$$x_{t+1}(\alpha) = \frac{1}{|\mathcal{N}_t(\alpha)|} \int_{\mathcal{N}_t} x_t(\beta) d\beta \tag{1.5.1}$$

in complete analogy with (1.1.1). This is meant to represent what happens when the number of agents grows to infinity, and in [3] the authors show that in a certain sense their model achieves this.

One way to get a better intuitive understanding of these dynamics is to, for some t, consider the derivative of the inverse of a function x_t: If $\frac{\mathrm{d}}{\mathrm{d}a}(x_t^{-1})$ is relatively small for opinions a in some interval this means that comparatively few agents hold these opinions. If the derivative is instead relatively large this means that the opinions are popular and will have a large impact on agents holding similar opinions. The function $\frac{\mathrm{d}}{\mathrm{d}a}(x_t^{-1})$ thus describes the density of agents over the opinion space.Figure 1.5 attempts to illustrate this correspondence between the function x_t and the density.[14]

One of the most basic results concerning the discrete model is that the system always freezes in finite time. It is a simple exercise to convince oneself that any discrete configuration can be represented in the continuous model using step functions, so clearly there are at least some configurations in the continuous agent model that freeze. It is also clear that many configurations where all opinions lie within a sufficiently small interval will reach consensus in a finite number of steps and thus freeze as well. A natural question to ask is whether there are in fact any configurations that do *not* freeze.

[13] See section 1.2.

[14] It should be noted that this construction is not intended to be rigorous but only to aid in the informal understanding. Even though all non-constant monotone functions can be assigned an “inverse” by letting discontinuities correspond to constant segments and vice versa, not all monotone functions have derivatives.

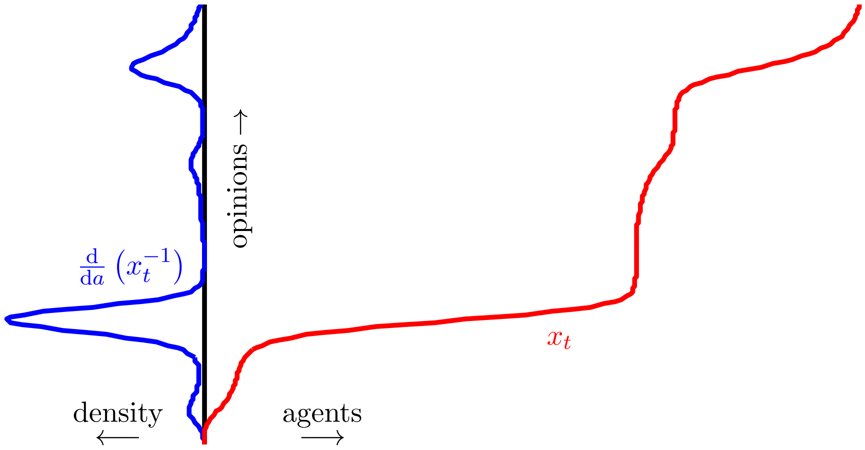

Figure 1.5: The function x_t is shown to the right and the density to the left. Both functions have opinions on the vertical axis. Flat areas in the opinion function correspond to opinions with a high density of agents, while steeper areas correspond less popular opinions.

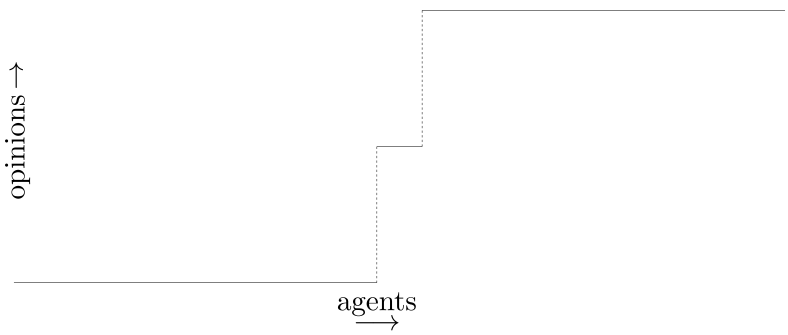

Figure 1.6: The initial opinion function consists of three plateaux, each with a near zero derivative, and two steep areas where the derivative is large. As the system is updated, we show that the plateaux will become flatter and the steep areas steeper.

We denote, in accordance with [2],[15] an opinion function as *regular* if it is everywhere continuous and piecewise continuously differentiable with uniform positive upper and lower bounds on its derivatives. Taking $I = [0, L]$, it is fairly straight-forward to show that x_{t+1} is regular if x_t is regular and $|x_t(L) - x_t(0)| \geq 2$. A regular function cannot be frozen, so to prove that a particular regular function x_0 will never freeze, it is enough to prove that $|x_t(L) - x_t(0)| \geq 2$ remains true for all t. In [3], the authors conjecture the existence of opinion functions that achieve this, and the main theorem of Paper I confirms this conjecture:

Theorem 1.5.1. *There exists a regular function $x_0 : [0, 1] \to \mathbb{R}$ such that, if the sequence $(x_t)_{t \in \mathbb{N}}$ is defined according to (1.5.1), then $x_t(1) - x_t(0) > 2$ for all t.*

The shape of the construction is depicted in Figure 1.6. The opinion function is antisymmetric about its centre and it is initialised very close to the continuous agent counterpart

[15] The original definition is slightly more general.

of a stable equilibrium.[16] The three plateaux seen in the picture correspond to three groups of agents whose opinions are pairwise too different for them to influence each other. The agents of the central small plateau will live in a rather "symmetric neighbourhood" and will not change much. The opinion function being regular, the large plateaux will see some few agents that lie in between themselves and the central plateau, and we show that their opinions are thus pulled towards the centre at a rate proportional to the amount of intermediate agents. We then show that the large plateaux wield enough influence over enough intermediate agents for the agents to become fewer at a rate that is at least exponential. Adding these two facts together tells us that the total movement will not exceed some sum of some exponentially decreasing numbers, and is thus bounded by some absolute constant. We show that this constant is small enough for the plateaux never to see each other, and this completes the proof.

1.5.2 Paper II

Before the results presented in Paper III were obtained, the freezing times of all known configurations of agents were at most directly proportional to the number of agents. To see that there are indeed such configurations, let $\mathcal{E}_n$ denote the configuration where each integer opinion from 1 to n is held by exactly one agent.[17] The initial evolution of this configuration is comparatively simple: In the first step, most of the agents will each have one neighbour on either side, both at distance 1, and thus will not move at all. As for the outermost agents, the extremists, they will have only one neighbour each apart from themselves and thus will move a distance of one half towards the centre. In the second time step, most agents will still "live in a symmetric neighbourhood" and remain unchanged. However, since both the extremists have now changed their position, the two penultimate agents will now move as well. In the third time step, all but three agents on each side will remain unchanged, and so on. By continuing in this way we reach the conclusion that there must be *at least* $\frac{n}{2}$ steps before the agents in the middle are affected, and we see that the time before the system freezes is *at least* proportional to the number of agents.

Though configurations like $\mathcal{E}_n$ have been known for many years (see for example [8]), not much more than what is explained in the previous paragraph was known about them. Several authors, including Ulrich Krause, had made the empirical observation that the evolution seemed to follow a very regular pattern. We make this observation precise and prove that the same behaviour occurs independently of the number of agents.

Theorem 1.5.2. *Let $n \geq 2$ be an integer, and write $n = 6k + l$ where $0 \leq l \leq 5$. Suppose that at $t = 0$ we have the opinion vector $\boldsymbol{\mathcal{E}}_n$ and we let it evolve according to (1.1.1). Then the following occurs:*

(i) after every fifth time step, a group of three agents will disconnect from either end of the receptivity graph and then collapse to a cluster in the subsequent time step.

(ii) the final, frozen configuration, will consist of $2k$ clusters of size 3 with opinions distributed symmetrically about $\frac{n+1}{2}$ plus, if $l > 0$, one cluster of size l with opinion $\frac{n+1}{2}$.

(iii) the configuration will freeze at time $t = 5k + \epsilon(l)$, where

$$\epsilon(l) = \begin{cases} l - 1, & \text{if } l \in \{2, 3\}, \\ l, & \text{if } l = 1, \\ l + 1, & \text{if } l \in \{0,\ 4,\ 5\}. \end{cases} \tag{1.5.2}$$

[16] See section 1.4.2.

[17] The letter $\mathcal{E}$ is for *equidistant*.

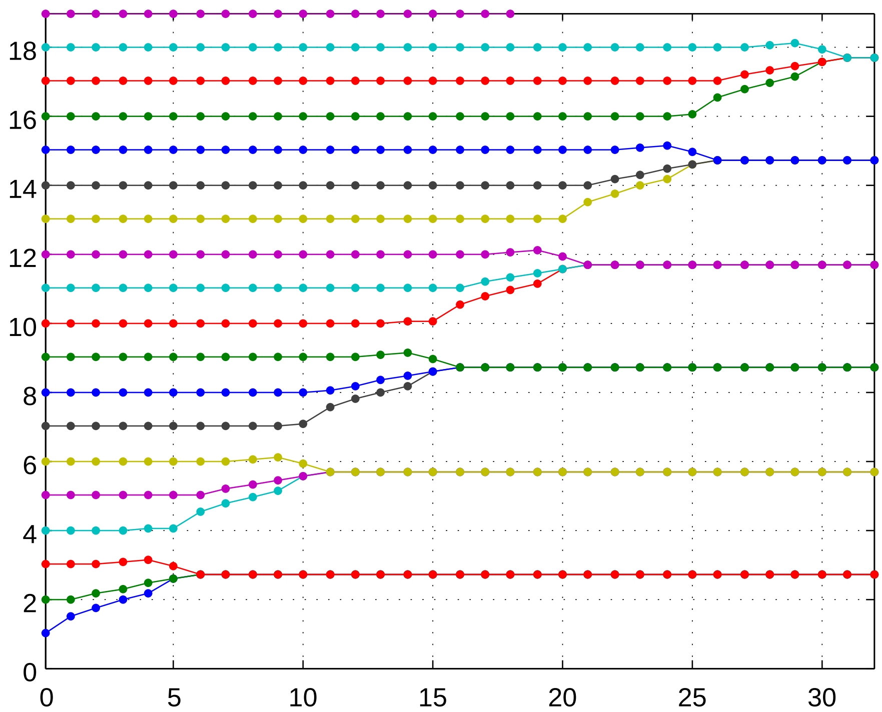

Figure 1.7: The edge of a long chain of equally spaced agents evolves in a periodic pattern. The vertical axis shows opinions while the horizontal axis shows time.

In the first part of Paper II, we prove this theorem. This is done by shifting focus from the opinions themselves to the sequence of distances from each agent to the next. This approach has the advantage that if we look five steps ahead and disregard the first three opinions or distances, the first elements in our new list will look approximately like the old ones, and not like the old ones plus some constant. In fact, the operator that updates the distances and then truncates the ends in this manner is linear. Formally, we assume n to be large, approximate $\mathcal{E}_n$ by a semi-infinite set, denoted $\mathcal{E}_\infty$, and consider what happens at the edge. After showing that the edge of $\mathcal{E}_\infty$ does stay similar to the initial configuration, we then show that the two sides of any finite configuration remain sufficiently similar to the edge of the semi-infinite case for the results to be transferred, reaching the conclusion that triplets of agents will indeed drop off on either side until fewer then six agents remain. Finally, we determine the exact number of steps needed by the remaining agents to reach local consensus, leaving the evolution of configurations $\mathcal{E}_n$ almost completely understood. Since six agents are lost every five steps approximately $\frac{5}{6}n$ steps are needed for $\mathcal{E}_n$ to freeze.

In the second part, we investigate what happens when n agents are placed not at consecutive integers, but at integer multiples of some real number $d \in (0,1]$. We denote such a sequence by $\mathcal{E}_{n,d}$ and note that $\mathcal{E}_{n,1} = \mathcal{E}_n$. A large number of simulations suggest that these configurations evolve in a manner that in some aspects closely resembles that of $\mathcal{E}_n$: After a small number of steps whose precise relation to d remains unclear, the agents who started out with opinions less than about 2.3 away from an extremist will disconnect on each side. This continues, in the sense that a fairly constant number of agents drop off fairly periodically, for as long as there are still agents left in the middle. Though this description is vague it is still the state of the art understanding of what is going on for small values of d. We present some numerical evidence making some of the many open questions a bit more precise, but do not prove any results for

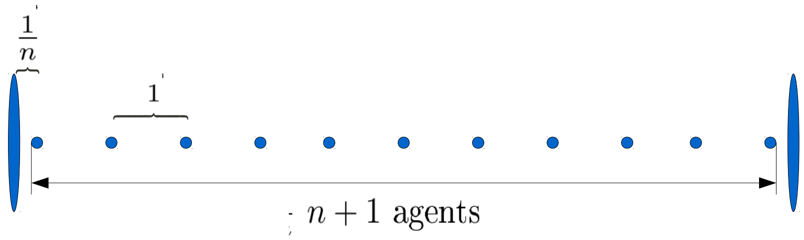

Figure 1.8: The construction resembles a dumbbell. Each cluster at the ends contain n agents each.

$d \leq \frac{1}{2}$.

For all values of d in the interval $(\frac{1}{2},1]$, we completely determine the evolution until the first disconnection. We then show that the same method that was used to prove periodicity in the evolution of $\mathcal{E}_{n,1}$ can be applied to other values of d as well. As an example, we show that for d close to 0.81 four agents on each side will disconnect every eighth time step. This example is important as on average one agent is lost in every step, so for a configuration of this type the freezing time may thus exceed the number of agents. When Paper II was written, this was the longest known freezing time for any configuration of n agents.

1.5.3 Paper III

The question of how many time steps a configuration with n agents can maximally take before it freezes has attracted much interest, and upper bounds for this number have been improved several times. In 2012–2013 [9] and [1] showed independently that the freezing time for any configuration of n agents is at most $O(n^3)$, setting the current record. However, no lower bound, other than the almost trivial linear bound discussed in Paper II, has been known. In this paper we present a configuration $\boldsymbol{\mathcal{D}}_n$ and prove the following theorem:

Theorem 1.5.3. *The configuration $\boldsymbol{\mathcal{D}}_n$ freezes after time $\Theta(n^2)$.*[18]

Our construction is depicted in Figure 1.8 and resembles a dumbbell where each agent or cluster of agents can see no further than the next cluster or agent. It can be constructed by taking a configuration $\mathcal{E}_{n+1}$ and attaching two clusters of n agents a distance of $\frac{1}{n}$ from either end. One can check that all the agents in the $\mathcal{E}_{n+1}$ are initially at equilibrium and will thus not move. The extremists on either side can only see one agent outside their cluster. Since there are so many extremists and the penultimate agents are so close, they will only move a distance of ${}^{\frac{1}{n}}/_{(n+1)} \sim \frac{1}{n^2}$ in the first step. The configuration will at first remain basically unchanged, so for a large number of updates the updates may be described by the same matrix, and the speed of the extremists will remain close to $\frac{1}{n^2}$ in subsequent steps as well.

By trying to stay at the average opinions of their neighbours, the agents in the $\mathcal{E}_{n+1}$ will, beginning near the edges, start moving towards the centre of the configuration. They will act as a "cushion" that distributes the contraction and prevent any one pair of agents from getting too close. We prove that this cushioning is enough to allow the extremists to move at least some

[18]In the article only the lower bound $\Omega(n^2)$ is stated in the theorem. The reason for this is that the proof of the upper bound uses the main result of Paper II, which has not yet been published.

constant distance that is independent of the number of agents in the configuration before they see their second neighbour. This is done by showing that the offset from the initial distances between neighbours is spread along the chain in the same way the distribution of the position of a certain kind of random walker spreads when time increases. As it happens, this spread also satisfies the heat equation, and this can be used to get an intuitive understanding of the proof: Imagine an isolated metal rod initially held at temperature 0. We now start heating it at the ends by adding a certain amount of energy in every time step. The heat will spread along the rod, keeping the temperature near the ends lower than the sum of the temperature that has been introduced. However, the ends will stay warmer than the rest of the rod, and their temperature will continue to increase. Like the heat in the rod, the offset from the initial distances will build up near the ends, until the outermost distances become too short. We show in the paper that this will happen after $\Theta(n^2)$ time steps. The extremists will finally see the penultimate agents in the chain, and these will immediately be pulled towards the edges and thereby disconnect from the rest of the chain. Because of symmetry this happens simultaneously on both sides.

To complete the proof, we note that what is left after the extremists absorb the outermost agents in the $\mathcal{E}_{n+1}$ is similar enough to an $\mathcal{E}_{n-3}$ for the rest of the evolution to proceed as is described in Paper II and terminate in $\sim \frac{5}{6}n$ time steps. Thus the total number of steps is on the order of n^2.

We conjecture that this is the slowest convergence possible for the classic HK-dynamics. See Section 1.4.3 for some further discussion of this problem.

1.6 Author's contributions to the papers

Paper I

The idea for the construction and most of the formulation and writing of the proof, apart from lemma 5 which replaces a much less elegant version of my own. All graphics.

Paper II

The idea to use the inter-agent distances instead of the explicit opinions, and thus reducing the problem to the study of the operator $\mathcal{T}$ of norm 1. All graphics and numerical simulations, including preliminary simulations.

Paper III

The idea for the construction of $\mathcal{D}$ and preliminary simulations.

Bibliography

[1] A. Bhattacharya, M. Braverman, B. Chazelle and H. L. Nguyen, "On the convergence of the Hegselmann-Krause system", *Proceedings of the 4th Innovations in Theoretical Computer Science conference (ICTS 2013)*, Berkeley CA, January 2013.

[2] V. D. Blondel, J. M. Hendrickx and J. N. Tsitsiklis, "On the 2R conjecture for multi-agent systems", *Proceedings of the European Control Conference 2007 (ECC 2007)*, Kos (Greece), July 2007, 874–881.

[3] V. D. Blondel, J. M. Hendrickx and J. N. Tsitsiklis, "On Krause's multi-agent consensus model with state-dependent connectivity", IEEE Trans. Automat. Control, vol. 54, pp. 2586–2597, Nov. 2009.

[4] S. Fortunato, "On the Consensus Threshold for the Opinion Dynamics of Krause-Hegselmann", International Journal of Modern Physics C, vol. 16, No. 2, pp. 259–270, 2005.

[5] P. Hegarty, A. Martinsson and E. Wedin, "The Hegselmann-Krause dynamics on the circle converge". Preprint at `http://arxiv.org/pdf/1410.7330v1.pdf`

[6] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence: models, analysis and simulations", Journal of Artificial Societies and Social Simulation, vol. 5, No. 3, 2002. Fulltext at `http://jasss.soc.surrey.ac.uk/5/3/2/2.pdf`

[7] U. Krause, *Soziale Dynamiken mit vielen Interakteuren, eine Problemskizze,* in: Modellierung und Simulation von Dynamiken mit vielen interagierenden Akteuren, U. Krause and M. Stöckler eds., Universität Bremen (1997).

[8] S. Martinez, F. Bullo, J. Cortes and E. Frazzoli, "On Synchronous Robotic Networks - Part I: Models, Tasks and Complexity", IEEE Trans. Automat. Control, vol. 52, pp. 2214–2226,Dec. 2007.

[9] S. Mohajer and B. Touri, "On convergence rate of scalar Hegselmann-Krause dynamics". Preprint at `http://arxiv.org/pdf/1211.4189v1.pdf`

Chapter 2

Paper I

The Hegselmann-Krause Dynamics for the Continuous-Agent Model and a Regular Opinion Function do not always lead to Consensus

Submitted to IEEE Transactions on Automatic Control. Preprint at `http://arxiv.org/abs/1402.7184`

Abstract

We present an example of a regular opinion function which, as it evolves in accordance with the discrete-time Hegselmann-Krause bounded confidence dynamics, always retains opinions which are separated by more than two. This confirms a conjecture of Blondel, Hendrickx and Tsitsiklis.

2.1 Introduction

There is a rapidly expanding vista for the application of mathematics to multi-agent systems, with applications ranging from engineering to the life and social sciences. One major theme of this effort is *emergence*, the name given to the idea that patterns in the collective behaviour of large groups of interacting agents may be explicable even if each individual is assumed to obey only rules which are both simple and *local*, the latter meaning that each agent is only influenced by its close neighbours, in some appropriate metric.

The field of *opinion dynamics* is concerned with how human agents modify their opinions on social issues as a result of the influence of others. This paper is a contribution to the study of a particularly elegant and well-known mathematical model, the bounded confidence model of Hegselmann and Krause [7], or simply the *HK-model* for brevity. In the simplest formulation of the model, we have a finite number, say N, of agents, indexed by the integers $1,2,\dots,N$. The opinion of agent i is represented by a real number $x(i)$, where the convention is that $x(i) \leq x(j)$ whenever $i \leq j$. The dynamics are as follows: There is a fixed parameter $r > 0$ such that, after each unit of time, every individual replaces their current opinion by the average of those which currently lie within distance r of themselves. This is summarised by the formula

$$x_{t+1}(i) = \frac{1}{|\mathcal{N}_t(i)|} \sum_{j \in \mathcal{N}_t(i)} x_t(j) \tag{2.1.1}$$

where $\mathcal{N}_t(i) = \{j : |x_t(j) - x_t(i)| \leq r\}$. As the dynamics is obviously unaffected by rescaling all opinions and the confidence bound r by a common factor, we can assume without loss of generality that $r = 1$.

Note that the HK-model seems to implicitly assume that each agent is aware of the opinions of all other agents, even if he chooses to ignore most of them when modifying his view. In one sense, this is a matter of interpretation. For example, a conservatively inclined Swedish citizen may switch the channel whenever a member of the Left party[1] is giving an interview, or may keep watching but shake his head and mutter under his breath. In other words, the agent adopts strategies which both filter out unwelcome opinions and prevent him from being aware of them in the first place. On the other hand, the HK-model clearly assumes that an agent is aware of all opinions within his current confidence range. There are no restrictions imposed by, for example, geography, which prevent certain agents from sharing opinions a priori. In other words, agents do not follow local rules in the sense described above, though this is the case when the HK-model is reinterpreted in terms of multi-agent rendezvous [9]. Other important features of the model are that it is fully deterministic and that all agents act simultaneously. Hence the model differs in important respects from other famous models of opinion dynamics such as classical voter models [11] or the Deffuant-Weisbuch model [5].

The update rule (2.1.1) is certainly simple to formulate, though the simplicity is deceptive. Associated to a given configuration $(x(1),\dots,x(N))$ of opinions is a *receptivity graph* G, whose nodes are the N agents and where an edge is placed between agents i and j whenever $|x(i) - x(j)| \leq 1$. In this case, agents i and j are said to be *neighbours*, alternatively that they *see* or *interact with* one another. The transition in the configuration from time t to time $t+1$ is determined by this graph at time t. However, it is clear from (2.1.1) that the dynamics will affect the graph, which in turn affects the dynamics. This feedback is the basic reason why many beautiful conjectures about the HK-model remain unresolved, as we shall now explain.

We begin with the necessary notation and terminology. The state space for a system of N agents obeying the HK-dynamics is the set of non-decreasing functions $x : \{1,2,\dots,N\} \to \mathbb{R}$,

[1] Vänsterpartiet in Swedish.

equivalently, the set of vectors $(x(1), \cdots, x(N)) \in \mathbb{R}^N$ such that $x(i) \leq x(j)$ whenever $i \leq j$. An *equilibrium* state is one such that $|x(i) - x(j)| > 1$ whenever $x(i) \neq x(j)$. Clearly, once an equilibrium state is reached, then the opinion of every agent will be frozen for all future time. It is also easy to see that the converse holds: if $x_{t+1}(i) = x_t(i)$ for all i, then x_t must be an equilibrium state. Any set of agents sharing a common opinion are referred to as a *cluster*. By a slight abuse of terminology, the term "cluster" may refer either to the set of agents with a certain opinion or the real number representing that opinion. Hence, a HK-system is in equilibrium if and only if no two clusters are within unit distance of each other. The simplest kind of equilibrium state is a *consensus*, in which there is only one cluster. Given a cluster $c \in \mathbb{R}$, its weight $w(c)$ is the number of agents sharing opinion c. A *stable equilibrium* is one in which, for any two clusters a and b,

$$|b - a| \geq 1 + \frac{\min\{w(a), w(b)\}}{\max\{w(a), w(b)\}}. \tag{2.1.2}$$

This last notion was introduced in [2], which is the paper that directly inspired the present work. The word "stability" refers to the fact that, if we extend the model to allow non-integer weights and add an agent of sufficiently small weight to an equilibrium configuration satisfying (2.1.2), then when the system is allowed to evolve again the new equilibrium will not differ much from the old one, no matter the opinion of the perturbing agent - see [2] for precise statements and proofs.

The two fundamental facts about the HK-model are the following:
(A) Any initial state will evolve to equilibrium within a finite time.
(B) Even if the receptivity graph is initially connected, the subsequent equilibrium state need not be a consensus.

Fact (A) seems to have been rediscovered several times over and there are a number of different proofs in the literature. Indeed, the same fact has been proven for a wide class of models of which HK is just one particularly simple example, see [4]. Some of the known proofs of (A) give effective bounds for the time taken to reach equilibrium, as a function of the number N of agents only. The best-known bound is $O(N^3)$, which was proven independently in [1] and [10]. It had been speculated that equilibrium is always reached within $O(N)$ steps, and that the worst-case scenario is given by the initial state $\mathcal{E}_N = (1, 2, \ldots, N)$. This is false, however. Recent work of the authors [8], [12] shows that $\mathcal{E}_N$ reaches equilibrium in $5N/6 + O(1)$ steps, whereas there exists a sequence of configurations which takes time $\Theta(N^2)$ to do so. It remains an important open problem in the field to determine the best-possible general upper bound.

Regarding (B), it is easy to see that consensus may not be achieved if the initial distribution of opinions is very uneven. For example, suppose we have 100 agents and the initial state is

$$x_0(i) = \begin{cases} -1, & 1 \leq i \leq 98, \\ 0, & i = 99, \\ +1, & i = 100. \end{cases}$$

At $t = 1$, the opinion of agent 99 will be pulled very close to -1, while agent 100 will only modify his opinion to $x_1(100) = 1/2$. Thus, agent 100 will now be isolated from everyone else and will form a cluster by himself in the equilibrium configuration. What is more interesting is that consensus may not emerge even when there is no such unevenness in the initial configuration. The simplest example is the initial state $\mathcal{E}_6$. A direct computation shows that the resulting equilibrium consists of clusters at $\frac{4613}{1728}$ and $\frac{7483}{1728}$, each of weight three. At this point it seems natural to ask what a "typical" equilibrium state looks like. In order to make this question precise, let us fix a parameter L and suppose that the initial opinions $x_1(0), \ldots, x_N(0)$ are chosen independently and uniformly at random from the interval $[0, L]$. The following two conjectures are supported by overwhelming numerical evidence:

Conjecture 2.1.1. *With opinions chosen initially as just described, let $p_{L,N}$ denote the probability that the resulting equilibrium is a consensus. Then there exists a critical value L_c*[2] *such that, as $N \to \infty$, $p_{L,N} \to 1$ whenever $L < L_c$ and $p_{L,N} \to 0$ whenever $L > L_c$.*

Conjecture 2.1.2. *With opinions chosen initially as described above, let $q_{L,N}$ denote the probability that the resulting equilibrium is stable, in the sense of (2.1.2). Then for any fixed L, $q_{L,N} \to 1$ as $N \to \infty$.*

We have not seen Conjecture 2.1.1 stated explicitly anywhere, though it is implicit in the statements of many different authors. Conjecture 2.1.2 is a special case of Conjecture 1 in [3]. They conjecture that the equilibrium state is almost surely stable under the weaker assumption that the initial opinions are chosen independently from any continuous and bounded pdf on $[0, L]$ with connected support, and not just the uniform distribution. Indeed, it is expected that, when the initial distribution is uniform, then the clusters at equilibrium will typically have about the same weight and hence, if (2.1.2) holds, will typically be separated by at least two. This hypothesis is referred to in the literature as the $2r$ *conjecture*. We are not aware of anyone having turned this hypothesis into a precise conjecture, however. The reason for this is probably that, at least as far as can be told from simulations to date, the distribution of cluster sizes arising from a uniform initial distribution of opinions appears to be quite subtle.

In an attempt to better understand the behaviour of the HK-model for a large number of agents, Blondel, Hendrickx and Tsitsiklis introduced in [2] a *continuous agent* version of the model. Here the uncountably many agents are indexed by numbers in the closed interval $[0, 1]$ and the state space consists of non-decreasing, bounded Lebesgue measurable functions $x : [0, 1] \to \mathbb{R}$. The analogue of (2.1.1) is

$$x_{t+1}(\alpha) = \frac{1}{|\mathcal{N}_t(\alpha)|} \int_{\mathcal{N}_t(\alpha)} x_t(\beta)\, d\beta, \tag{2.1.3}$$

where $\mathcal{N}_t(\alpha) = \{\beta : |x_t(\beta) - x_t(\alpha)| \le 1\}$ and $|\cdot|$ denotes the length of an interval. Note that, if x_t is non-decreasing then $\mathcal{N}_t(\alpha)$ is indeed always an interval, justifying this notation. It is also clear that if x_t is non-decreasing then so is x_{t+1}, so our choice of state-space also makes sense. An *equilibrium state* in this setting is a function attaining only finitely many values, such that the difference between any two such values exceeds one whenever both are attained on sets of positive measure. Stable equilibrium can be defined as in (2.1.2), where now $w(c) = |x^{-1}(c)|$, and the inequality is required to hold whenever both clusters have positive weight. In particular, consensus means a constant function, whereas any equilibrium state which is not a consensus is represented by a discontinuous function. Note that, even if x_t is continuous then x_{t+1} may not be, if x_t is constant on an interval of positive measure. For example, if

$$x_0(\alpha) = \begin{cases} 0, & 0 \le \alpha \le 1/2, \\ 4\alpha - 2, & 1/2 \le \alpha \le 1, \end{cases}$$

then one may check that

$$x_1(\alpha) = \begin{cases} 1/6, & 0 \le \alpha \le 1/2, \\ 2\frac{(\alpha-1/4)^2}{\alpha+1/4}, & 1/2 \le \alpha \le 3/4, \\ 2(\alpha - 1/4), & 3/4 < \alpha \le 1. \end{cases}$$

[2] In [6], simulations are presented which suggest that L_C is close to 5.

There is now a discontinuity at $\alpha = 3/4$, since $\lim_{\alpha \to 3/4^+} x_1(\alpha) = 1 > 1/2 = x_1(3/4)$. This is caused by the fact that, if $\alpha \leq 3/4$ then $x_0(\alpha) \leq 1$ and so $\mathcal{N}_0(\alpha)$ reaches all the way down to zero, whereas if $\alpha > 3/4$ then $\mathcal{N}_0(\alpha) \subset]1/2,\, 1]$.

It is reasonable to restrict attention to initial states x_0 which are injective. Indeed, x_0^{-1} should correspond to the cdf in Conjecture 2.1.2 above and Conjecture 1 of [3]. We shall assume henceforth that the initial state is *regular*, by which we mean that it is almost everywhere C^1, with strictly positive lower and upper bounds on its derivative where it exists. This is a slight strengthening of the notion of regularity as defined in [3].

In contrast to the discrete case, it is not clear which initial states will reach equilibrium in finite time. In [2], it is conjectured that a regular initial state x_0 will converge almost everywhere to a stable equilibrium, that is: there is a stable equilibrium x_∞ such that, for each $\varepsilon > 0$ there is a $T_\varepsilon > 0$ such that $\mu(\{\alpha : |x_t(\alpha) - x_\infty(\alpha)| > \varepsilon\}) < \varepsilon$ for all $t > T_\varepsilon$, where μ denotes Lebesgue measure. They prove a weaker statement in [3], but this fundamental conjecture about the continuous agent model remains open.

It is also proven in Lemma 4 of [2] (see also Proposition 3 of [3]) that if x_0 is regular and if

$$x_t(1) - x_t(0) \geq 2 \;\text{ for all } t \tag{2.1.4}$$

then x_t will also be regular for all t. Hence in such a situation x_t will not reach equilibrium in finite time, nor will it converge to a consensus. This brings us to perhaps the most curious aspect of the continuous agent model, namely it is not obvious that there is any regular initial state which does not reach consensus. The existence of regular initial states whose updates satisfy (2.1.4) was conjectured in [2], but they could give no example with proof. Our contribution here will be to prove this conjecture:

Theorem 2.1.3. *There exists a regular function $x_0 : [0,1] \to \mathbb{R}$ such that, if the sequence $(x_t)_{t\in\mathbb{N}}$ is defined according to (2.1.3), then $x_t(1) - x_t(0) > 2$ for all t.*

Section 2.2 contains a proof of this result and Section 2.3 contains a discussion of some open problems.

2.2 Proof of Main Theorem

The opinion function to be described below will converge pointwise to a non-regular stable state with 3 clusters of positive weight, and the construction can be extended to allow convergence to (at least) any odd number of such clusters.

Since scaling in the agent space I does not affect the dynamics, we will loosen the definition of $I = [0,1]$ and let I be a longer interval. This is done to facilitate some computations at the end of the proof. To construct our initial state x_0, we first partition the set I of agents into successive closed intervals A, B, C, D and E, each intersecting the next in exactly one point. We choose a small positive ε and let these intervals have the lengths $|A| = |E| = 1$, $|B| = |D| = \varepsilon^4$, and $|C| = \varepsilon^2$, so that the endpoints of the intervals lie symmetrically around the centre of C, which we denote by $c = 1+\varepsilon^4+\frac{\varepsilon^2}{2}$. The proof will go through for any sufficiently small ε, but we will fix $\varepsilon = \frac{1}{100}$ which will certainly be small enough. In an analogous manner, we partition the opinion space into closed intervals $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ and $\mathcal{E}$. We will consistently use Roman capitals for intervals in the agent space and script capitals for intervals in the opinion space. We take $|\mathcal{A}| = |\mathcal{C}| = |\mathcal{E}| = \varepsilon$ and $|\mathcal{B}| = |\mathcal{D}| = d = \frac{3}{2}$. The choice of d is somewhat arbitrary, but depends on the choice of ε and must always lie in the open interval $]1, 2[$. To have some co-ordinates to work with, we place the origin at the lower endpoints of the intervals A and $\mathcal{A}$, and we thus have $I = [0, 2c] = [0,\ 2 + \varepsilon^2 + 2\varepsilon^4] = [0,\ 2.00010002]$ and opinions ranging from 0 to $2d + 3\varepsilon = 3.03$.

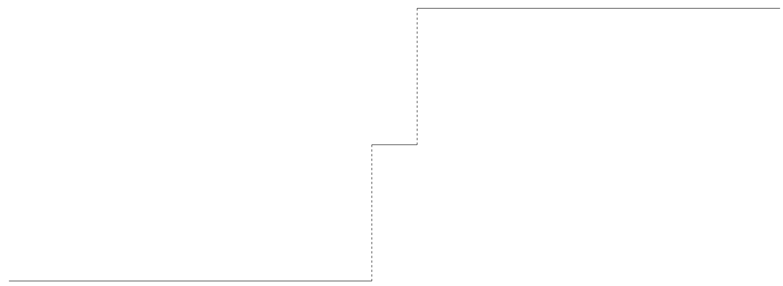

Figure 2.1: The shape to which the opinion functions will converge, which consists of two major clusters, with a smaller cluster in the middle.

We now define x_0 to be linear on each of the subsets A–E, in such a way that $x_0(A) = \mathcal{A}$, $x_0(B) = \mathcal{B}$, $x_0(C) = \mathcal{C}$, $x_0(D) = \mathcal{D}$ and $x_0(E) = \mathcal{E}$. This will force x_0 to stay within the "boxes" illustrated in Fig. 2.2. With this definition, the derivatives of x_0 on A and B will be $x_0'|_A = e_0 = \frac{\varepsilon}{|A|} = \varepsilon$, respectively $x_0'|_B = s_0 = \frac{d}{|B|} = \frac{3d}{2\varepsilon^4}$. The function x_0 is anti-symmetric about c, in particular $x_0'|_E = x_0'|_A$ and $x_0'|_D = x_0'|_B$. It is clear that $x_t(c) = x_0(c)$ for all time steps t, and that the anti-symmetry remains.

We define $A_t = x_t^{-1}([0, 2\varepsilon])$ and $B_t = x_t^{-1}([2\varepsilon, \varepsilon + d])$ to be the sets of agents with opinions in $[0, 2\varepsilon]$ and $[2\varepsilon, \varepsilon+d]$, respectively, at time t, see Fig. 2.3. Note that A is a proper subset of A_0. The reason for using 2ε instead of just ε will become clear in the proof of Lemma 2.2.4 below. We also let $\mathcal{A}_t = x_t(A)$. Again, Roman capitals denote sets of agents while script capitals are reserved for sets of opinions. We will let $\bar{A}_t$ denote the average opinion on A_t at time t.

We also define sequences $(e_t)_{t\geq 0}$ and $(s_t)_{t\geq 0}$: For $t = 0$ we will use the previously defined numbers $e_0 = \varepsilon$ and $s_0 = \frac{d}{\varepsilon^4}$, and then recursively set $e_{t+1} = \frac{2e_t}{s_t}$ and $s_{t+1} = \frac{\varepsilon}{2}\frac{s_t}{e_t}$.

Using the above definitions, we can note that when $t = 0$ the following properties hold:

I: $\mathcal{A}_t \subseteq \mathcal{A}$.

II: $B \supseteq B_t$.

III: $x_t'|_A \leq e_t$.

IV: $x_t'|_{B_t} \geq s_t$.

V: $\varepsilon - \bar{A}_t \geq 4\varepsilon^2$.

Indeed, I–IV hold at $t = 0$ for any $\varepsilon > 0$, and it is easy to check that, for $d = \frac{3}{2}$,

$$\bar{A}_0 \leq \frac{\varepsilon}{2} + \varepsilon^6 \tag{2.2.1}$$

and hence $\varepsilon = \frac{1}{100}$ is enough for V to hold as well.

We will prove by strong induction that properties I–V hold at all time steps. Note that Theorem 2.1.3 will follow immediately, since if property I holds for all t, then the opinions of agents in A will never exceed ε and, by symmetry, the opinions of agents in E will remain above $2d + 2\varepsilon$.

Throughout the proofs below we will use without further comment the fact that if property I holds for all $t \leq T$ then, in particular, $x_t(2c) - x_t(0) \geq 2$ for all $t \leq T$ and hence, by Proposition 3 of [3], the functions x_t remain regular up to time T.

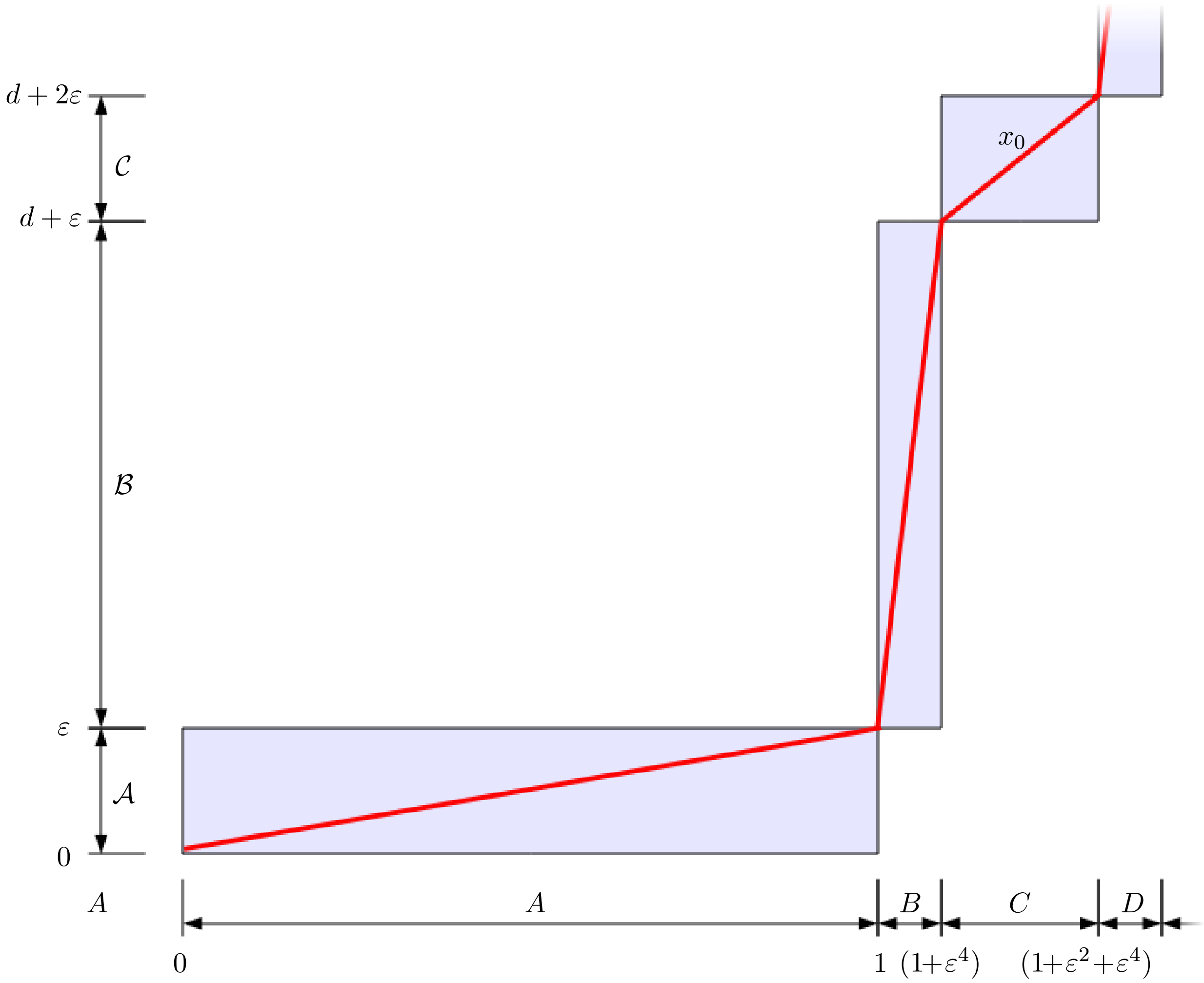

Figure 2.2: Some time invariant subspaces of I and the opinion space for a piecewise linear initial opinion function along with some of the end points. Note that this figure is not to scale.

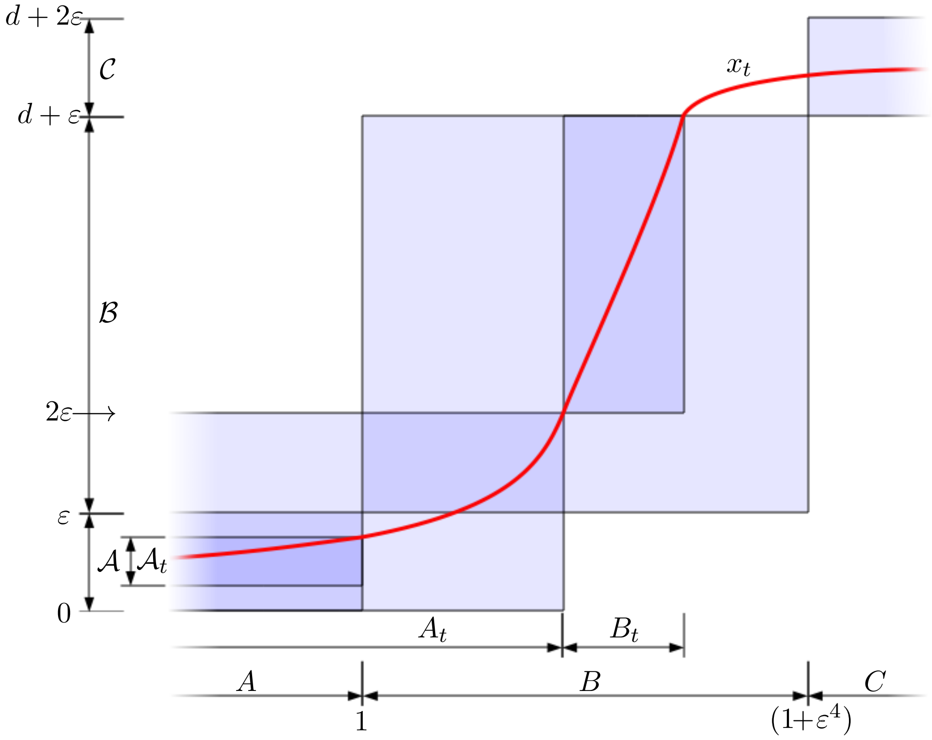

Figure 2.3: Overview of the subsets of the agent space and the opinion space at time t.

Lemma 2.2.1. *Assume properties I, II and V hold at all $t \leq T$. Then*

i) $B_{T+1} \subseteq x_T^{-1}(\mathcal{A}_T + 1)$ *where* $\mathcal{A}_T + 1 = \{\alpha + 1 : \alpha \in \mathcal{A}_T\}$.

ii) $\mathcal{A}_{T+1} \subseteq \mathcal{A}$.

iii) $A_{T+1} \supseteq A_T$.

iv) $B_{T+1} \subseteq B_T$.

Proof. Let

$$\beta_T = x_T^{-1}(\min\ \mathcal{A}_T + 1), \qquad \gamma_T = x_T^{-1}(\max\ \mathcal{A}_T + 1).$$

denote the two extreme agents in $x_T^{-1}(\mathcal{A}_T + 1)$. We will show that

$$x_{T+1}(\beta_T) \leq \varepsilon, \qquad x_{T+1}(\gamma_T) \geq d + \varepsilon \tag{2.2.2}$$

which together with monotonicity is easily checked to imply all four parts of the lemma.

By definition of γ_T, the leftmost agent he can see is the rightmost agent of A. Property I implies that $x_T(\gamma_T) \in [1, 1+\varepsilon]$, and hence the rightmost agent he can see has an opinion between 2 and $2+\varepsilon$ at time T. Because of property II and symmetry and since $d = \frac{3}{2}$ and $\varepsilon < \frac{1}{4}$, no agent in C will have an opinion above 2, so γ_T can see all agents in C. By property I and symmetry the agents in E all have opinions larger than $2d + 2\varepsilon$ at time T, and hence γ_T cannot see any agents in E. Thus $B \cup C \subseteq \mathcal{N}_T(\gamma_T) \subseteq B \cup C \cup D$. The integral of the function

$$\tilde{x}_T(\alpha) = \begin{cases} \frac{3}{2}\varepsilon + d, & \text{if } \alpha \in C, \\ 0, & \text{else} \end{cases}$$

is smaller than or equal to that of x_T over all intervals containing C, since the average opinion on C will always be $\frac{3}{2}\varepsilon + d$ by symmetry. We thus have that

$$x_{T+1}(\gamma_T) \geq \frac{1}{|\mathcal{N}_T(\gamma_T)|} \int_{\mathcal{N}_T(\gamma_T)} \tilde{x}_T(\alpha) d\alpha \geq \frac{|C|(\frac{3}{2}\varepsilon + d)}{|B| + |C| + |D|} = \frac{|C|(\frac{3}{2}\varepsilon + d)}{|C|(2\varepsilon^2 + 1)} \geq \varepsilon + d$$

using $\varepsilon = \frac{1}{100}$ and $d = \frac{3}{2}$.

Next, consider the agent β_T. By definition this agent can see every agent in A. Property I implies again that $x_T(\beta_T) \in [1, 1+\varepsilon]$, and hence the same argument as above implies that β_T can also see all agents in B and C, but no agents in E. The integral of the function

$$\hat{x}_T(\alpha) = \begin{cases} \bar{A}_T, & \text{if } \alpha \in A_T, \\ 2\varepsilon + 2d, & \text{else} \end{cases}$$

over $\mathcal{N}_T(\beta_T)$ is thus greater than or equal to that of x_T over the same set. Since property I implies $A_T \supseteq A$ we can thereby use $\hat{x}_T$ to get the following bound:

$$\begin{aligned} x_{T+1}(\beta_T) &\leq \frac{1}{|\mathcal{N}_T(\beta_T)|} \int_{A \cup B \cup C \cup D} \hat{x}_T(\alpha) d\alpha \leq \frac{|A_T|\bar{A}_T + (2\varepsilon + 2d)|B \cup C \cup D|}{|A_T|} \leq \\ &\leq \bar{A}_T + (2\varepsilon + 2d)(\varepsilon^2 + 2\varepsilon^4) \leq \bar{A}_T + 4\varepsilon^2 \leq \varepsilon \end{aligned} \tag{2.2.3}$$

where the last inequality is true since we assume property V. We have now established the inequalities in (2.2.2) so the proof of the lemma is complete. □

Lemma 2.2.2. *Assume properties I, II and V hold at all $t \leq T$. Then the increase in the mean opinion from A_T at time T to A_{T+1} at time $T+1$ is at most linear in $|B_T|$. More precisely, $\bar{A}_{T+1} - \bar{A}_T \leq 4|B_T|$.*

Proof. Lemma 2.2.1 tells us that $A_{T+1} \supseteq A_T$, and this allows us to write $A_{T+1} = A_T \sqcup (A_{T+1} \setminus A_T)$, a disjoint union of two sets.

As for the first of these sets, recall that at time T agents in A_T have opinions in $[0, 2\varepsilon]$, and agents in B_T have opinions in $[2\varepsilon, \varepsilon + d]$. Hence with $\varepsilon = \frac{1}{100}$ all agents in A_T can see one another, together with some agents in B_T whose opinions are all at most $1 + 2\varepsilon$. Thus the average of x_{T+1} over A_T at time $T+1$ will be

$$\overline{(A_T)}_{T+1} \leq \frac{|A_T|\bar{A}_T + |B_T|(1+2\varepsilon)}{|A_T| + |B_T|} \leq \bar{A}_T + \frac{|B_T|(1+2\varepsilon)}{|A_T|} \leq \bar{A}_T + 2|B_T|$$

where the last inequality uses that $A \subseteq A_T$, which follows from property I.

The average of x_{T+1} over $A_{T+1} \setminus A_T$ at time $T+1$ is certainly at most 2ε, by definition of the set A_{T+1}. It also follows immediately from Lemma 2.2.1(iv) and monotonicity that $A_{T+1} \setminus A_T \subseteq B_T$, and we thereby get the total average

$$\bar{A}_{T+1} \leq \overline{(A_T)}_{T+1} + \frac{|B_T|(2\varepsilon)}{|A_T|} \leq \bar{A}_T + 2|B_T| + |B_T|(2\varepsilon) \leq \bar{A}_T + 4|B_T|.$$

□

Lemma 2.2.3. *Assume properties I-IV hold at all $t \leq T$. Then $x'_{T+1}|_A \leq e_{T+1}$.*

Lemma 2.2.4. *Assume properties I-IV hold at all $t \leq T$. Then $x'_{T+1}|_{B_{T+1}} \geq s_{T+1}$ and $|B_{T+1}| \leq \frac{d}{s_{T+1}}$.*

In the proofs of Lemmas 2.2.3 and 2.2.4 the following additional lemma will be used:

Lemma 2.2.5. *Let x_t be a regular opinion function on I such that $x_t(2c)-x_t(0) > 2$ and define the functions u_t, v_t, $w_t : I \to I$ as follows:*

$$u_t(\alpha) = \begin{cases} 0, & \text{if } x_t(\alpha) \le x_t(0) + 1 \\ x_t^{-1}(x_t(\alpha) - 1), & \text{otherwise} \end{cases}$$

$$v_t(\alpha) = \begin{cases} 2c, & \text{if } x_t(\alpha) \ge x_t(2c) - 1 \\ x_t^{-1}(x_t(\alpha) + 1), & \text{otherwise} \end{cases}$$

$$w_t(\alpha) = v_t(\alpha) - u_t(\alpha)$$

In words $u_t(\alpha)$ and $v_t(\alpha)$ are the leftmost and rightmost agents, respectively, that interact with agent α at time t, and $w(\alpha)$ is the length of the set of neighbours of α at time t.

Then the updated function x_{t+1} is regular with derivative, where it exists, given by

$$x'_{t+1}(\alpha) = \frac{1}{w_t(\alpha)} \left[\, u'_t(\alpha) \cdot (1 + x_{t+1}(\alpha) - x_t(\alpha)) + v'_t(\alpha) \cdot (1 + x_t(\alpha) - x_{t+1}(\alpha)) \,\right].$$

Proof. That x_{t+1} is regular was proven in Proposition 3 of [3]. Assuming x_{t+1}, u_t, v_t and w_t are differentiable at α, we can compute as follows:

We first use the definition in (2.1.3) along with the product rule for derivatives to get

$$x'_{t+1}(\alpha) = \frac{-w'_t(\alpha)}{(w_t(\alpha))^2} \int_{u_t(\alpha)}^{v_t(\alpha)} x_t(\beta) d\beta + \frac{1}{w_t(\alpha)} \frac{d}{d\alpha} \left(\int_{u_t(\alpha)}^{v_t(\alpha)} x_t(\beta) d\beta \right). \tag{2.2.4}$$

The first term simplifies to

$$\frac{[u'_t(\alpha) - v'_t(\alpha)] x_{t+1}(\alpha)}{w_t(\alpha)} \tag{2.2.5}$$

and, by the chain rule, the second term can be rewritten as

$$\frac{1}{w_t(\alpha)} \left[x_t(v_t(\alpha)) \cdot v'_t(\alpha) - x_t(u_t(\alpha)) \cdot u'_t(\alpha) \right]. \tag{2.2.6}$$

But by definition of the functions u_t and v_t, we have

$$x_t(v_t(\alpha)) = \begin{cases} x_t(\alpha) + 1, & \text{if } x_t(\alpha) \le x_t(2c) - 1 \\ 2c, & \text{otherwise} \end{cases}$$

$$x_t(u_t(\alpha)) = \begin{cases} x_t(\alpha) - 1, & \text{if } x_t(\alpha) \ge x_t(0) + 1 \\ 0, & \text{otherwise.} \end{cases}$$

We would like to substitute the values $x_t(v_t(\alpha)) = x_t(\alpha) + 1$ and $x_t(u_t(\alpha)) = x_t(\alpha) - 1$ into (2.2.6). The former doesn't hold when $x_t(\alpha) > x_t(2c) - 1$, but in this range $v_t(\alpha) = 2c$ so $v'_t(\alpha) = 0$, so the substitution can be made in any case. A similar reasoning applies to the latter substitution. Hence the right-hand side of (2.2.6) simplifies to

$$\frac{v'_t(\alpha) \cdot (x_t(\alpha) + 1) - u'_t(\alpha) \cdot (x_t(\alpha) - 1)}{w_t(\alpha)}. \tag{2.2.7}$$

Substituting (2.2.5) and (2.2.7) into (2.2.4) leads after a little computation to (2.2.4). $\square$

Proof of Lemma 2.2.3. By property I, for agents α in A we have that $u_T(\alpha) = 0$, so $u'_T(\alpha) = 0$. Using Lemma 2.2.5 this gives

$$x'_{T+1}(\alpha) = \frac{1 + x_T(\alpha) - x_{T+1}(\alpha)}{w_T(\alpha)}\, v'_T(\alpha) \tag{2.2.8}$$

for all $\alpha \in A$. We also know from property I that, at time T, all agents in A can see each other, so $w_T(\alpha) \geq |A| = 1$.

To get a bound on $v'_T(\alpha)$, we use the definition of v_t, the chain rule, and the formula for the derivative of an inverse function:

$$v'_T(\alpha) = \frac{d}{d\alpha} x_T^{-1}(x_T(\alpha)+1) = x'_T(\alpha) \cdot \frac{1}{x'_T(x_T^{-1}(x_T(\alpha)+1))} = \frac{x'_T(\alpha)}{x'_T(v_T(\alpha))}.$$

To bound this, first note that $\alpha \in A$ implies $x'_T(\alpha) \leq e_T$ by property III. Second, note that since we assume $\mathcal{A}_T \subseteq \mathcal{A}$, it follows that $x_T(v_T(\alpha)) \in [1, 1+\varepsilon]$. In particular, $v_T(\alpha) \in B_T$, and thus $x'_T(v_T(\alpha)) \geq s_T$ by property IV. Putting this together results in the bound

$$v'_T(\alpha) \leq \frac{e_T}{s_T}. \tag{2.2.9}$$

Finally we observe that $1 + x_T(\alpha) - x_{T+1}(\alpha) \leq 2$ holds trivially, and we can now insert this and (2.2.9) into (2.2.8) to obtain

$$x'_{T+1}(\alpha) \leq 2\frac{e_T}{s_T} = e_{T+1}$$

as desired. □

Proof of Lemma 2.2.4. First observe that since both the terms within brackets in (2.2.4) are positive, only one of them is needed to construct a lower bound for the derivative:

$$x'_{T+1}(\alpha) \geq \frac{1}{w_T(\alpha)} u'_T(\alpha)[1 + x_{T+1}(\alpha) - x_T(\alpha)]. \tag{2.2.10}$$

We know from assuming property I and symmetry that no agent in B_T can see as far as E, so $w_T(\alpha) \leq |A \cup B \cup C \cup D| \leq 2|A| = 2$. Lemma 2.2.1(i) and the definition of B_T together assure us that $(1 + x_{T+1}(\alpha) - x_T(\alpha)) \geq \varepsilon$: All the agents in B_{T+1} must have had opinions in $\mathcal{A}_T + 1$ at time T, according to Lemma 2.2.1(i). This is the motivation for using 2ε in the definitions of A_T and B_T. It also lets us use e_T and s_T in a way similar to what was done in the proof of Lemma 2.2.3 to get that $u'_T(\alpha) = \frac{x'_T(\alpha)}{x'_T(u_T(\alpha))} \geq \frac{s_T}{e_T}$. Applying these inequalities to (2.2.10) gives the result.

The upper bound on the size of B_{T+1} simply comes from multiplying the inverse of the bound on the derivative with the height of B_{T+1}, which we know is constantly $d - \varepsilon < d$ by construction. □

Proof of Theorem 2.1.3. We would like properties I–V to hold for all time steps, for then we would be done.

By Lemmas 2.2.1, 2.2.3 and 2.2.4, if properties I-V hold for all $t \leq T$, then I-IV will still hold at time $T+1$. To complete the induction it remains to show that V still holds at time $T+1$.

Lemma 2.2.2 allows us to bound each of the increments $\bar{A}_{t+1} - \bar{A}_t$, yielding

$$\bar{A}_{T+1} = \bar{A}_0 + \sum_{t=0}^{T} \left(\bar{A}_{t+1} - \bar{A}_t\right) \leq \bar{A}_0 + 4\sum_{t=0}^{T} |B_t| \leq \bar{A}_0 + 4\varepsilon^4 + 4\sum_{t=1}^{T} \frac{d}{s_t}, \tag{2.2.11}$$

where the last inequality follows from Lemma 2.2.4 plus the fact that $B_0 \subseteq B$ and $|B| = \varepsilon^4$. We have $s_1 = \frac{d}{2\varepsilon^4}$ and it can easily be checked that the sequence (e_t) is decreasing. Hence, for $t \geq 1$,

$$s_{t+1} = \frac{\varepsilon}{2e_t} s_t \geq \frac{\varepsilon}{2e_1} s_t = \frac{d}{4\varepsilon^5} s_t$$

and hence, by iteration

$$s_t \geq \left(\frac{d}{4\varepsilon^5}\right)^{t-1} s_1$$

Substituting this into (2.2.11) and using (2.2.1) we get

$$\bar{A}_{T+1} \leq \frac{\varepsilon}{2} + \varepsilon^6 + 4\varepsilon^4 + \frac{4d}{s_1} \sum_{t=1}^{T} \left(\frac{4\varepsilon^5}{d}\right)^{t-1} \leq \frac{\varepsilon}{2} + \varepsilon^6 + 4\varepsilon^4 + \frac{8\varepsilon^4}{1 - \left(\frac{4\varepsilon^5}{d}\right)} \leq \varepsilon - 4\varepsilon^2$$

where the last inequality holds for $\varepsilon = \frac{1}{100}$ and $d = \frac{3}{2}$. This completes the proof. □

2.3 Discussion

When thinking about how to construct an example to prove Theorem 2.1.3, we first considered a "single-S" shape, without the narrow plateau in the middle, but with the height of the narrow strip connecting the two tails still being above two. We could not prove that the updates of such an initial state would also satisfy (2.1.4), though we suspect this is the case. In fact, what we think happens when the function is updated is that a narrow plateau will form in the middle, thus yielding the "double-S" shape of the function in Section 2.2 as an intermediate step in the evolution.

In any case, there should be even simpler examples of regular functions which satisfy (2.1.4). Indeed, Conjecture 2.1.1 suggests the following corresponding hypothesis for the continuous agent model:

Conjecture 2.3.1. *Let $x_0 : [0,1] \to \mathbb{R}$ be given by $x_0(\alpha) = L\alpha$. Then there is a critical value L_c^* such that the updates x_t satisfy (2.1.4) whenever $L > L_c^*$, whereas x_0 will evolve to consensus when $L < L_c^*$. Moreover, $L_c^* = L_c$, the critical value in Conjecture 2.1.1.*

In fact, we also conjecture there will not be evolution to consensus at the critical value L_c^*. Intuitively, the reason for this is as follows. For as long as the updates x_t are continuous and non-constant, the ranges $x_t(1) - x_t(0)$ will be strictly decreasing with t. The "$2r$ conjecture" suggests that, given a linear x_0, there will eventually be consensus if and only if the range of opinions shrinks to strictly below two at some point. Hence, at $L = L_c^*$, we should converge almost everywhere to an equilibrium consisting of two clusters of equal measure and separated by exactly two.

This leads in turn to another obvious remaining question, namely whether it is possible for a regular initial state to fail to satisfy (2.1.4) and yet never reach consensus.

Acknowledgment

We thank the referees for their very careful reading of the original manuscript, which led to substantial improvements in the revised version.

Bibliography

[1] A. Bhattacharya, M. Braverman, B. Chazelle and H. L. Nguyen, "On the convergence of the Hegselmann-Krause system", *Proceedings of the 4th Innovations in Theoretical Computer Science conference (ICTS 2013)*, Berkeley CA, January 2013.

[2] V. D. Blondel, J. M. Hendrickx and J. N. Tsitsiklis, "On the 2R conjecture for multi-agent systems", *Proceedings of the European Control Conference 2007 (ECC 2007)*, Kos (Greece), July 2007, 874–881.

[3] V. D. Blondel, J. M. Hendrickx and J. N. Tsitsiklis, "On Krause's multi-agent consensus model with state-dependent connectivity", IEEE Trans. Automat. Control, vol. 54, pp. 2586–2597, Nov. 2009.

[4] B. Chazelle, "The total s-energy of a multiagent system", SIAM J. Control Optim., vol. 49, No. 4, pp. 1680–1706, 2011.

[5] G. Deffuant, D. Neau, F. Amblard and G. Weisbuch, "Mixing beliefs among interacting agents", Adv. Complex Syst., vol. 3, pp. 87–98, 2000.

[6] S. Fortunato, "On the Consensus Threshold for the Opinion Dynamics of Krause-Hegselmann", International Journal of Modern Physics C, vol. 16, No. 2, pp. 259–270, 2005.

[7] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence: models, analysis and simulations", Journal of Artificial Societies and Social Simulation, vol. 5, No. 3, 2002. Fulltext at `http://jasss.soc.surrey.ac.uk/5/3/2/2.pdf`

[8] P. Hegarty and E. Wedin, "The Hegselmann-Krause dynamics for equally spaced agents", Preprint at `http://arxiv.org/abs/1406.0819`

[9] S. Martinez, F. Bullo, J. Cortes and E. Frazzoli, "On Synchronous Robotic Networks - Part I: Models, Tasks and Complexity", IEEE Trans. Automat. Control, vol. 52, pp. 2214–2226, Dec. 2007.

[10] S. Mohajer and B. Touri, "On convergence rate of scalar Hegselmann-Krause dynamics". Preprint at `http://arxiv.org/pdf/1211.4189v1.pdf`

[11] `http://en.wikipedia.org/wiki/Voter_model`

[12] E. Wedin and P. Hegarty, "A quadratic lower bound for the convergence rate in the one-dimensional Hegselmann-Krause bounded confidence dynamics", Preprint at `http://arxiv.org/abs/1406.0769`

Chapter 3

Paper II

The Hegselmann-Krause dynamics for equally spaced agents

Preprint at `http://arxiv.org/abs/1406.0819`

Abstract

We consider the Hegselmann-Krause bounded confidence dynamics for n equally spaced opinions on the real line, with gaps equal to the confidence bound r, which we take to be 1. We prove rigorous results on the evolution of this configuration, which confirm hypotheses previously made based on simulations for small values of n. Namely, for every n, the system evolves as follows: after every 5 time steps, a group of 3 agents become disconnected at either end and collapse to a cluster at the subsequent step. This continues until there are fewer than 6 agents left in the middle, and these finally collapse to a cluster, if n is not a multiple of 6. In particular, the final configuration consists of $2\lfloor n/6 \rfloor$ clusters of size 3, plus one cluster in the middle of size $n \pmod 6$, if n is not a multiple of 6, and the number of time steps before freezing is $5n/6 + O(1)$. We also consider the dynamics for arbitrary, but constant, inter-agent spacings $d \in [0, 1]$ and present three main findings. Firstly we prove that the evolution is periodic also at some other, but not all, values of d, and present numerical evidence that for all d something "close" to periodicity nevertheless holds. Secondly, we exhibit a value of d at which the behaviour is periodic and the time to freezing is $n + O(1)$, hence slower than that for $d = 1$. Thirdly, we present numerical evidence that, as $d \to 0$, the time to freezing may be closer, in order of magnitude, to the diameter $d(n-1)$ of the configuration rather than the number of agents n.

3.1 Introduction

The Hegselmann-Krause (HK) bounded confidence model of opinion dynamics, in its original one-dimensional setting introduced in [3], works as follows. We have a finite number n of agents, indexed by the integers $1, 2, \ldots, n$. Time is measured discretely and the opinion of agent i at time $t \in \mathbb{N} \cup \{0\}$ is represented by a real number $x_t(i)$, where the convention is that $x_t(i) \leq x_t(j)$ whenever $i \leq j$. There is a fixed parameter $r > 0$ such that the dynamics are given by

$$x_{t+1}(i) = \frac{1}{|\mathcal{N}_t(i)|} \sum_{j \in \mathcal{N}_t(i)} x_t(j), \tag{3.1.1}$$

where $\mathcal{N}_t(i) = \{j : |x_t(j) - x_t(i)| \leq r\}$. As the dynamics are obviously unaffected by rescaling all opinions and the confidence bound r by a common factor, we can assume without loss of generality that $r = 1$.

Let $(x(1), \ldots, x(n))$ be a vector of opinions. We say that agents i and j *agree* if $x(i) = x(j)$. A maximal set of agents that agree is called a *cluster*, and the number of agents in a cluster is called its *size*. The configuration $\boldsymbol{x} = (x(1), \ldots, x(n))$ is said to be *frozen* if $|x(i) - x(j)| > 1$ whenever $x(i) \neq x(j)$. It is easy to see that, if $\boldsymbol{x}_t$ and $\boldsymbol{x}_{t+1}$ are related as in (3.1.1), then $\boldsymbol{x}_{t+1} = \boldsymbol{x}_t$ if and only if $\boldsymbol{x}_t$ is frozen. Thus once opinions obeying the HK-dynamics become frozen, they will remain so for all future time.

Perhaps the most fundamental result about the HK-dynamics is that any configuration of opinions will freeze in finite time. There are multiple proofs of this in the literature, but the same fact is true for a wide class of models of which HK is just one particularly simple example, see [2]. More interestingly, the time taken for a configuration to freeze is bounded by a universal function of the number n of agents. Currently, the best upper bound is $O(n^3)$, due to [1]. An important open problem in the field is to find the optimal bound.

It was noted in [1], and even earlier in [6], that one definitely cannot do better than an $O(n)$ bound. For suppose we start from the configuration $\boldsymbol{\mathcal{E}}_n = (1, 2, \ldots, n)$, so opinions are equally spaced with gaps equal to the confidence bound. Then it is not hard to see that, as the configuration updates, if $i < n/2$ then the opinions of agents i and $(n+1) - i$ will remain constant as long as $t < i$, while both will change at $t = i$. Hence, the time it takes for the configuration $\boldsymbol{\mathcal{E}}_n$ to freeze is at least $n/2$.

Intuitively, $\boldsymbol{\mathcal{E}}_n$ seems like a good candidate for a configuration which converges as slowly as possible simply because, at the outset, opinions are placed as far apart as they can be while retaining an unbroken chain of influence. As it turns out, this intuition is badly wrong - in a companion paper [7] we will exhibit configurations of n agents which, as $n \to \infty$, take time $\Omega(n^2)$ to freeze. Configurations of equally spaced agents nevertheless remain interesting for other reasons. Krause [4] has observed, based on simulations for values of n up to 100 or so, that the configuration $\boldsymbol{\mathcal{E}}_n$ seems to evolve in a very regular manner. Our main result confirms, and makes completely precise, Krause's hypotheses. Before stating it, we introduce some graph-theoretic terminology. Let $(x(1), \ldots, x(n))$ be a vector of opinions. We can define a *receptivity graph* G, whose nodes are the n agents and where an edge is placed between agents i and j whenever $|x(i) - x(j)| \leq 1$. We say that agents i and j are *connected* if they are in the same connected component of the receptivity graph. Observe that every connected component of G is an interval of agents and that i is disconnected from $i+1$ if and only if $x(i+1) > x(i) + 1$. We can now state our theorem:

Theorem 3.1.1. *Let $n \geq 2$ be an integer, and write $n = 6k + l$ where $0 \leq l \leq 5$. Suppose that at $t = 0$ we have the opinion vector $\boldsymbol{\mathcal{E}}_n$ and we let it evolve according to (3.1.1). Then the following occurs:*

(i) after every fifth time step, a group of three agents will disconnect from either end of the receptivity graph and then collapse to a cluster in the subsequent time step.

(ii) the final, frozen configuration, will consist of $2k$ clusters of size 3 with opinions distributed symmetrically about $\frac{n+1}{2}$ plus, if $l > 0$, one cluster of size l with opinion $\frac{n+1}{2}$.

(iii) the configuration will freeze at time $t = 5k + \epsilon(l)$, where

$$\epsilon(l) = \begin{cases} l-1, & \text{if } l \in \{2,3\}, \\ l, & \text{if } l = 1, \\ l+1, & \text{if } l \in \{0, 4, 5\}. \end{cases} \tag{3.1.2}$$

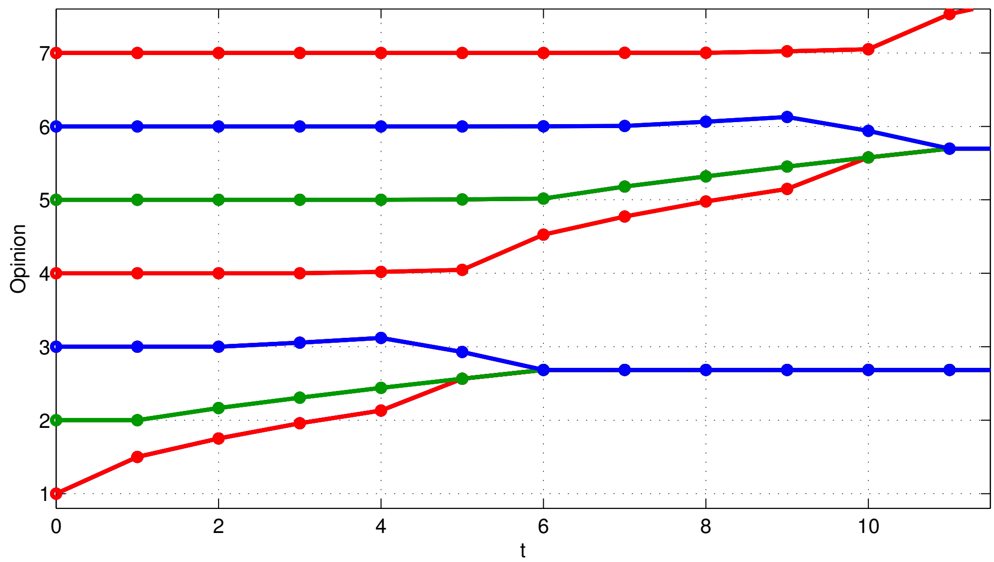

Figure 3.1: After the first 3 agents become disconnected, what is left is very similar to a shorter version of the original chain.

Remark 3.1.2. *The formula for the freezing time for n agents can be written as*

$$T(n) = 1 + 5\lfloor \frac{n+2}{6} \rfloor + \frac{1}{3}\left(\sqrt{3}\sin\left(\frac{2\pi(n-1)}{3}\right) - \cos\left(\frac{\pi(n-1)}{3}\right) - (-1)\right). \tag{3.1.3}$$

This formula is given in [5], but without proof.

Theorem 3.1.1 will be proven in Section 2. In Section 3 we investigate more generally the evolution of a configuration of equally spaced agents, when the inter-agent spacing is an arbitrary number $d \in (0, 1]$. We will show that the evolution is periodic also for some other values of d, though not all, while numerical and heuristic evidence suggests nevertheless that something "close" to periodicity might hold for arbitrary d. We will show that, for a small range of values of d slightly above 0.8, the behaviour is periodic, with groups of four agents disconnecting after every eighth time step, which leads to a freezing time of $n + O(1)$. Thus $\boldsymbol{\mathcal{E}}_n$ does not even converge most slowly amongst equally spaced configurations. We conjecture, however, that any equally spaced configuration of n agents will freeze in time $n + O(1)$, and present numerical evidence suggesting that, as $d \to 0$, the freezing time may be closer, in order of magnitude, to the diameter $d(n-1)$ of the configuration rather than the number n of agents.

3.2 The case of equal spacings $d = 1$

For $n \in \mathbb{N} \cup \{0\}$ set

$$\mathbb{R}^n_* = \{(x(1), \dots, x(n)) \in \mathbb{R}^n : x(i) \le x(j) \text{ whenever } i \le j\},$$
$$\mathbb{R}^n_+ = \{(x(1), \dots, x(n)) \in \mathbb{R}^n : x(i) \ge 0 \text{ for each } i\}.$$

If $n \ge 2$, there is a natural map $\phi : \mathbb{R}^n_* \to \mathbb{R}^{n-1}_+$ given by

$$\phi\,[(x(1), \dots, x(n))] = (x(2) - x(1), \dots, x(n) - x(n-1)).$$

The update rule (3.1.1) can be written in matrix form as

$$\boldsymbol{x}_{t+1} = A_t\, \boldsymbol{x}_t,$$

where $\boldsymbol{x}_t, \boldsymbol{x}_{t+1} \in \mathbb{R}^n_*$ and A_t is a row-stochastic $n \times n$ matrix. Note that the entries of A_t depend only on the receptivity graph G_t at time t. If $n \ge 2$ then, setting $\boldsymbol{y}_t = \phi(\boldsymbol{x}_t)$, we can just as well write the update rule as

$$\boldsymbol{y}_{t+1} = B_t\, \boldsymbol{y}_t,$$

where $\boldsymbol{y}_t, \boldsymbol{y}_{t+1} \in \mathbb{R}^{n-1}_+$ and B_t is an $(n-1) \times (n-1)$ matrix with non-negative entries, though the row sums will no longer equal one in general. As before, the entries of B_t depend only on the graph G_t. We will find it more convenient to work with (3.2) rather than (3.2), in other words to replace a vector $\boldsymbol{x}$ of opinions by a vector $\boldsymbol{y} = \phi(\boldsymbol{x})$ of gaps between opinions. Observe that $\phi(\boldsymbol{\mathcal{E}}_n) = (1, 1, \dots, 1)$, a vector which we denote $\mathbf{1}_{n-1}$. More generally, if the receptivity graph corresponding to $\boldsymbol{x}$ is connected, then the entries of $\phi(\boldsymbol{x})$ are bounded above by one.

To get a feeling for Theorem 3.1.1, we look at $n = 11$ as an example. What is important is what happens during the first five time steps. At $t = 0$ it is obvious that

$$\boldsymbol{y}_0 = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1),$$
$$E(G_0) = \{\{i, i+1\} : 1 \le i \le 10\},$$
$$B_0(i,j) = \begin{cases} 1/6, & \text{if } i = j = 1 \text{ or } i = j = 10, \\ 0, & \text{if } |i - j| > 1, \\ 1/3, & \text{otherwise.} \end{cases}$$

Thus,

$$\boldsymbol{y}_1 = B_0\, \boldsymbol{y}_0 = (0.5, 1, 1, 1, 1, 1, 1, 1, 1, 0.5), \tag{3.2.1}$$

from which it is in turn clear that $G_1 = G_0$ and $B_1 = B_0$. The entries of all subsequent vectors $\boldsymbol{y}_t$ will of course all be rational numbers, but we will write approximations to four decimal places so as to make it easier to keep track of magnitudes. At the next time step we have

$$\boldsymbol{y}_2 = B_1\, \boldsymbol{y}_1 = B_0^2\, \boldsymbol{y}_0 \approx (0.4167, 0.8333, 1, 1, 1, 1, 1, 1, 0.8333, 0.4167). \tag{3.2.2}$$

Thus $G_2 = G_0$ and $B_2 = B_0$ still and so

$$\boldsymbol{y}_3 = B_0^3\, \boldsymbol{y}_0 \approx (0.3472, 0.75, 0.9444, 1, 1, 1, 1, 0.9444, 0.75, 0.3472). \tag{3.2.3}$$

Still $G_3 = G_0$ and so

$$\boldsymbol{y}_4 = B_0^4\, \boldsymbol{y}_0 \approx$$
$$(0.3079, 0.6806, 0.8981, 0.9815, 1, 1, 0.9815, 0.8981, 0.6806, 0.3079). \tag{3.2.4}$$

Now finally something happens. Since $\boldsymbol{y}_4(1) + \boldsymbol{y}_4(2) < 1$, agents 1 and 3 are now connected, and similarly with agents 9 and 11. Thus $E(G_4) = E(G_0) \cup \{\{1, 3\}, \{9, 11\}\}$ which leads to

$$B_4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{12} & \frac{1}{6} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{5}{12} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{5}{12} & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{6} & \frac{1}{12} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.2.5}$$

Hence,

$$\begin{aligned} &\boldsymbol{y}_5 = B_4 \, \boldsymbol{y}_4 \approx \\ &(0, 0.3636, 1.1186, 0.9599, 0.9938, 0.9938, 0.9599, 1.1186, 0.3636, 0). \end{aligned} \tag{3.2.6}$$

Here $\boldsymbol{y}_5(1) = \boldsymbol{y}_5(10) = 0$ exactly, which means that agents 1 and 2 agree, as do agents 10 and 11. Moreover, $\boldsymbol{y}_5(3) = \boldsymbol{y}_5(8) > 1$, which means that agent 3 (resp. 8) has become disconnected from agent 4 (resp. 9). This is in accordance with part (i) of Theorem 3.1.1. It is easy to continue and check that agents 1, 2, 3 will collapse to a cluster at $t = 6$, as will agents 9, 10, 11, whereas the remaining agents 4, 5, 6, 7, 8 will collapse to a cluster at $t = 11$, in accordance with (3.1.2).

We can now give a rough outline of the proof of Theorem 3.1.1. Given $n \geq 11$, we proceed as follows:

Step 1: Show that the receptivity graph G_t, and hence the transition matrix B_t, is constant for $t = 0, 1, 2, 3$, whereas G_4 contains the two extra edges $\{1, 3\}$ and $\{n-2, n\}$. Thus $\boldsymbol{y}_5 = B_4 \, B_0^4 \, \boldsymbol{y}_0$. Then show that $\boldsymbol{y}_5(3) = \boldsymbol{y}_5(n-3) > 1$, which implies that three agents become disconnected at each end. If $\boldsymbol{y}_5(1) + \boldsymbol{y}_5(2) < 1$, then each group of three agents will collapse to a cluster at $t = 6$.

Note that we have already completed this step. The calculations above verify it for $n = 11$ and it is clear that, if we then increase n, it will not affect how the opinions of the first or last four agents evolve over the first five time steps. Indeed, the value of n cannot be "felt" before every agent has changed their opinion at least once which, as we remarked earlier, will not happen while $t < n/2$.

Step 2: Thus at $t = 5$, three agents break free from each end of the configuration. This leaves us with $n - 6$ agents and a corresponding vector of gaps $\overline{\boldsymbol{y}}_5 \in \mathbb{R}_+^{n-7}$. We now reset time to zero and consider $\overline{\boldsymbol{y}}_5$ as the new initial configuration. For example, with $n = 11$ we have, by (3.2.6),

$$\overline{\boldsymbol{y}}_5 \approx (0.9599, 0.9938, 0.9938, 0.9599). \tag{3.2.7}$$

The entries of $\overline{\boldsymbol{y}}_5$ will lie in $(0, 1]$. The idea is to show that they lie sufficiently close to 1 such that, if still $n - 6 \geq 11$, then the evolution of the receptivity graph over the next five time steps will be exactly the same as if all entries equalled 1. To complete the proof, we have to be able to iterate this procedure, and finally verify the theorem directly for $n \leq 10$ - indeed, we need to verify for these values of n that the behaviour is unaffected if the starting values in $\boldsymbol{y}_0$ are all in (0, 1] and sufficiently close to 1.

To complete Step 2, it is convenient to extend the HK-model to an infinite sequence of agents, more precisely to a well-ordered infinite sequence so that geometrically there is an agent furthest to the left. Indeed, (3.2) and (3.2) make perfect sense if we regard $\boldsymbol{x}_t$ and $\boldsymbol{y}_t$ as elements of $\mathbb{R}^\infty$, the vector space consisting of all infinite, well-ordered sequences of real numbers, and A_t, B_t as appropriate linear operators on this space. Let $\boldsymbol{\mathcal{E}}_\infty = (1,\, 2, \dots)$ denote the element of $\mathbb{R}^\infty$ representing a sequence of equally spaced agents with gaps of one. The obvious analogue of Theorem 3.1.1 would be the following:

Theorem 3.2.1. *The evolution of the configuration $\boldsymbol{\mathcal{E}}_\infty$ under (3.1.1) is periodic, namely after every fifth time step a group of three agents will disconnect on the left and then collapse to a cluster at the subsequent time step.*

Our strategy will be to first prove Theorem 3.2.1 and then argue that the behaviour is essentially unaffected when we go back to finite sequences.

To begin with, we define precisely the machinery we need, using the standard notation and terminology of functional analysis. An element of $\mathbb{R}^\infty$ will be denoted by a well-ordered sequence $\boldsymbol{x} = (x(i))_{i=1}^\infty$ of real numbers. Recall that l^∞ denotes the subspace of $\mathbb{R}^\infty$ consisting of bounded sequences. It is a Banach space with norm $||\boldsymbol{x}|| = \sup_i |x(i)|$. We let $\mathbf{1}_\infty = (1,\, 1, \dots)$ denote the element of l^∞ consisting entirely of ones. For a linear operator $T : l^\infty \to l^\infty$, its norm is defined as $||T|| = \sup_{||\boldsymbol{x}||=1} ||T(\boldsymbol{x})||$. One says that T is bounded if $||T|| < \infty$ and $B(l^\infty)$ denotes the Banach space of all bounded linear operators on l^∞. Let $B = (b(i,j))_{i,j=1}^\infty$ be a doubly-infinite matrix and set

$$s = \sup_i \sum_{j=1}^\infty |b(i,j)|. \tag{3.2.8}$$

If s is finite then the map T given by

$$(T(\boldsymbol{x}))(i) = \sum_{j=1}^\infty b(i,j)\, x(j), \quad \boldsymbol{x} \in l^\infty,$$

is a well-defined element of $B(l^\infty)$ of norm s. The map can be written as a matrix product $T(\boldsymbol{x}) = B\boldsymbol{x}$, when $\boldsymbol{x}$ is written as an infinite column.

We now consider a specific collection of operators $\widetilde{B}_t$, $t = 0, \dots, 4$ defined by matrices satisfying (3.2.8). In all cases, the elements $b(i,j)$ will be non-negative and there will be only finitely many non-zero entries in each row, i.e.: for each i we have $b(i,j) = 0$ for all $j \gg_i 0$. Set

$$\widetilde{B}_0(i,j) = \begin{cases} 1/6, & \text{if } (i,j) = (1,1), \\ 1/3, & \text{if } |i-j| \le 1 \text{ and } (i,j) \neq (1,1), \\ 0, & \text{otherwise}, \end{cases}$$

$$\widetilde{B}_3 = \widetilde{B}_2 = \widetilde{B}_1 = \widetilde{B}_0,$$

$$\widetilde{B}_4(i,j) = \begin{cases} B_4(i,j), & \text{if } i \le 3 \text{ and } j \le 10, \\ 1/3, & \text{if } i > 3 \text{ and } |i-j| \le 1, \\ 0, & \text{otherwise}. \end{cases} \quad (B_4 \text{ as in (3.2.5)}).$$

Let $S_3 : l^\infty \to l^\infty$ be a threefold leftward shift, i.e.:

$$(S_3\boldsymbol{x})(i) = x(i+3), \quad \forall\, i \ge 1, \tag{3.2.9}$$

and finally let $\mathcal{T} : l^\infty \to l^\infty$ be the composition

$$\mathcal{T} = S_3 \circ \widetilde{B}_4 \circ (\widetilde{B}_0)^4. \tag{3.2.10}$$

The point is that, firstly, the evolution under (3.1.1) of an infinite sequence of equally spaced agents is described for $t \leq 4$ by

$$\boldsymbol{y}_{t+1} = \widetilde{B}_t \, \boldsymbol{y}_t, \quad \boldsymbol{y}_0 = \mathbf{1}_\infty. \tag{3.2.11}$$

Secondly, at $t = 5$, the first three agents become disconnected. Hence, the map $\mathbf{1}_\infty \mapsto \mathcal{T}\mathbf{1}_\infty$ describes what happens if we take a sequence of equally spaced agents, run the HK-dynamics over 5 time steps, remove the first three agents which have become disconnected from the others, and reset time to zero. These assertions follow from (3.2.1)-(3.2.4) and (3.2.6), together with the observation that, whenever $\boldsymbol{y}_0$ is a multiple of $\mathbf{1}_\infty$, so that all its entries are equal, then $\boldsymbol{y}_t(i) = \boldsymbol{y}_0(i)$ for all $i \geq 6$ and all $t \leq 5$. Now we need to show two things:

CLAIM 1: If $||\boldsymbol{y}|| \leq 1$ and $||\mathbf{1}_\infty - \boldsymbol{y}|| < \varepsilon$ for some sufficiently small $\varepsilon > 0$, then $\boldsymbol{y} \mapsto \mathcal{T}\boldsymbol{y}$ still describes, as above, the evolution over 5 time steps of a sequence of agents with initial inter-agent spacings given by $\boldsymbol{y}$.

CLAIM 2: For all $n \in \mathbb{N}$, $||\mathcal{T}^n \mathbf{1}_\infty - \mathbf{1}_\infty||$ is sufficiently small so that Claim 1 can be applied.

Verifying these two claims will not immediately allow us to complete Step 2 above. Given a finite sequence of n equally spaced agents, these two claims allow us to iterate as in Step 2 up as far as $t \approx n/2$, i.e.: as long as agents in the middle have not yet been affected, because up to that point the evolution of either half of the finite sequence is exactly the same as for the corresponding initial segment of the infinite sequence. In order to continue the iteration beyond this time, we will use the precise quantitative estimates obtained in the proofs of the two claims below. Basically, the idea is that the entries of $\mathcal{T}^n \mathbf{1}_\infty$ remain much closer to one than is needed for Claim 1 to hold, so that even though the entries of the finite and infinite sequences diverge when $t > n/2$, the accumulated divergence never becomes so large so that we cannot apply Claim 1. The precise argument will follow the verification of the two claims.

Claim 1 is actually a statement about the operators $\widetilde{B}_t$, $0 \leq t \leq 4$. We prove the following:

Proposition 3.2.2. *Let $\boldsymbol{y}_0 \in l^\infty$ represent the gaps between consecutive agents in an infinite sequence of agents, indexed by $1, 2, \dots,$ and let G_0 be the corresponding receptivity graph. For each $t = 0, 1, \dots, 4$, let $\boldsymbol{y}_{t+1} = \widetilde{B}_t \, \boldsymbol{y}_t$, and let G_{t+1} be the corresponding graph. If $||\boldsymbol{y}_0|| \leq 1$ and $||\mathbf{1}_\infty - \boldsymbol{y}_0|| < \frac{7}{79}$ then the following hold:*

(i) for each $0 \leq t \leq 3$, G_t contains exactly the edges $\{i, i+1\}$, $i \in \mathbb{N}$,

(ii) G_4 contains the edges of G_3 plus the additional edge $\{1, 3\}$,

(iii) G_5 contains all the edges of G_4 except for the edge $\{3, 4\}$. Moreover, $\boldsymbol{y}_5(1) = 0$.

Hence Claim 1 holds with $\varepsilon = \frac{7}{79}$.

Proof. We already know that (i)-(iii) hold when $\boldsymbol{y}_0 = \mathbf{1}_\infty$. Now take $\boldsymbol{z}_0 = \frac{72}{79}\mathbf{1}_\infty$. As remarked earlier, since $\boldsymbol{z}_0$ is a multiple of $\mathbf{1}_\infty$, all its entries from the sixth onwards will remain unchanged for $t \leq 5$. So, when considering the evolution of the graph G_t, it suffices to consider the first five entries of each vector $\boldsymbol{z}_t$. Since every entry of $\boldsymbol{z}_0$ lies between 1/2 and 1, it is immediate that G_0 is as claimed in (i). We have $\boldsymbol{z}_t = \frac{72}{79}\boldsymbol{y}_t$ for all t, where the $\boldsymbol{y}_t$ are as in (3.2.1)-(3.2.4) and (3.2.6) except that we have an infinite sequence of ones from the sixth position onwards. Numerically,

$$\boldsymbol{z}_1 \approx (0.4557, \, 0.9114, \, 0.9114, \, 0.9114, \, 0.9114, \, \dots), \tag{3.2.12}$$

$$\boldsymbol{z}_2 \approx (0.3791,\ 0.7594,\ 0.9114,\ 0.9114,\ 0.9114,\ \dots), \tag{3.2.13}$$

$$\boldsymbol{z}_3 \approx (0.3165,\ 0.6835,\ 0.8608,\ 0.9114,\ 0.9114,\ \dots), \tag{3.2.14}$$

$$\boldsymbol{z}_4 \approx (0.2806,\ 0.6203,\ 0.8186,\ 0.8945,\ 0.9114,\ \dots), \tag{3.2.15}$$

$$\boldsymbol{z}_5 \approx (0,\ 0.3314,\ 1.0195,\ 0.8748,\ 0.9058,\ \dots). \tag{3.2.16}$$

Note that $\boldsymbol{z}_3(1) + \boldsymbol{z}_3(2) = 1$ exactly, as may be readily checked. For sequences $\boldsymbol{u}, \boldsymbol{v} \in l^\infty$, write $\boldsymbol{u} < \boldsymbol{v}$ if $u(i) < v(i)$ for every i and observe that, if M is a doubly-infinite matrix satisfying (3.2.8), having non-negative entries and at least one non-zero entry in each row, then $\boldsymbol{u} < \boldsymbol{v} \Rightarrow M\boldsymbol{u} < M\boldsymbol{v}$. Now if $||\boldsymbol{w}_0|| \leq 1$ and $||\mathbf{1}_\infty - \boldsymbol{w}_0|| < \frac{7}{79}$, then $\boldsymbol{z}_0 < \boldsymbol{w}_0 \leq \mathbf{1}_\infty$. Hence $\boldsymbol{z}_t < \boldsymbol{w}_t \leq \boldsymbol{y}_t$ would hold for all $t \leq 5$, but for the fact that the first entry of each vector is zero at $t = 5$. By comparing (3.2.1)-(3.2.4), (3.2.6) with (3.2.12)-(3.2.16), and noting in particular that both $\boldsymbol{w}_3(1) + \boldsymbol{w}_3(2) > 1$ and $\boldsymbol{w}_5(3) > 1$, it follows immediately that (i)-(iii) all hold for the $\boldsymbol{w}_t$. $\square$

We now turn to Claim 2. First of all, let us write out $\mathcal{T}$ explicitly as a doubly-infinite matrix. The upper-left 3×11 block is

$$\begin{pmatrix} \frac{47}{972} & \frac{59}{486} & \frac{5}{27} & \frac{17}{81} & \frac{5}{27} & \frac{10}{81} & \frac{5}{81} & \frac{5}{243} & \frac{1}{243} & 0 & 0 \\ \frac{1}{54} & \frac{5}{81} & \frac{10}{81} & \frac{5}{27} & \frac{17}{81} & \frac{5}{27} & \frac{10}{81} & \frac{5}{81} & \frac{5}{243} & \frac{1}{243} & 0 \\ \frac{1}{243} & \frac{5}{243} & \frac{5}{81} & \frac{10}{81} & \frac{5}{27} & \frac{17}{81} & \frac{5}{27} & \frac{10}{81} & \frac{5}{81} & \frac{5}{243} & \frac{1}{243} \end{pmatrix}. \tag{3.2.17}$$

Every other entry in the first three rows is zero, and every row from the fourth onwards is just a rightward shift of the third row, i.e.:

$$\mathcal{T}(i,j) = \begin{cases} 0, & \text{if } i \leq 3 \text{ and } j \geq 12, \\ 0, & \text{if } i \geq 4 \text{ and } j = 1, \\ \mathcal{T}(i-1, j-1), & \text{if } i \geq 4 \text{ and } j \geq 2. \end{cases} \tag{3.2.18}$$

Note that the sum of the entries in every row from the third onwards equals 1. Let

$$a = \sum_{j=1}^{9} \mathcal{T}(1,j) = \frac{311}{324} \approx 0.9599,$$

$$b = \sum_{j=1}^{10} \mathcal{T}(2,j) = \frac{161}{162} \approx 0.9938.$$

Proposition 3.2.3. *Let* $\boldsymbol{y}_0 := \mathbf{1}_\infty$ *and for all* $\tau \geq 0$, $\boldsymbol{y}_{\tau+1} := \mathcal{T}\boldsymbol{y}_\tau$. *Then*

(i) $\boldsymbol{y}_{\tau+1} \leq \boldsymbol{y}_\tau$ *for all* τ.

(ii) Let $\gamma := 0.1117$. *Then for all* $\tau \geq 1$,

$$1 - a \leq 1 - \boldsymbol{y}_\tau(1) < (1+\gamma)(1-a), \tag{3.2.19}$$

$$1 - b \leq 1 - \boldsymbol{y}_\tau(2) < \gamma(1-a) + (1+\gamma)(1-b), \tag{3.2.20}$$

$$1 - \boldsymbol{y}_\tau(i) < \gamma^{i-2}(1-a), \quad \forall\, i \geq 3. \tag{3.2.21}$$

(iii) the sequence $(\boldsymbol{y}_\tau)_{\tau=0}^\infty$ *converges in* l^∞ *to a fixed point of* $\mathcal{T}$.

Remark 3.2.4. *The reason for using* τ *instead of* t *is that multiplication by* $\mathcal{T}$ *is to be thought of as representing the evolution of a configuration of agents over* 5 *time steps. Hence, one can informally imagine that "*$\tau = 5t$*".*

Proof. Since the sum of the entries in every row of the matrix of $\mathcal{T}$ is at most one it is immediate that $\mathcal{T}(\mathbf{1}_\infty) \le \mathbf{1}_\infty$. Since $\mathcal{T}$ has non-negative entries, it follows by induction that $\boldsymbol{y}_{\tau+1} \le \boldsymbol{y}_\tau$ for every τ, which proves (i). Note that (iii) follows from (i) and (ii), so it remains to prove (ii).

The inequalities in (3.2.19)-(3.2.21) are obviously satisfied when $\tau = 1$. Hence, by (i), the left-hand inequalities in (3.2.19) and (3.2.20) will be satisfied for all $\tau \ge 1$. For the right-hand inequalities, we proceed by induction on τ. First consider $i = 1$. We have

$$\boldsymbol{y}_{\tau+1}(1) = \sum_{j=1}^{9} \mathcal{T}(1,j)\, \boldsymbol{y}_\tau(j) \Rightarrow 1 - \boldsymbol{y}_{\tau+1}(1) = (1-a) + \sum_{j=1}^{9} \mathcal{T}(1,j)\,(1 - \boldsymbol{y}_\tau(j)).$$

Assuming (3.2.19)-(3.2.21) hold at step τ, it follows that (3.2.19) holds at step $\tau + 1$ if and only if $f_1(\gamma) \le 0$, where

$$\begin{aligned} f_1(\gamma) = (1-a) &\left[-\gamma + \mathcal{T}(1,1)\,(1+\gamma) + \mathcal{T}(1,2)\,\gamma + \sum_{j=3}^{9} \mathcal{T}(1,j)\,\gamma^{j-2} \right] \\ &+ (1-b)\,\mathcal{T}(1,2)\,(1+\gamma). \end{aligned}$$

We checked with `Matlab` that $f_1(\gamma) = 0$ has two solutions in the interval $[0,1]$, at $\gamma_1 \approx 0.1116$ and $\gamma_2 \approx 0.9624$, and that $f_1(\gamma) < 0$ for $\gamma \in (\gamma_1, \gamma_2)$. Thus (3.2.19) holds also at step $\tau + 1$, provided $\gamma = 0.1117$.

Next consider $i = 2$. Analogous calculations lead to the requirement that $f_2(\gamma) \le 0$, where

$$\begin{aligned} f_2(\gamma) = (1-a) &\left[-\gamma + \mathcal{T}(2,1)\,(1+\gamma) + \mathcal{T}(2,2)\,\gamma + \sum_{j=3}^{10} \mathcal{T}(2,j)\,\gamma^{j-2} \right] \\ &+ (1-b)\left[-\gamma + \mathcal{T}(2,2)\,(1+\gamma) \right]. \end{aligned}$$

One checks that $f_2(\gamma) = 0$ has one solution in $[0,1]$, at $\gamma_3 \approx 0.03$, and that $f_2(\gamma) < 0$ for all $\gamma \in (\gamma_3, 1]$. Thus (3.2.20) holds also at step $\tau + 1$, provided $\gamma = 0.1117$.

For $i = 3$, we are led to the requirement that $f_3(\gamma) \le 0$, where

$$\begin{aligned} f_3(\gamma) = (1-a) &\left[-\gamma + \mathcal{T}(3,1)\,(1+\gamma) + \mathcal{T}(3,2)\,\gamma + \sum_{j=3}^{11} \mathcal{T}(3,j)\,\gamma^{j-2} \right] \\ &+ (1-b)\,\mathcal{T}(3,2)\,(1+\gamma). \end{aligned}$$

One checks that $f_3(\gamma) < 0$ for all $\gamma \in (\gamma_4,\, \gamma_5)$, where $\gamma_4 \approx 0.008$ and $\gamma_5 \approx 0.9965$.

For $i = 4$, using (3.2.18) we are led to the requirement that $f_4(\gamma) \le 0$ where

$$f_4(\gamma) = (1-a) \left[-\gamma^2 + \mathcal{T}(3,1)\,\gamma + \sum_{j=2}^{11} \mathcal{T}(3,j)\,\gamma^{j-1} \right] + (1-b)\,\mathcal{T}(3,1)\,(1+\gamma).$$

One checks that $f_4(\gamma) < 0$ for all $\gamma \in (\gamma_6,\, \gamma_7)$, where $\gamma_6 \approx 0.0430$ and $\gamma_7 \approx 0.9996$.

Finally, every $i \ge 5$ will lead to the condition that $f_5(\gamma) \le 0$ where

$$f_5(\gamma) = -\gamma^2 + \sum_{j=1}^{11} \mathcal{T}(3,j)\,\gamma^{j-1}.$$

One checks that $f_5(\gamma) < 0$ for all $\gamma \in (\gamma_8,\, 1]$, where $\gamma_8 \approx 0.0786$. This completes the induction step. $\square$

Claim 2 follows from Proposition 3.2.3. Indeed, for each $i \geq 1$, let $g_i = g_i(a,b,\gamma)$ be the bound on the appropriate right-hand side in (3.2.19)-(3.2.21). Since the g_i form a decreasing sequence, we have proven that $||\mathcal{T}^n \mathbf{1}_\infty - \mathbf{1}_\infty|| < g_1 \approx 0.0446 < \frac{7}{79}$ for all n. Together with Proposition 3.2.2, this already suffices to prove Theorem 3.2.1.

Turning to finite sequences of agents, we are now ready to complete the proof of our main result.

Proof of Theorem 3.1.1. Let $\boldsymbol{x}_0$ be any finite or infinite vector of equally spaced opinions, and $\boldsymbol{y}_0$ the corresponding vector of gaps. We start with a couple of observations:

(a) Suppose $\boldsymbol{x}_0$ is finite of length n, and thus $\boldsymbol{y}_0$ has length $n-1$. Given the updating rule (3.1.1), at all times t we will have $x_t(i) = x_t(n+1-i)$ for $i = 1, \dots, \lfloor \frac{n+1}{2} \rfloor$ and similarly $y_t(i) = y_t(n-i)$ for $i = 1, \dots, \lfloor \frac{n}{2} \rfloor$. Hence in order to understand how the configuration evolves, it suffices to understand the vectors $\left(y_t(1), \dots, y_t\left(\lfloor \frac{n}{2} \rfloor\right)\right)$ for all t.

(b) Suppose the receptivity graph G_t is as in Proposition 3.2.2 for all $t \leq 5$. Then, for any $k \in \mathbb{N}$ in the case of infinite vectors, or for any $k \leq n/2$ in the case of finite vectors, the values of $y_t(i)$ for all $0 \leq t \leq 5$, $1 \leq i \leq k$, only depend on $y_0(j)$, for $1 \leq j \leq k+5$.

Now let $n \geq 11$ be given. Write $\lfloor n/2 \rfloor := r$. For each $0 \leq \tau \leq m$, where m is a bound to be determined in a moment, we define a sequence of "replacement" operators $\mathcal{R}_\tau : l^\infty \to l^\infty$ as follows:

$$\begin{cases} (\mathcal{R}_\tau \boldsymbol{x})(i) = x(i), & \text{if } i \leq r - 3\tau \text{ or } i > r - 3\tau + 5, \\ (\mathcal{R}_\tau \boldsymbol{x})(r - 3\tau + j) = x(r - 3\tau - j), & \text{if } 1 \leq j \leq 5 \text{ and } n \text{ is even}, \\ (\mathcal{R}_\tau \boldsymbol{x})(r - 3\tau + j) = x(r - 3\tau + j + 1), & \text{if } 1 \leq j \leq 5 \text{ and } n \text{ is odd}. \end{cases} \tag{3.2.22}$$

The definition makes sense as long as $\tau \leq m$, where

$$m = \begin{cases} \lfloor \frac{r-6}{3} \rfloor, & \text{if } n \text{ is even}, \\ \lfloor \frac{r-5}{3} \rfloor, & \text{if } n \text{ is odd}. \end{cases} \tag{3.2.23}$$

Note that each $\mathcal{R}_\tau$ is linear and of norm one, since every entry in $\mathcal{R}_\tau \boldsymbol{x}$ is also an entry in $\boldsymbol{x}$.

Set $\boldsymbol{y}_0 = \boldsymbol{z}_0 = \mathbf{1}_\infty$. For each $0 \leq \tau \leq m$, let $\Theta_\tau := \mathcal{T} \circ \mathcal{R}_\tau$ and define inductively

$$\boldsymbol{y}_{\tau+1} = \mathcal{T} \boldsymbol{y}_\tau, \quad \boldsymbol{z}_{\tau+1} = \Theta_\tau \, \boldsymbol{z}_\tau, \quad \delta_\tau := ||\boldsymbol{z}_\tau - \boldsymbol{y}_\tau||. \tag{3.2.24}$$

We are interested in bounding the δ_τ. Clearly, $\delta_0 = 0$. Since $||T|| = 1$ we have

$$\delta_{\tau+1} = ||\mathcal{T}(\mathcal{R}_\tau \, \boldsymbol{z}_\tau) - \mathcal{T}(\boldsymbol{y}_\tau)|| \leq ||\mathcal{R}_\tau(\boldsymbol{z}_\tau) - \boldsymbol{y}_\tau||. \tag{3.2.25}$$

Next, by the triangle inequality and the properties of $\mathcal{R}_\tau$,

$$\begin{aligned} ||\mathcal{R}_\tau(\boldsymbol{z}_\tau) - \boldsymbol{y}_\tau|| &= ||\mathcal{R}_\tau(\boldsymbol{z}_\tau) - \mathcal{R}_\tau(\boldsymbol{y}_\tau) + \mathcal{R}_\tau(\boldsymbol{y}_\tau) - \boldsymbol{y}_\tau|| \leq \\ &\leq ||\mathcal{R}_\tau(\boldsymbol{z}_\tau) - \mathcal{R}_\tau(\boldsymbol{y}_\tau)|| + ||\mathcal{R}_\tau(\boldsymbol{y}_\tau) - \boldsymbol{y}_\tau|| = \\ = ||\mathcal{R}_\tau(\boldsymbol{z}_\tau - \boldsymbol{y}_\tau)|| + ||\mathcal{R}_\tau(\boldsymbol{y}_\tau) - \boldsymbol{y}_\tau|| &\leq ||\boldsymbol{z}_\tau - \boldsymbol{y}_\tau|| + ||\mathcal{R}_\tau(\boldsymbol{y}_\tau) - \boldsymbol{y}_\tau|| = \\ &= \delta_\tau + ||\mathcal{R}_\tau(\boldsymbol{y}_\tau) - \boldsymbol{y}_\tau||. \end{aligned}$$

Thus we have the recurrence

$$\delta_{\tau+1} \leq \delta_\tau + ||\mathcal{R}_\tau(\boldsymbol{y}_\tau) - \boldsymbol{y}_\tau||. \tag{3.2.26}$$

We now use Proposition 3.2.3 to bound the second term on the right of (3.2.26). It follows immediately from the definition of $\mathcal{R}_\tau$ and the fact that the numbers g_i are decreasing with i that

$$||\mathcal{R}_\tau(\boldsymbol{y}_\tau) - \boldsymbol{y}_\tau|| \leq \begin{cases} g_{r-3\tau-5}, & \text{if } n \text{ is even}, \\ g_{r-3\tau-4}, & \text{if } n \text{ is odd}. \end{cases}$$

Hence, for all $\tau \leq m$,

$$\delta_\tau \leq \delta_\infty := \sum_{k=1}^{\infty} g_{1+3k} = (1-a)\left(\frac{\gamma^2}{1-\gamma^3}\right) \approx 0.0005. \tag{3.2.27}$$

Thus for all $\tau \leq m$, and using Proposition 3.2.2 again,

$$||\boldsymbol{z}_\tau - \mathbf{1}_\infty|| \leq ||\boldsymbol{z}_\tau - \boldsymbol{y}_\tau|| + ||\boldsymbol{y}_\tau - \mathbf{1}_\infty|| \leq \delta_\infty + g_1 < \frac{7}{79}. \tag{3.2.28}$$

Since the operator $\mathcal{R}_\tau$ just replaces some elements of $\boldsymbol{z}_\tau$ with others, we also have $||\mathcal{R}_\tau(\boldsymbol{z}_\tau) - \mathbf{1}_\infty|| < 7/79$ for all $\tau \leq m$. Thus Proposition 3.2.2 holds for each vector $\mathcal{R}_\tau(\boldsymbol{z}_\tau)$. The point is that this is exactly what we need in order to deduce that a finite sequence of n equally spaced agents, with initial gaps of one, will evolve as claimed in Theorem 3.1.1, up to $\tau = m+1$, in other words up to time $t = 5(m+1)$, where m is related to n by (3.2.23). This is a direct consequence of observations (a) and (b) above.

To complete the proof of the theorem, it just remains to consider what happens from time $5(m+1)$ onwards. At this time, a total of $6(m+1)$ agents will have become disconnected, and by the next time step will have all collapsed into $2(m+1)$ clusters of size 3 each. We will be left, at time $5(m+1)$, with a group of somewhere between 5 and 10 agents in the middle, depending on the value of $n \pmod 6$. The gaps between these remaining agents will still be less than and close to one, indeed we can use a bound similar to (3.2.27). We will need to add the $k = 0$ term, but can also start the sum from either g_1, g_2 or g_3, depending on $n \pmod 6$. Set

$$\delta_{1,\infty} := \sum_{k=0}^{\infty} g_{1+3k} = (1-a)\left(1 + \gamma + \frac{\gamma^2}{1-\gamma^3}\right) \approx 0.0451,$$
$$\delta_{2,\infty} := \sum_{k=0}^{\infty} g_{2+3k} = (1-a)\left(\gamma + \frac{\gamma^3}{1-\gamma^3}\right) + (1-b)(1+\gamma) \approx 0.0114,$$
$$\delta_{3,\infty} := \sum_{k=0}^{\infty} g_{3+3k} = (1-a)\left(\frac{\gamma}{1-\gamma^3}\right) \approx 0.0045.$$

One can check exhaustively that the remaining middle component of $G_{5(m+1)}$, depending on $n \pmod 6$, will satisfy the following:

Case 1: $n \equiv 5 \pmod 6$.

We have 5 agents left, with gaps represented by the vector (y_1, y_2, y_2, y_1), where

$$y_1 \geq 1 - \delta_{1,\infty} - g_1 \approx 0.9103,$$
$$y_2 \geq 1 - \delta_{1,\infty} - g_2 \approx 0.9435.$$

Case 2: $n \equiv 0 \pmod 6$.

We have 6 agents left, with gaps represented by the vector $(y_1, y_2, y_3, y_2, y_1)$, where y_1, y_2 satisfy the same inequalities as in Case 1 and

$$y_3 \geq 1 - \delta_{1,\infty} - g_3 \approx 0.9504.$$

Case 3: $n \equiv 1 \pmod 6$.

We have 7 agents left, with gaps represented by the vector $(y_1, y_2, y_3, y_3, y_2, y_1)$, where

$$\begin{aligned} y_1 &\geq 1 - \delta_{2,\infty} - g_1 \approx 0.9440, \\ y_2 &\geq 1 - \delta_{2,\infty} - g_2 \approx 0.9773, \\ y_3 &\geq 1 - \delta_{2,\infty} - g_3 \approx 0.9841. \end{aligned}$$

Case 4: $n \equiv 2 \pmod 6$.

We have 8 agents left, with gaps represented by the vector $(y_1, y_2, y_3, y_4, y_3, y_2, y_1)$, where y_1, y_2, y_3 satisfy the same inequalities as in Case 3 and

$$y_4 \geq 1 - \delta_{2,\infty} - g_4 \approx 0.9881.$$

Case 5: $n \equiv 3 \pmod 6$.

We have 9 agents left, with gaps represented by the vector $(y_1, y_2, y_3, y_4, y_4, y_3, y_2, y_1)$, where

$$\begin{aligned} y_1 &\geq 1 - \delta_{3,\infty} - g_1 \approx 0.9509, \\ y_2 &\geq 1 - \delta_{3,\infty} - g_2 \approx 0.9842, \\ y_3 &\geq 1 - \delta_{3,\infty} - g_3 \approx 0.9910, \\ y_4 &\geq 1 - \delta_{3,\infty} - g_4 \approx 0.9950. \end{aligned}$$

Case 6: $n \equiv 4 \pmod 6$.

We have 10 agents left, with gaps represented by the vector $(y_1, y_2, y_3, y_4, y_5, y_4, y_3, y_2, y_1)$, where y_1, y_2, y_3, y_4 satisfy the same inequalities as in Case 5 and

$$y_5 \geq 1 - \delta_{3,\infty} - g_5 \approx 0.9954.$$

When considering the further evolution of one of these six configurations, we can adopt the same strategy as in the proof of Proposition 3.2.2. We look on the one hand at what happens when each y_i has the minimum value allowed by the inequalities, on the other at what happens when each $y_i = 1$, and then "interpolate" between these two extremes. We performed all computations in `Matlab` and it turns out that the following occurs:

Case 1: The five agents will always collapse to a single cluster after 6 time steps, though there are two different possibilities for the sequence of receptivity graphs. One possibility is that agent 3 will become connected to agents 1 and 5 after three steps. If that happens, agents 1 and 2 will agree after four steps, as will agents 4 and 5. However, these pairs will still be at distance greater than one from one another, so the final collapse to a cluster will require two further time steps. The other possibility is that agent 3 will become connected to 1 and 5 for the first time after four steps. In that case, the merged pair $\{1, 2\}$ will be within distance one of the merged pair $\{4, 5\}$ at step five and thus we collapse to a cluster again at step six.

In all remaining cases, there is only one possible sequence of receptivity graphs.

Case 2: Agents 1 and 3 will become connected after four steps, as will agents 4 and 6. At step five, agents 1, 2, 3 will disconnect from 4, 5, 6, and each group of three will collapse to a cluster at step six.

CASE 3: At step five, we will split into three components, consisting of $\{1, 2, 3\}$, $\{4\}$ and $\{5, 6, 7\}$. Each boundary component collapses to a cluster at step six.

CASE 4: At step five, we will split into three components, consisting of $\{1, 2, 3\}$, $\{4, 5\}$ and $\{6, 7, 8\}$. Each component collapses to a cluster at step six.

CASE 5: At step five, we will split into three components, consisting of $\{1, 2, 3\}$, $\{4, 5, 6\}$ and $\{7, 8, 9\}$. Each boundary component will collapse to a cluster at step six, whereas the middle component will collapse at step seven.

CASE 6: At step five, we will split into three components, consisting of $\{1, 2, 3\}$, $\{4, 5, 6, 7\}$ and $\{8, 9, 10\}$. Each boundary component collapses to a cluster at step six, but the middle component will only collapse at step ten.

The above analysis serves to verify that the values of $\varepsilon(l)$ in (3.1.2) are correct for every $0 \le l \le 5$, and thus completes the proof of Theorem 3.1.1.

3.3 The general case of equally spaced opinions

In this section we will study the evolution of a finite or infinite sequence of opinions, updating according to (3.1.1) and initially equally spaced with gaps equal to d, where d is an arbitrary real number in the interval $(0, 1]$. Theorem 3.1.1 describes a striking regularity in the evolution when $d = 1$. One motivation for looking at other values of d is the observation that there exists some $\epsilon > 0$ such that Theorems 3.1.1 and 3.2.1 hold verbatim for all $d \in (1 - \epsilon, 1]$. Indeed, this follows immediately from an examination of the argument in Section 2. To begin with, observe that the initial receptivity graph is identical for all $d > 1/2$, namely it is just a chain. Next, let's consider an infinite sequence of agents. Proposition 3.2.2 tells us that the evolution up to $t = 5$ is identical for all $d > 72/79$. In Proposition 3.2.3 we just replace $\mathbf{1}_\infty$ by a multiple $d\mathbf{1}_\infty$ of itself and rescale all inequalities - for example, the analogue of (3.2.19) will read: $d(1-a) \le d - \boldsymbol{y}_\tau(1) < d(1+\gamma)(1-a)$. Clearly, for d sufficiently close to 1, such bounds will still be good enough to be able to apply Proposition 3.2.2, even allowing for the additional perturbations introduced when returning to finite sequences, which we dealt with at the end of Section 3.2.

The obvious question then is to what extent something like Theorems 3.1.1 and 3.2.1 continues to hold as d decreases. Figure 3.2 below suggests that something very interesting may be going on. On the horizontal axis, we plot d on a logarithmic scale. Starting from 1 and decreasing to 0.005, we ran for each value of d a simulation of $\lfloor \frac{30}{d} \rfloor$ agents. [1] This gives the configuration a diameter of ≈ 30, so that for all times $t < 15$ the edge of the configuration look like the edge of an infinite system with the same d. The blue dots indicate the first time $t = L(d)$ at which the receptivity graph disconnects. The green triangles record the number $M(d)$ of agents that disconnect at time $L(d)$ from each end of the graph. Note that $M(d)$ remains much smaller than n which means that, for every d we simulated, the behaviour up to time $L(d)$ would be identical for any sufficiently large number (depending on d) of agents and hence, in particular, for an infinite sequence of agents, except that in that case the graph will

[1] The precise expression for the vector of values of d is `0.005.^(linspace(1,0,100)).^2`

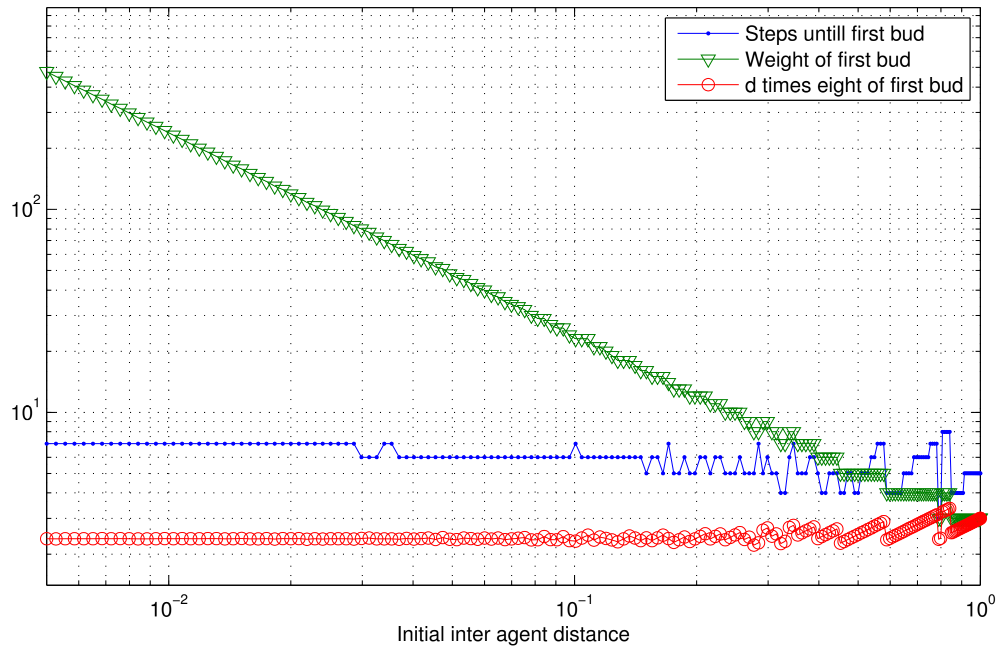

Figure 3.2: Data for the time of first disconnection.

only disconnect on the left. This explains our notation. The red circles record the products $d \cdot M(d)$.

There are two obviously striking features in Figure 3.2. Firstly, $L(d)$ does not seem to be increasing as d decreases. It is not constant - recall that Theorem 3.2.1 says that $L(1) = 5$, while the figure shows that $L(d)$ can attain any value among $\{3, \ldots, 9\}$. Indeed, all these values are already attained in the interval $d \in \left(\frac{1}{2}, 1\right]$, see Table 3.1 below. Moreover, $L(d) = 6$ for "most" values of d simulated with $log_{10}(d) \approx -1$, whereas $L(d) = 7$ for most values with $log_{10}(d) \approx -2$. This may suggest logarithmic growth. When d gets really small then simulations become impractical. However, it is obvious to ask

Question 3.3.1. *Is there an absolute positive constant L such that $L(d) \leq L$ for all $d \in (0, 1]$? If not, is it at least true that $L(d) = O\left(\log \frac{1}{d}\right)$?*

The second striking feature is that the products $d \cdot M(d)$ seem to be hardly changing $d \to 0$. In fact, they are all close to 2.38. We do not know if they converge to a limit, however.

Question 3.3.2. *Is it true that $d \cdot M(d)$ converges to a limit as $d \to 0$? If so, what is the limit? If not, is it at least true that $d \cdot M(d) = \Theta(1)$ for all $d \in (0, 1]$?*

These two questions concern only the behaviour up to the first disconnection in the receptivity graph. The meat in Theorems 3.1.1 and 3.2.1 is that this behaviour is then repeated, forever in the case of an infinite sequence of agents and until there are less than $2 \cdot M(1) = 6$ agents left in the finite case. It turns out that such simple periodicity does not hold for arbitrary d, though there is a lot of evidence that the behaviour is always "close" to periodic. The rest of this section will be concerned with developing this assertion. If we accept it for the moment, then following on from Questions 3.3.1 and 3.3.2 we can ask about the freezing time for an arbitrary configuration of equally spaced agents. If the evolution were perfectly periodic for all d then, as $d \to 0$, Question 3.3.2 would imply that the freezing time is $O[L(d) \cdot (dn)]$ for $n \gg_d 0$. Notice that $d(n-1)$ is the diameter of the configuration. As Question 3.3.1 suggests $L(d)$ grows very slowly, if at all, this would imply that, for general d, the diameter is a much better measure of the freezing time than the number of agents. Indeed, a universal bound on $L(d)$ would in turn suggest an affirmative answer to the following:

Question 3.3.3. *Does there exist a universal positive constant κ such that, for any finite configuration of equally spaced agents obeying (3.1.1), the freezing time is at most $\kappa \cdot D$, where D is the diameter of the configuration?*

To appreciate how far we are from being able to answer any of the questions posed so far, we cannot even prove that, for all d and $n \gg_d 0$, the receptivity graph must actually disconnect at all ! Nor can we prove even an $O(n)$ bound for the freezing time, independent of d. On the other hand, the intuition that equal spacings of $d = 1$ should yield the most slowly converging configuration turns out to be false, even amongst equally spaced configurations. We will prove below that for a short interval of d-values slightly above 0.8, the freezing time is $n + O(1)$. We conjecture that this is the worst-case scenario:

Question 3.3.4. *Is it true that any configuration of n equally spaced opinions, which evolve according to (3.1.1), will freeze by time $n + c$, where $c > 0$ is an absolute constant?*

Let's now go into more detail. In what follows, we denote $\boldsymbol{\mathcal{E}}_{n,\,d} := (1,\, 1{+}d, \dots,\, 1{+}(n{-}1)d) \in \mathbb{R}^n$ and $\boldsymbol{\mathcal{E}}_{\infty,\,d} := (1,\, 1+d,\, \dots) \in \mathbb{R}^\infty$. Proposition 3.2.2 says that the behaviour up to the first disconnection in the receptivity graph is identical for all d in a half-open interval to the left of $d = 1$. Basically, the reason for this is that, before the disconnection, edges are only added to the graph, not removed. Consider a fixed edge $\{i, j\}$ with $i < j$ and a fixed $d = d^*$. If this edge is added to the graph at time t it means that (1) $x_t(j) - x_t(i) \leq 1$, whereas (2) $x_{t-1}(j) - x_{t-1}(i) > 1$. If we now decrease d and assume that the evolution of the graph is identical up to time $t - 1$, then clearly (1) will still hold, while (2) will hold for all $d \in (d^* - \varepsilon,\, d^*]$ for some $\varepsilon > 0$. Hence, if edges are only added, never deleted, before the first disconnection, then the evolution of the graph during this period will be identical for all $d \in (d^* - \varepsilon^*,\, d^*]$, where ε^* will in general depend on d^* - in particular, it will likely be smaller if the time $L(d^*)$ at which the disconnection occurs is greater. We may ask whether this is what indeed always happens:

Question 3.3.5. *Is it true that, for every $d \in (0, 1]$, edges are only added to the receptivity graph, never deleted, before the time $t = L(d)$ at which it disconnects? Consequently, is it true that there is a decreasing sequence*

$$1 = d_0 > d_1 > \cdots \tag{3.3.1}$$

such that the evolution of the graph up to the first disconnection, and hence the values of the functions $L(d)$ and $M(d)$, are constant on each interval $d \in (d_{i+1},\, d_i]$? Moreover, does the sequence d_i tend to zero?

By exhaustive computation we have constructed the sequence (3.3.1) down to $d = 1/2$, which is a natural threshold as it marks the point at which the receptivity graph is more than just a chain at $t = 0$. Table 3.1 below shows the 12 different possibilities for the evolution prior to disconnection of an infinite sequence of agents for all $d > 1/2$. The rightmost column in the table describes the precise evolution. Here $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ are operators on l^∞, considered as bi-infinite matrices. $\mathcal{A}$ is what we called $\widetilde{B}_0$ in Section 3.2, the operator corresponding to when the receptivity graph is just a chain. $\mathcal{B}$ is what we previously called $\widetilde{B}_4$, corresponding to addition of the edge $\{1, 3\}$. For $\mathcal{C}$ we in turn add the edges $\{i, 4\}$, $1 \leq i \leq 2$ and for $\mathcal{D}$ we then add the edges $\{j, 5\}$, $1 \leq j \leq 3$. $\mathcal{S}$ denotes a shift operator. We take $\boldsymbol{y}_0 = d\,\mathbf{1}_\infty$ and the map $\boldsymbol{y}_0 \mapsto \mathcal{T}\boldsymbol{y}_0$ describes the evolution up to time $L(d)$, followed by removing the segment of size $M(d)$ which has become disconnected.

i	d_i	$L(d_i)$	$M(d_i)$	$\mathcal{T} = \mathcal{T}_i$
0	1	5	3	$\mathcal{S}_3 \cdot \mathcal{B} \cdot \mathcal{A}^4$
1	$\frac{72}{79} \approx 0.9114$	4	3	$\mathcal{S}_3 \cdot \mathcal{B} \cdot \mathcal{A}^3$
2	$\frac{864}{1027} \approx 0.8413$	9	4	$\mathcal{S}_4 \cdot \mathcal{C} \cdot \mathcal{B}^5 \cdot \mathcal{A}^3$
3	$\frac{31104}{36979} \approx 0.8411$	8	4	$\mathcal{S}_4 \cdot \mathcal{C} \cdot \mathcal{B}^4 \cdot \mathcal{A}^3$
4	$\frac{4}{5}$	3	3	$\mathcal{S}_3 \cdot \mathcal{B} \cdot \mathcal{A}^2$
5	$\frac{48}{61} \approx 0.7869$	7	4	$\mathcal{S}_4 \cdot \mathcal{C} \cdot \mathcal{B}^4 \cdot \mathcal{A}^2$
6	$\frac{1296}{1735} \approx 0.7470$	6	4	$\mathcal{S}_4 \cdot \mathcal{C} \cdot \mathcal{B}^3 \cdot \mathcal{A}^2$
7	$\frac{54}{79} \approx 0.6835$	5	4	$\mathcal{S}_4 \cdot \mathcal{C} \cdot \mathcal{B}^2 \cdot \mathcal{A}^2$
8	$\frac{2}{3}$	5	4	$\mathcal{S}_4 \cdot \mathcal{C} \cdot \mathcal{B}^3 \cdot \mathcal{A}$
9	$\frac{9}{14} \approx 0.6429$	4	4	$\mathcal{S}_4 \cdot \mathcal{C} \cdot \mathcal{B}^2 \cdot \mathcal{A}$
10	$\frac{1440}{2459} \approx 0.5856$	7	5	$\mathcal{S}_5 \cdot \mathcal{D} \cdot \mathcal{C}^3 \cdot \mathcal{B}^2 \cdot \mathcal{A}$
11	$\frac{69120}{125359} \approx 0.5514$	6	5	$\mathcal{S}_5 \cdot \mathcal{D} \cdot \mathcal{C}^2 \cdot \mathcal{B}^2 \cdot \mathcal{A}$

Table 3.1: The possible "states" prior to first disconnection for all $d > 1/2$.

To verify that this list is complete can be reduced to a finite computation. To get a flavour of how this works, consider $d_7 = 54/79$. If $\boldsymbol{y}_0 := d_7 \mathbf{1}_\infty$ and $\boldsymbol{y}_4 := \mathcal{B}^2 \cdot \mathcal{A}^2 \cdot \boldsymbol{y}_0$ one may verify that $\boldsymbol{y}_4(1) = 0$ and $\boldsymbol{y}_4(2) + \boldsymbol{y}_4(3) = 1$. This means that the edges $\{1, 4\}$ and $\{2, 4\}$ will be added to the graph at $t = 4$ when $d = d_7$ but not when $d > d_7$, so at $d = d_7$ the behaviour changes. If now $\boldsymbol{y}_5 := \mathcal{C} \cdot \boldsymbol{y}_4$ then $\frac{1}{d_7} \boldsymbol{y}_5(4) = \frac{14189}{8640} \approx \frac{1}{0.6089}$, which means that adding the edges $\{1, 4\}$ and $\{2, 4\}$ at $t = 4$ would result in the removal of edge $\{4, 5\}$, and hence a disconnection of the graph, at $t = 5$ for any $d \in \left(\frac{8640}{14189}, \frac{54}{79}\right]$. However, this is a faithful description of the evolution only down to $d_8 = 2/3$, because then the behaviour changes already at $t = 1$. For if instead $\boldsymbol{y}_0 := d_8 \mathbf{1}_\infty$ and $\boldsymbol{y}_1 := \mathcal{A} \cdot \boldsymbol{y}_0$ then $\boldsymbol{y}_1(1) + \boldsymbol{y}_1(2) = 1$ so the edge $\{1, 3\}$ is already added at $t = 1$.

Now notice from the Table that the quotient $L(d)/M(d)$ is not maximised at $d_0 = 1$, it attains greater values of $9/4$ and $8/4 = 2$ at d_2 and d_3 respectively. For values of d in the very narrow interval $(d_3, d_2]$, however, the subsequent evolution is not periodic, rather after a finite amount of time, the behaviour will hop to that exhibited in the interval $(d_4, d_3]$. We will return to this later. However we can prove that, for d lying in some subinterval of $(d_4, d_3]$, the behaviour is periodic.

Theorem 3.3.6. *There exists a non-empty open subinterval $I \subseteq (d_4, d_3]$ such that for all $d \in I$, the following holds:*

(i) If the initial configuration $\boldsymbol{\mathcal{E}}_{\infty, d}$ evolves according to (3.1.1) then the evolution is periodic: after every eighth time step, a group of four agents disconnect from the left-hand end of the receptivity graph and collapse to a cluster at the subsequent time step.

(ii) there are positive constants C, C' such that the finite configuration $\boldsymbol{\mathcal{E}}_{n, d}$ will evolve in an analogous manner, with a group of four agents becoming disconnected at each end after every eighth time step, until there are at most C agents left in the middle. These will collapse to a cluster after at most C' further time steps.

Note that the statement in the finite case is slightly less precise than in Theorem 3.1.1. This is because, already for $n = 7$, the evolution of the graph is not constant for $d \in (d_4, d_3]$. As may be verified by direct computation, three different things can happen and the freezing time can be 8, 9 or 10. In any case, the important point about Theorem 3.3.6(ii) is that the freezing time is $n + O(1)$.

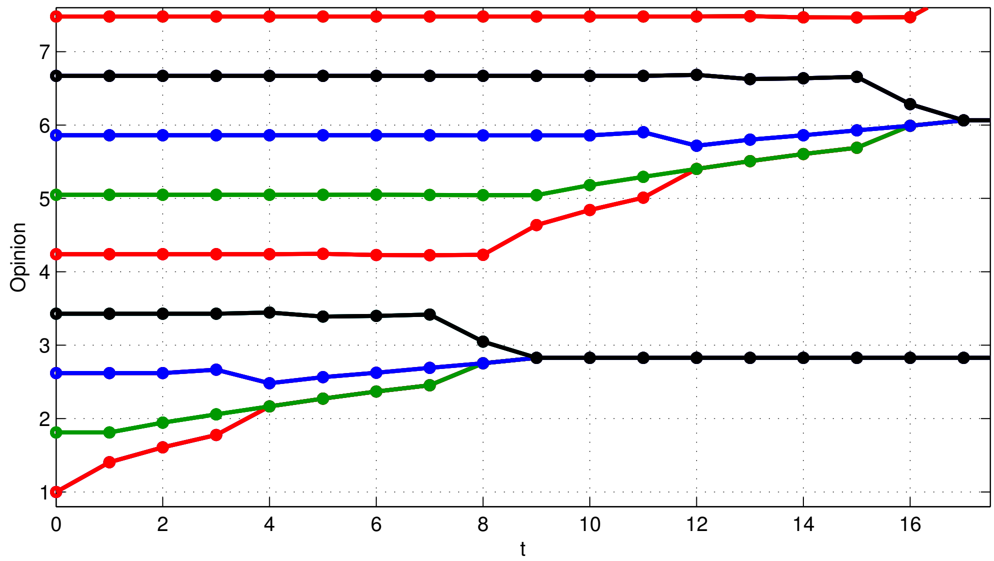

Figure 3.3: With $d \approx 0.81$ four agents on each side are disconnected from the central connected component every eighth time step. As with the case $d = 1$, what is left is very similar to a shorter version of the original chain.

Proof. The proof is completely analogous to that given in Section 3.2 so we only present a sketch. Table 3.1 yields an immediate analogue of Proposition 3.2.2, namely: for an infinite sequence of agents with initial spacings given by $\boldsymbol{y}_0 \in l^\infty$, if $d_4\mathbf{1}_\infty < \boldsymbol{y}_0 \leq d_3\mathbf{1}_\infty$, then the evolution up to $t = 8$ is identical to that described in the $i = 3$ row of the table, i.e.:

$$\boldsymbol{y}_t = \mathcal{A}^t\boldsymbol{y}_0 \text{ for } 0 \leq t \leq 3, \qquad \boldsymbol{y}_t = \mathcal{B}^{t-3}\boldsymbol{y}_3 \text{ for } 3 \leq t \leq 7, \qquad \boldsymbol{y}_8 = \mathcal{C}\boldsymbol{y}_7.$$

Secondly, consider the bi-infinite matrix $\mathcal{T}_3$. Analogously to (3.2.17) and (3.2.18), all the "action" takes place in the upper-left 5×17 block, and every row from the sixth onwards is just a shift to the right of the previous row. The sum of the entries in every row from the fifth onwards equals one. Let a, b, c, d respectively denote the sums of the entries in the first four rows. One can check that

$$a_3 = \frac{375281}{373248} \approx 1.0054, \quad b_3 = \frac{281497}{279936} \approx 1.0056, \tag{3.3.2}$$

$$c = \frac{7787}{7776} \approx 1.0014, \quad d = \frac{4373}{4374} \approx 0.9998. \tag{3.3.3}$$

One can prove the following analogues of (3.2.19)-(3.2.21), the proof reduced as before to solving (in `Matlab`) a finite collection of polynomial equations:

Let $\boldsymbol{y}_0 := \mathbf{1}_\infty$ and for all $\tau \geq 0$, $\boldsymbol{y}_{\tau+1} := \mathcal{T}_3\,\boldsymbol{y}_\tau$. Let $\gamma_3 := 0.3107$. Then for all $\tau \geq 1$,

$$a_3 - 1 \leq \boldsymbol{y}_\tau(1) - 1 \leq (1+\gamma_3)(b_3 - 1), \tag{3.3.4}$$

$$b_3 - 1 \leq \boldsymbol{y}_\tau(2) - 1 \leq (1+\gamma_3)(b_3 - 1), \tag{3.3.5}$$

$$c - 1 \leq \boldsymbol{y}_\tau(3) - 1 \leq \gamma_3(b_3 - 1), \tag{3.3.6}$$

$$|\boldsymbol{y}_\tau(i) - 1| \leq \gamma_3^{i-2}(b_3 - 1), \quad \forall\, i \geq 4. \tag{3.3.7}$$

Note that, according to (3.3.2), there are row sums both above and below one, and thus the sequence $\boldsymbol{y}_\tau$ is not monotonic as in Proposition 3.2.3(i). Thus the left-hand inequalities in (3.3.4) and (3.3.5) also contribute to the collection of polynomial equations to be solved here. In addition, because of the lack of monotonicity, it is not immediately obvious that $(\boldsymbol{y}_\tau)$ converges in l^∞ to a fixed point of $\mathcal{T}_3$. Though computations suggest this is definitely the case, we have not tried to prove it as we don't need the result.

Now it follows from (3.3.4)-(3.3.7) that, for all $\tau \geq 0$ and $d \in (d_4\, d_3]$, $||d\mathbf{1}_\infty - d\boldsymbol{y}_\tau|| \leq d_3(1+\gamma_3)(b_3-1) \approx 0.0061$. This is much less than half the length of the interval $(d_4,\, d_3]$. So as long as we don't choose d too close to the edges of the interval, the behaviour described in part (i) of Theorem 3.3.6 for an infinite sequence of agents has been established. In the case of a finite sequence of n agents there will be additional "perturbations" once agents in each half of the sequence affect one another, and these can be analysed by introducing operators $\mathcal{R}_{3,\tau}$, $\Theta_{3,\tau} = \mathcal{T}_3\,\mathcal{R}_{3,\tau}$ and vectors $\boldsymbol{z}_\tau = \boldsymbol{z}_{3,\tau}$ analogous to (3.2.22) and (3.2.24). In (3.2.22), $\mathcal{R}_\tau = \mathcal{R}_{0,\tau}$ replaced elements of an input $\boldsymbol{x}$ in groups of five and moved this window three steps to the left as τ increased. In the present case, $\mathcal{R}_{3,\tau}$ will replace elements eight at a time and move the window four steps to the left each time. In imitating the argument on page 43, we must exercise a little caution since $||\mathcal{T}_3|| = b_3 > 1$, which in (3.2.25) would lead to an exponentially growing bound for δ_τ. However, since $\mathcal{T}_3$ is essentially a "band matrix" and the sum of the entries in each row from the fifth onwards is one, multiplying by $\mathcal{T}_3$ will not magnify errors until $\tau \geq m_3 - O(1)$, in fact until $\tau \geq m_3 - 4$, where now

$$m_3 = \begin{cases} \lfloor \frac{r-9}{4} \rfloor, & \text{if } n \text{ is even,} \\ \lfloor \frac{r-8}{4} \rfloor, & \text{if } n \text{ is odd.} \end{cases} \qquad (r = n/2).$$

Indeed we still have much more margin for error here than in Section 3.2 so we can continue replacement up to $\tau = m_3 + 1$ and obtain the following analogues of (3.2.28) and (3.2.27):

$$||\boldsymbol{z}_{m_3+1} - \mathbf{1}_\infty|| \leq \delta^*_\infty + g^*_1,$$

where

$$g^*_1 = d_3(1+\gamma_3)(b_3-1) \approx 0.0061$$

and

$$\delta^*_\infty = d_3 \cdot b_3^4 \cdot \sum_{k=0}^{\infty} |g^*_{1+4k}| = d_3 \cdot b_3^4 \cdot \left[g^*_1 + (b-1)\left(\frac{\gamma_3^3}{1-\gamma_3^4}\right)\right] \approx 0.0054,$$

Thus $||\boldsymbol{z}_{m_3+1} - \mathbf{1}_\infty|| < 0.0116$. This is still less than $\frac{1}{2}(\delta_3 - \delta_4)$, which proves that the evolution of the configuration of n agents up to time $t = 8(m_3+1)$ will be periodic, at least for all d in some sufficiently small interval around the centre of $(d_4,\, d_3]$. At time $8(m_3+1)$ there will be somewhere between 8 and 15 agents left, depending on $n \pmod 8$, so the proof of the theorem is complete. □

Theorems 3.2.1 and 3.3.6 prove that, at least for some values of $d > 1/2$, the evolution of the configuration $\boldsymbol{\mathcal{E}}_{\infty,d}$ is periodic, with $M(d)$ agents becoming disconnected on the left after every $L(d)$ time steps. Such simple periodicity does not hold for arbitrary d. Indeed, if d is close to some d_i, $i > 0$, and hence close to the boundary between the two intervals $(d_i,\, d_{i-1}]$ and $(d_{i+1},\, d_i]$, then the evolution can jump from one "state" to another after a finite time. This is why the ratio $L(d_2)/M(d_2) = 9/4 > 8/4$ does not yield a configuration which converges even more slowly than in Theorem 3.3.6. Recall from Table 3.1 that the interval $(d_3,\, d_2]$ is very narrow. One can check that, starting with $d = d_2$, a group of 4 agents will disconnect at $t = 9$, as in Table 3.1, but at this point the gap between the first two remaining agents will be $\frac{13349725}{15971904} \approx 0.8358 < d_3$, which will suffice for the evolution to jump into a new state whereby the next group of 4 agents disconnect after 8 further steps (at $t = 17$). In fact, the system will remain in that state forever, as can be proven by explicit computation and following the proof of Theorem 3.3.6.

In fact, a system can take arbitrarily long to jump from one state to another. To see this, we consider values of d slightly above $d_1 = 72/79$. For $n \geq 0$ let $\boldsymbol{y}_n := \mathcal{T}_0^n \mathbf{1}_\infty$, $\boldsymbol{z}_n := \mathcal{A}^3 \boldsymbol{y}_n$ and $d^{n+1} := \frac{1}{\boldsymbol{z}_n(1)+\boldsymbol{z}_n(2)}$. By Proposition 3.2.3, we know that the sequence $(\boldsymbol{y}_n)$ is monotonically decreasing in l^∞ and converging toward a fixed point $\boldsymbol{y}_\infty$ of $\mathcal{T}_0$. Since, as one can readily check, the matrix $\mathcal{A}^3$ has non-negative entries and row sums at most one, the same is true of the sequence $(\boldsymbol{z}_n)$. Hence, (d^n) is an increasing sequence, starting from $d^1 = d_1 = 72/79$ and converging to a limit $d^\infty = \frac{1}{\boldsymbol{z}_\infty(1)+\boldsymbol{z}_\infty(2)}$, which numerically is about 0.921776.... It is easy to see that, if $d \in (d^{n-1}, d^n]$, then starting from the configuration $\boldsymbol{\mathcal{E}}_{\infty,\,d}$, it will happen n times that 3 agents disconnect after 5 time steps, whereas on the $(n+1)$:st occasion, 3 agents will disconnect after 4 steps instead. Indeed, one can check (though the computations become extremely messy) that the system will then forever remain in the latter state. This leads us to our final question:

Question 3.3.7. *Is it true that, for any $d \in (0, 1]$, the evolution of the system $\boldsymbol{\mathcal{E}}_{\infty,\,d}$ is ultimately periodic, in the sense that both the time between successive disconnections and the number of agents which disconnect are constant from some point onwards?*

Given the ideas introduced in this paper, answering this last question for $d > 1/2$ at least could perhaps be reduced to a finite, if extremely messy computation. However, this would not yield much insight into what happens as $d \to 0$, which is the main theme behind all the questions posed in this section. The proofs in this paper all relied heavily on explicit numerical estimates (as in (3.2.19)-(3.2.21) and (3.3.4)-(3.3.7)). To push the work further, it seems that a deeper qualitative understanding of the evolution of sequences of equally spaced agents and the associated linear operators on l^∞ will be needed.

Bibliography

[1] A. Bhattacharya, M. Braverman, B. Chazelle and H. L. Nguyen, *On the convergence of the Hegselmann-Krause system*, Proceedings of the 4th Innovations in Theoretical Computer Science conference (ICTS 2013), Berkeley CA, January 2013.

[2] B. Chazelle, *The total s-energy of a multiagent system*, SIAM J. Control Optim. **49**, No. 4, 1680–1706 (2011).

[3] R. Hegselmann and U. Krause, *Opinion dynamics and bounded confidence: models, analysis and simulations*, Journal of Artificial Societies and Social Simulation **5**, No. 3. (2002) Fulltext at `http://jasss.soc.surrey.ac.uk/5/3/2/2.pdf`

[4] U. Krause, *Private communication.*

[5] S. Kurz, *How long does it take to consensus in the Hegselmann-Krause model?* Preprint at `http://arxiv.org/abs/1405.5757`

[6] S. Martinez, F. Bullo, J. Cortes and E. Frazzoli, *On Synchronous Robotic Networks - Part II: Time Complexity of Rendezvous and Deployment Algorithms*, IEEE Trans. Automat. Control **52**, No. 12, 2214–2226 (2007).

[7] E. Wedin and P. Hegarty, *A quadratic lower bound for the convergence rate in the one-dimensional Hegselmann-Krause bounded confidence dynamics*. Preprint at `http://arxiv.org/abs/1406.0769`

Chapter 4

Paper III

A quadratic lower bound for the convergence rate in the one-dimensional Hegselmann-Krause bounded confidence dynamics

To appear in Discrete and Computational Geometry. Preprint at `http://arxiv.org/abs/1406.0769`

Abstract

Let $f_k(n)$ be the maximum number of time steps taken to reach equilibrium by a system of n agents obeying the k-dimensional Hegselmann-Krause bounded confidence dynamics. Previously, it was known that $\Omega(n) = f_1(n) = O(n^3)$. Here we show that $f_1(n) = \Omega(n^2)$, which matches the best-known lower bound in all dimensions $k \geq 2$.

4.1 Introduction

The field of opinion dynamics is concerned with how human agents influence one another in forming opinions, say on social and political issues (though in principle on anything). Mathematical modelling in this area has increased rapidly in recent years, as technology has improved the prospects for running computer simulations. Rigorous results remain rare, however, and mainly confined to the simplest properties of the simplest models. One such simple model which has proven immensely popular is the so-called *Hegselmann-Krause bounded confidence model* (HK-model for brevity). It was introduced in [6], though the paper usually cited is [5], which at the time of writing has 935 citations on Google scholar, mostly from non-mathematicians. The model works as follows. We have a finite number n of agents, indexed by the integers $1, 2, \dots, n$. Time is measured discretely and the opinion of agent i at time $t \in \mathbb{N} \cup \{0\}$ is represented by a real number $x_t(i) \in \mathbb{R}$. There is a fixed parameter $r > 0$ such that the dynamics are given by

$$x_{t+1}(i) = \frac{1}{|\mathcal{N}_t(i)|} \sum_{j \in \mathcal{N}_t(i)} x_t(j), \tag{4.1.1}$$

where $\mathcal{N}_t(i) = \{j : |x_t(j) - x_t(i)| \leq r\}$. Thus each agent is only willing to compromise at any time with those whose opinions lie within his so-called *confidence interval*, and he updates to the average of these opinions, including his own. Moreover, the width of this interval, $2r$, is the same for all agents. Since the dynamics are obviously unaffected by rescaling all opinions and the confidence bound r by a common factor, we can assume without loss of generality that $r = 1$.

Two important qualitative features of the HK-model are that agents act synchronously and in a completely deterministic manner. This is in contrast to some other famous opinion dynamics models such as voter models [10] or the Deffuant-Weisbuch model [3]. Its popularity is probably due to the simplicity of its formulation, which nevertheless seems "natural". Mathematically, it is very tantalising. The update rule (4.1.1) is linear, but clearly the transition matrix is in general time-dependent, which is the key point. The HK-model has many elegant features which are still either partly understood or have only been observed in simulations. For a more comprehensive survey of the theoretical challenges, see for example the introduction to [11].

In this paper, we will focus on one particular question which has been the subject of much attention, namely how long it takes for opinions obeying the HK-dynamics to stabilise. First, some notation and terminology. Let $(x(1), \dots, x(n))$ be a configuration of opinions. We say that agents i and j *agree* if $x(i) = x(j)$. A maximal set of agents that agree is called a *cluster*, and the number of agents in a cluster is called its *size*. The configuration is said to be *frozen*[1] if $|x(i) - x(j)| > 1$ whenever $x(i) \neq x(j)$. Clearly, if the configuration is frozen then $x_{t+1}(i) = x_t(i)$ for all i, and it is easy to see that the converse also holds.

Perhaps the most fundamental result about the HK-dynamics is that any configuration of opinions will freeze in a finite number of time steps, which moreover is universally bounded by a function of the number n of agents only. Indeed, the same is true of a wide class of models including HK as a simple prototype, see [2]. Let $f_1(n)$ denote the maximum number of time steps taken to freeze by a configuration of n agents obeying (4.1.1). For the HK-model, the bound given in [2] is $f_1(n) = n^{O(n)}$. However, it is known that $f_1(n)$ is bounded by a polynomial function of n. The first such bound of $O(n^5)$ was established in [9] and the current record is $O(n^3)$, due to [1].

Lower bounds for $f_1(n)$ have received less attention, perhaps due to the difficulty in finding explicit examples of configurations which take a long time to freeze. A natural example to look

[1]Other terms used in the literature are "in equilibrium" or "has converged". We think our term captures the point with the least possible room for misinterpretation, however.

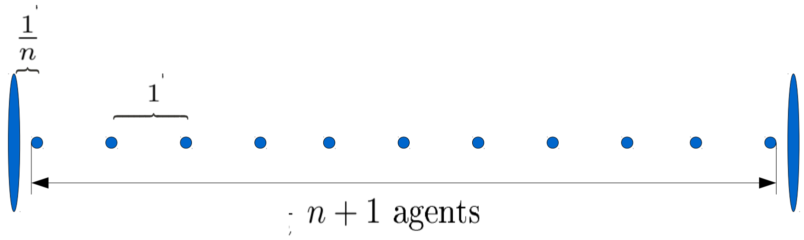

Figure 4.1: Schematic representation of the configuration $\mathcal{D}_n$. Each dumbbell has weight n.

at is the configuration $\mathcal{E}_n = (1, 2, \ldots, n)$, in which opinions are equally spaced with gaps equal to the confidence bound. Thus, agents are placed as far apart as possible to begin with, without being split into two isolated groups. It is not hard to see that, as this configuration updates, if $i < n/2$ then the opinions of agents i and $(n+1) - i$ will remain constant as long as $t < i$, while both will change at $t = i$. Hence, the time taken for the configuration $\mathcal{E}_n$ to freeze is at least $n/2$. In fact, this configuration freezes in time $5n/6 + O(1)$, see [4].

Thus, $f_1(n) = \Omega(n)$, an observation that was already made in [9]. In this paper, we will prove that $f_1(n) = \Omega(n^2)$ by exhibiting an explicit sequence $\mathcal{D}_n$ of configurations which take this long to freeze. In fact, we shall abuse notation slightly. Though we could define a suitable configuration for any number n of agents, in order to simplify the appearance of certain formulas we will assume that n is even and let $\mathcal{D}_n$ denote a certain configuration on $3n + 1$ agents. Our construction basically combines the chain $\mathcal{E}_n$ with an example of Kurz [7], and is defined as follows:

Definition 4.1.1. *Let n be a positive, even integer. The configuration $\mathcal{D}_n$ consists of $3n + 1$ agents whose opinions are given by*

$$x(i) = \begin{cases} -\frac{1}{n}, & \text{if } 1 \le i \le n, \\ i - (n+1), & \text{if } n+1 \le i \le 2n+1, \\ n + \frac{1}{n}, & \text{if } 2n+2 \le i \le 3n+1. \end{cases} \tag{4.1.2}$$

The configuration is represented pictorially in Figure 4.1. It has the shape of a dumbbell. Indeed, someone familiar with the theory of Markov chains might consider this a natural candidate for maximising the freezing time[2]. There is a subtlety, however. Along the "bar" of the dumbbell, opinions are equally spaced at distance one, whereas the two dumbbell clusters themselves are positioned much closer, at distance $1/n$, to the ends of the bar. The latter is what raises the freezing time from $\Theta(n)$ to $\Theta(n^2)$, as will become evident from the proof below. In fact, this is just one of at least three ways of considering our construction as a modification of others previously known which all freeze in linear time. A second way would be to think of it as starting from $\mathcal{E}_n$, which freezes in time $O(n)$, and then adding the dumbbells. A third would be to start from the configuration in [7], which consists of the two dumbbells placed at distance $1/n$ from their respective solitary agents, but then without the long intermediate chain[3]. Kurz

[2] In the general theory of irreducible Markov chains on graphs, dumbbell-like graphs are known to have the longest mixing times. See, for example, [8].

[3] In fact, in the Markov chain literature, this configuration is commonly termed a dumbbell, whereas ours would be referred to as a "dumbbell with a chain in between". We hope the reader is not confused !

showed that his configuration took time $\Omega(n)$ to freeze and as a by-product of our method, it can be easily shown to freeze in time $O(n)$.

Let us now formally state our result.

Theorem 4.1.2. *The configuration $\boldsymbol{\mathcal{D}}_n$ freezes after time $\Omega(n^2)$.*

The proof will be given in the next section. One important feature of our result is that it matches the best-known lower bound for the freezing time of the multi-dimensional HK-model. The latter refers to the fact that rule (4.1.1) makes sense if opinions $x_t(i)$ are considered as vectors in $\mathbb{R}^k$ for any fixed k and neighborhoods $\mathcal{N}_t(i)$ are defined with respect to Euclidean distance. The sociological interpretation would be that there are k "issues", and that agents will compromise if and only if their opinions are sufficiently close on *all* issues. Let $f_k(n)$ denote the maximum number of time steps taken to freeze by a configuration of n agents with opinions in $\mathbb{R}^k$ and obeying (4.1.1). It turns out that $f_k(n)$ is bounded by a universal polynomial function of n and k. This was also established in [1], who gave the bound $f_k(n) = O(n^{10}k^2)$. Note, though, that this is much worse than the best bound $O(n^3)$ in one dimension. Indeed, the proof of the latter in [1] uses a different argument which does not seem to generalise to higher dimensions[4].

Already in two dimensions, however, a quadratic lower bound was also proven in [1]. Their example, which we denote $\boldsymbol{\mathcal{F}}_n$, places the n agents at the vertices of a regular n-gon of side-length one, and they show that the system requires at least $n^2/28$ steps to freeze. [5] The configuration $\boldsymbol{\mathcal{F}}_n$ seems, at least in hindsight, like a natural "two-dimensional version" of $\boldsymbol{\mathcal{E}}_n$. It is not really clear how far one can push this idea, however, as the upper bound of $O(n^{10}k^2)$ for all dimensions makes immediately clear. Indeed, there is no example known in dimensions $k \geq 3$ which takes longer to freeze than $\boldsymbol{\mathcal{F}}_n$, now considered as a configuration on a plane in $\mathbb{R}^k$. The configurations $\boldsymbol{\mathcal{D}}_n$ discussed in this paper are also quite different from the $\boldsymbol{\mathcal{F}}_n$.

We finish this section by giving some more fairly standard terminology to be used below. Let $(x(1), \ldots, x(n))$ be a configuration of one-dimensional opinions, obeying the convention that $x(i) \leq x(j)$ whenever $i \leq j$. We can define a *receptivity graph* G, whose nodes are the n agents and where an edge is placed between agents i and j whenever $|x(i) - x(j)| \leq 1$. We say that agents i and j are *connected* if they are in the same connected component of the receptivity graph. Observe that every connected component of G is an interval of agents and that i is disconnected from $i+1$ if and only if $x(i+1) > x(i) + 1$.

4.2 Proof of Theorem 4.1.2

Lemma 4.2.1. *Let $n \geq 2$ and let $\mathcal{P}_n$ denote the path on n vertices, indexed from left-to-right by the integers $1, \ldots, n$. Let $X_0, X_1, \ldots$ be a random walk on $\mathcal{P}_n$ with transition probabilities $p_{i,j}$ given by*

$$p_{i,j} = \begin{cases} 2/3, & \text{if } (i,j) = (1,1) \text{ or } (n,n), \\ 1/3, & \text{otherwise and if } |i-j| \leq 1, \\ 0, & \text{otherwise.} \end{cases} \tag{4.2.1}$$

[4] An important fact which makes the one-dimensional model much simpler to analyse is that, as soon as an agent becomes isolated, he will remain so forever. This is not always the case in higher dimensions. As an example in $\mathbb{R}^2$, consider three agents a, b, c initially placed at $(0, -0.5)$, $(0, 0.5)$ and $(1, 0)$ respectively. At $t = 0$, only a and b will interact, but this first interaction will bring them both to $(0, 0)$ where they are close enough to c to interact at $t = 1$.

[5] By symmetry, it is clear that all agents will end up in agreement in this case.

For any i, j and $t \geq 0$, let $h_{i,j}(t)$ denote the expected number of times a walk started at i will hit j up to and including time t, i.e.:

$$h_{i,j}(t) = \mathbb{E}[\#s : X_s = j,\ 0 \leq s \leq t \mid X_0 = i].$$

Then $h_{1,1}(t) \leq c_1 \cdot \sqrt{t}$ for all $1 \leq t \leq n^2$, where $c_1 > 0$ is an absolute constant, independent of n.

Proof. This result surely follows from standard textbook facts about random walks on graphs, but since we cannot point to a reference for the precise result, we shall outline a proof in any case.

Let us consider instead a cycle $\mathcal{C}_{2n}$ of length $2n$, with vertices indexed clockwise by $1, 2, \ldots, 2n$, and a random walk on the cycle for which the transition probabilities are $p'_{i,j} = 1/3$ if $|i - j| \pmod{2n} \leq 1$ and $p'_{i,j} = 0$ otherwise. Let $h'_{i,j}(t)$ denote the expected number of times a walk on $\mathcal{C}_{2n}$ started at node i hits node j up to and including time t.

Claim 1: (i) $h'_{1,2n}(t) \leq h'_{1,1}(t)$.
(ii) $h_{1,1}(t) = h'_{1,1}(t) + h'_{1,2n}(t) \leq 2h'_{1,1}(t)$.

To prove (i) first note that, by the symmetry of the transition rules on the cycle, the function $h'_{i,i}(t)$ is independent of i. Let τ be the random time at which a walk started at 1 first hits $2n$. Then

$$h'_{1,2n}(t) = \sum_{s=0}^{t} \mathbb{P}(\tau = s) \cdot h'_{2n,2n}(t-s) = \sum_{s=0}^{t} \mathbb{P}(\tau = s) \cdot h'_{1,1}(t-s) \leq$$
$$\leq \sum_{s=0}^{t} \mathbb{P}(\tau = s) \cdot h'_{1,1}(t) \leq h'_{1,1}(t),$$

where we have used the obvious fact that the functions $h'_{i,j}(t)$ are all non-decreasing in t.

The right-hand inequality in (ii) follows from (i). For the left-hand equality, we identify the nodes of $\mathcal{C}_{2n}$ in pairs as

$$v_1 = \{1, 2n\},\ \ v_2 = \{2, 2n-1\},\ \ \ldots,\ \ v_n = \{n, n+1\}.$$

A random walk on $\mathcal{C}_{2n}$ can be identified with a random walk on the path $\mathcal{P}_n$ whose vertices from left-to-right are $v_1, \ldots, v_n$, where any step in the former which remains inside the same subset v_i is considered as standing still at the same vertex in the latter. It is also easy to see that if the transition probabilities on the cycle are $p'_{i,j}$, then on the path they become $p_{i,j}$. The equality in (ii) follows immediately from these observations.

By Claim 1, it suffices to prove that $h'_{1,1}(t) = O(\sqrt{t})$ for all $1 \leq t \leq n^2$. We go one step further. Let $q(t)$ denote the probability that the walk on $\mathcal{C}_{2n}$, started at node 1, is also at node 1 at time t. By linearity of expectation, it suffices to prove that $q(t) = O(1/\sqrt{t})$ for all $1 \leq t \leq n^2$.

So fix a time $t \geq 1$. Any walk consists of steps of three types: clockwise, anticlockwise and standing still. The walk will be back at node 1 at time t if and only if the numbers of clockwise and anticlockwise steps among the first t steps are congruent modulo $2n$. The expected number of standing still steps is $t/3$ and, up to an error of order $e^{-\alpha t}$, where $\alpha > 0$ is an absolute constant, we can ignore all walks where the number of standing still steps is greater than $t/2$ say. Conditioned on the number l of such steps and their timings, there are 2^{t-l} possible walks. The number of these which have c clockwise steps is $\binom{t-l}{c}$, which is less than $\frac{2^{t-l}}{\sqrt{t-l}}$ for any c and maximised at $c = \lfloor \frac{t-l}{2} \rfloor$. Since we're assuming $l \le t/2$, it follows that every binomial coefficient is less than $2^{t-l}\sqrt{\frac{2}{t}}$. The ones that contribute to $q(t)$ are those such that $2c \equiv t - l \pmod{2n}$. The gap between any two such values of c is at least n which, since $t \le n^2$, is at least $\lceil \sqrt{t}\, \rceil$.

Claim 2: There is a real number $\kappa \in (0, 1)$ such that, for all integers $m \ge 2$ and $r \ge 1$,

$$\binom{m}{\lfloor m/2 \rfloor + r\lceil \sqrt{m}\, \rceil} \le \kappa^r \binom{m}{\lfloor m/2 \rfloor}. \tag{4.2.2}$$

Once again, we will prove this directly, rather than appealing to some textbook fact. For $0 \le k < m$, let $f(m, k) := \binom{m}{k+1}/\binom{m}{k} = \frac{m-k}{k+1}$. The function $f(m, k)$ is decreasing in k as long as $k \ge \lfloor m/2 \rfloor$, thus it suffices to prove (4.2.2) for $r = 1$. If we put $k = \lfloor m/2 \rfloor + \lfloor \frac{1}{2}\sqrt{m} \rfloor$ then, for sufficiently large m, $f(m, k) \le 1 - \frac{1}{\sqrt{m}}$. Thus, for sufficiently large m,

$$\frac{\binom{m}{\lfloor m/2 \rfloor + \lceil \sqrt{m}\, \rceil}}{\binom{m}{\lfloor m/2 \rfloor}} = \prod_{j=1}^{\lceil \sqrt{m}\, \rceil} f(m, \lfloor m/2 \rfloor + j) \le \left(1 - \frac{1}{\sqrt{m}}\right)^{\frac{1}{2}\sqrt{m}} \le e^{-1/2}. \tag{4.2.3}$$

So, for m sufficiently large, (4.2.2) holds with $\kappa = e^{-1/2}$ Hence it holds for some $\kappa < 1$ and all $m \ge 2$, since for all such m, the first quotient in (4.2.3) is strictly less than one. This proves Claim 2.

Claim 2 implies that, conditioned on l, the contributions to $q(t)$ from different values of c decrease exponentially as one moves away from $\lfloor \frac{t-l}{2} \rfloor$, and hence the total contribution is bounded by an absolute constant times the largest one which, as previously stated, is at most $\sqrt{\frac{2}{t}}$. Unwinding our argument, what we have shown is that, provided $1 \le t \le n^2$ and conditioning on the number and timing of all standing still steps up to time t, the probability of the walk being back at node 1 is $O(1/\sqrt{t}) + O(e^{-\alpha t}) = O(1/\sqrt{t})$. Hence, $q(t) = O(1/\sqrt{t})$, as desired. ☐

Lemma 4.2.2. *Let $n \in \mathbb{N}$, $\kappa \in \mathbb{Q}_{>0}$ and, for $t \ge 0$, let $\boldsymbol{\delta}_t = (\delta_{1,t}, \ldots, \delta_{n,t})$ be a sequence of vectors in $\mathbb{Q}^n_{\ge 0}$ defined recursively as follows:*

$$\begin{aligned} \boldsymbol{\delta}_0 &= (0, \ldots, 0), \\ \delta_{1,t+1} &= \kappa + \frac{2}{3}\delta_{1,t} + \frac{1}{3}\delta_{2,t}, \\ \delta_{n,t+1} &= \kappa + \frac{2}{3}\delta_{n,t} + \frac{1}{3}\delta_{n-1,t}, \\ \delta_{i,t} &= \frac{1}{3}\left(\delta_{i-1,t} + \delta_{i,t} + \delta_{i+1,t}\right), \quad \forall\, 2 \le i \le n-1. \end{aligned}$$

Then there is an absolute constant $c_2 > 0$ such that $\delta_{i,t} \le c_2 \cdot \kappa \cdot \sqrt{t}$ for all i and all $t \le n^2$.

Proof. For any t, it is clear that $\delta_{i,t} = \delta_{(n+1)-i,t}$ and that $\delta_{i,t} \ge \delta_{i+1,t}$ for all $i < n/2$. It thus suffices to prove that $\delta_{1,t} = O(\kappa\sqrt{t})$ for all $t \le n^2$.

The recursion can be written in matrix form as

$$\boldsymbol{\delta}_0 = \mathbf{0}, \tag{4.2.4}$$
$$\boldsymbol{\delta}_{t+1} = \boldsymbol{v} + P \cdot \boldsymbol{\delta}_t, \tag{4.2.5}$$

where $\boldsymbol{v} = (\kappa, 0, 0, \ldots, 0, \kappa)^T$ and $P = (p_{i,j})$ is the transition matrix of (4.2.1). It follows easily from (4.2.4) and (4.2.5) that, for any $t > 0$,

$$\boldsymbol{\delta}_t = (I + P + \cdots + P^{t-1})\boldsymbol{v}.$$

Hence,

$$\delta_{1,t} = \kappa \cdot (h_{1,1}(t) + h_{1,n}(t)) \leq 2\kappa \cdot h_{1,1}(t), \tag{4.2.6}$$

where the last inequality can be proven in a similar manner to part (i) of Claim 1 in the proof of Lemma 4.2.1. Hence, Lemma 4.2.2 follows from (4.2.6) and Lemma 4.2.1. $\square$

Proof. of Theorem 4.1.2. For simplicity (see (4.2.7) below), we assume $n \geq 3$. Let $\boldsymbol{x}_0 = \mathcal{D}_n \in \mathbb{R}^{3n+1}$ and for all $t > 0$ let the updates $\boldsymbol{x}_t = (x_t(1), \dots, x_t(3n+1))$ be generated according to (4.1.1). So $\boldsymbol{x}_t$ represents the positions of the agents at time t. We will find it more convenient to work instead with the vectors of gaps $\boldsymbol{y}_t = (y_{0,t}, \dots, y_{n+1,t}) \in \mathbb{R}^{n+2}$ given by

$$y_{i,t} = x_t(n+1+i) - x_t(n+i), \quad 0 \leq i \leq n+1.$$

Observe that $\boldsymbol{y}_0 = \left(\frac{1}{n}, 1, \dots, 1, \frac{1}{n}\right)$. Let G_t denote the receptivity graph at time t. For as long as $G_t = G_0$, it is easily checked that $\boldsymbol{y}_{t+1} = M \cdot \boldsymbol{y}_t$ where $M = (m_{i,j})$ is an $(n+2) \times (n+2)$ matrix whose upper left 2×3 block is

$$\begin{pmatrix} \frac{n}{(n+1)(n+2)} & \frac{1}{n+2} & 0 \\ \frac{n}{n+2} & \frac{2n+1}{3(n+2)} & \frac{1}{3} \end{pmatrix},$$

which is symmetric about its midpoint, i.e.:

$$m_{i,j} = m_{(n+3)-i,\,(n+3)-j}$$

and which, for $3 \leq i \leq n$, satisfies

$$m_{i,j} = \begin{cases} 1/3, & \text{if } |i-j| \leq 1, \\ 0, & \text{otherwise.} \end{cases} \tag{4.2.7}$$

We define auxiliary vectors $\boldsymbol{\delta}_t = (\delta_{0,t}, \dots, \delta_{n+1,t})$ as follows:

$$y_{i,t} =: \frac{1}{n} - \frac{\delta_{i,t}}{n^2}, \quad \text{if } i = 0 \text{ or } i = n+1, \tag{4.2.8}$$

$$y_{i,t} =: 1 - \frac{\delta_{i,t}}{n^2}, \quad \text{for } 1 \leq i \leq n. \tag{4.2.9}$$

Observe that $\boldsymbol{\delta}_0 = \mathbf{0}$ and $\delta_{i,t} = \delta_{(n+1)-i,t}$ for all i and t. As long as $G_t = G_0$ one checks that the following recursion is satisfied:

$$0 \leq \delta_{0,t+1} \leq 1 + \frac{1}{n}\left(\delta_{0,t} + \delta_{1,t}\right), \tag{4.2.10}$$

$$0 \leq \delta_{1,t+1} \leq \delta_{0,t} + \frac{2}{3}\delta_{1,t} + \frac{1}{3}\delta_{2,t}, \tag{4.2.11}$$

$$0 \leq \delta_{i,t+1} = \frac{1}{3}\left(\delta_{i-1,t} + \delta_{i,t} + \delta_{i+1,t}\right) \quad \text{for } 2 \leq i \leq n-1. \tag{4.2.12}$$

Applying Lemma 4.2.2 with $\kappa = 2$ it is easy to deduce that, for some absolute constant $c_3 > 0$ and all $t \leq c_3 \cdot n^2$, the solution to (4.2.10)-(4.2.12) with initial condition $\boldsymbol{\delta}_0 = \mathbf{0}$ will satisfy

$$\delta_{0,t} \leq 2, \quad \delta_{n+1,t} \leq 2, \qquad \delta_{i,t} < n-2 \text{ for } 1 \leq i \leq n.$$

But this in turn implies, from (4.2.8) and (4.2.9), that $y_{i,t} + y_{i+1,t} > 1$ for all $0 \leq i \leq n$ and all $t \leq c_3 \cdot n^2$, hence indeed it is true that $G_t = G_0$ for all such t. In particular, agent $n+2$ will not be visible to the cluster on the left before time $c_3 \cdot n^2$, which proves that the configuration will take at least this long to freeze. $\square$

Remark 4.2.3. *One can prove that the configuration does indeed freeze in time* $\Theta(n^2)$*. First, we can turn the above argument around somewhat and deduce instead from the above relations that* $\delta_{0,t} \geq 1/2$ *for all* $t > 0$ *and hence, instead of (4.2.11), that*

$$\delta_{1,t+1} \geq \frac{1}{4} + \frac{2}{3}\delta_{1,t} + \frac{1}{3}\delta_{2,t}.$$

The argument in Lemma 4.2.2 can then be turned on its head to deduce that $\delta_{1,t} = \Omega(h_{1,1}(t))$*, while it is almost trivial that* $h_{1,1}(t) = \Omega(\frac{t}{n})$*. What all of this implies is that agent* $n+2$ *will indeed become visible to the cluster on the left at time* $t^* = \Theta(n^2)$*, and it will then immediately disconnect from agent* $n+3$*. We then just need to consider the subsequent evolution of the chain* $\mathcal{C}$ *of agents* $n+3, \ldots, 2n-2$*. Since* $\delta_{i,t^*} = O(n)$ *for every* i*, it follows from (4.2.8) and (4.2.9) that the gaps between consecutive agents in* $\mathcal{C}$ *are all greater than* $1 - O(1/n)$*. Hence the chain will freeze in time* $5n/6 + O(1)$*. This last deduction follows from unpublished results in [4], more precisely from Theorem 1.1 and remarks at the outset of Section 3 in that paper.*

Given that the configuration $\boldsymbol{\mathcal{D}}_n$ *freezes in time* $\Theta(n^2)$*, one can try to compute the constant factor accurately. We have not done so, but a combination of simulations and the Ockham's razor principle lead us to believe that the freezing time for* $\boldsymbol{\mathcal{D}}_n$ *is* $(1 + o(1))\frac{n^2}{4}$*. The factor of* $4 = 2^2$ *comes from the fact that the numbers* $\delta_{1,t}$ *in (4.2.9) seem to grow like* $2\sqrt{t}$*.*

Note that, if our hypothesis is correct, then the freezing time of the configuration $\boldsymbol{\mathcal{D}}_n$ *still grows more slowly, at least for* $n \gg 0$*, than that of the two-dimensional configuration* $\boldsymbol{\mathcal{F}}_{3n+1}$*. These are also two quite different types of configurations. It remains unclear what the right estimate for the function* $f_k(n)$ *might be in higher dimensions.*

4.3 Acknowledgements

We thank Sascha Kurz and Anders Martinsson for helpful discussions.

Bibliography

[1] A. Bhattacharya, M. Braverman, B. Chazelle and H. L. Nguyen, *On the convergence of the Hegselmann-Krause system*, Proceedings of the 4th Innovations in Theoretical Computer Science conference (ICTS 2013), Berkeley CA, January 2013.

[2] B. Chazelle, *The total s-energy of a multiagent system*, SIAM J. Control Optim. **49**, No. 4, 1680–1706 (2011).

[3] G. Deffuant, D. Neau, F. Amblard and G. Weisbuch, *Mixing beliefs among interacting agents*, Adv. Complex. Syst. **3**, 87–98, (2000).

[4] P. Hegarty and E. Wedin, *The Hegselmann-Krause dynamics for equally spaced agents.* Preprint at `http://arxiv.org/abs/1406.0819`

[5] R. Hegselmann and U. Krause, *Opinion dynamics and bounded confidence: models, analysis and simulations*, Journal of Artificial Societies and Social Simulation **5**, No. 3, (2002). Fulltext at `http://jasss.soc.surrey.ac.uk/5/3/2/2.pdf`

[6] U. Krause, *Soziale Dynamiken mit vielen Interakteuren, eine Problemskizze,* in: Modellierung und Simulation von Dynamiken mit vielen interagierenden Akteuren, U. Krause and M. Stöckler eds., Universität Bremen (1997).

[7] S. Kurz, *How long does it take to consensus in the Hegselmann-Krause model ?*. Preprint at `http://arxiv.org/abs/1405.5757`

[8] D. A. Levin, Y. Peres and E. L. Wilmer, *Markov chains and mixing times. With a chapter by James G. Propp and David B. Wilson.* American Mathematical Society, RI, xviii + 371 pp, Providence, (2009).

[9] S. Martinez, F. Bullo, J. Cortes and E. Frazzoli, *On Synchronous Robotic Networks - Part II: Time Complexity of Rendezvous and Deployment Algorithms*, IEEE Trans. Automat. Control **52**, No. 12, 2214–2226, (2007).

[10] `http://en.wikipedia.org/wiki/Voter_model`

[11] E. Wedin and P. Hegarty, *The Hegselmann-Krause dynamics for continuous agents and a regular opinion function do not always lead to consensus.* Preprint at `http://arxiv.org/abs/1402.7184`