THESIS FOR THE DOCTORAL DEGREE OF PHILOSOPHY

On the mathematics of the one-dimensional Hegselmann-Krause model

EDVIN WEDIN

UNIVERSITY OF GOTHENBURG

Division of Algebra and Geometry
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
Gothenburg, Sweden 2022

On the mathematics of the one-dimensional Hegselmann-Krause model
Edvin Wedin
Gothenburg 2022
ISBN 978-91-8009-634-8 (PRINT)
ISBN 978-91-8009-635-5 (PDF)
Available at http://hdl.handle.net/2077/70275

Division of Algebra and Geometry
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Sweden
Telephone +46 (0)31 772 1000

Typeset with LaTeX
Printed by Stema
Printed in Borås, Sweden 2022

On the mathematics of the one-dimensional Hegselmann-Krause model

Edvin Wedin

Division of Algebra and Geometry
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg

Abstract

This thesis contains three papers, which all in different ways concern the asymptotics of the Hegselmann-Krause model in opinion dynamics. In this model a set of conformist agents repeatedly and synchronously replace their real-valued opinions by the average of those opinions that are within unit distance of their own. The resulting dynamics become exceedingly subtle, and are sensitive to the precise initial configuration.

In Paper I, my supervisor and I investigate the case when the system is initialised with a chain of equidistant opinions. This includes determining the precise evolution for every (possibly infinite) chain with inter-agent distance 1 or ≈ 0.81.

In Paper II, I expand the current theory for the evolution of systems where a very large number of opinions are independently drawn at random from the uniform distribution on an interval. I develop a method for approximating the updates of configurations with arbitrarily many agents by updates of smaller ones which can be explicitly computed. I use this method to show rigorously for the first time that for some interval lengths, one asymptotically almost surely reaches a consensus. This gives theoretical support for an observation by Lorenz, that sometimes a group that would not otherwise reach a consensus could do so by individually becoming more closed-minded.

In Paper III, we show that if opinions are taken to be points on a circle with perimeter larger than 2 instead of being real numbers, the resulting sequence of updates must converge pointwise from any possible initial configuration.

Keywords: Opinion dynamics, Hegselmann-Krause model, bounded confidence, consensus, piecewise linear dynamics.

List of papers

This doctoral thesis is based on the following papers:

Paper I. P. Hegarty and E. Wedin, *The Hegselmann-Krause dynamics for equally spaced agents*, J. Difference Equ. Appl. **22** (2016), No. 11, 1621–1645.

Paper II. E. Wedin, *A rigorous formulation of and partial results on Lorenz's "consensus strikes back" phenomenon for the Hegselmann-Krause model*, (2021) arXiv: 2107.12906.

Paper III. P. Hegarty, A. Martinsson and E. Wedin, *The Hegselmann-Krause dynamics on the circle converge*, J. Difference Equ. Appl. **22** (2016) No. 11, 1720–1731.

Contributions

Paper I. The idea to work with gaps instead of opinions directly, and to formulate the problem in terms of a single linear operator. Preliminary simulations, including finding the slower convergence at $d \approx 0.81$. Partaking in the development of the proofs. Reworking the paper after review.

Paper II. Choosing the problem, developing the underlying theory and proof strategy, stating and proving all the theorems. Got some help from my supervisor in writing the paper.

Paper III. The idea to work with inter-agent distances on the circle. Discussing the problem.

Papers not included in the thesis

E. Wedin and P. Hegarty, *A quadratic lower bound for the convergence rate in the one-dimensional Hegselmann-Krause bounded confidence dynamics*, Discrete Comput. Geom. **53** (2015), No. 2, 478–486.

E. Wedin and P. Hegarty, *The Hegselmann-Krause dynamics for the continuous-agent model and a regular opinion function do not always lead to consensus*, IEEE Trans. Automat. Control **60** (2015), No. 9, 2416–2421.

P. Hegarty, A. Martinsson, E. Wedin, *The "No Justice in the Universe" phenomenon: why honesty of effort may not be rewarded in tournaments*, (2018) arXiv: 1803.00823.

Acknowledgements

The list of people who in one way or another deserve to be acknowledged for supporting me in my getting here is long, too long for all the names to fit within these pages and for none to be forgotten.

At the top of the list, without question, sits my supervisor Peter Hegarty, who has guided me through the world of research since I was a master student in 2013. For his tireless listening to new ideas, even those that didn't work out, his advice and for *always* finding time for me. For teaching me about mathematical rigour, about the world of academia, to keep my mathematical mind open and to never trust the obvious. In short, for teaching me how to be a mathematician.

There are a few other people at the department who deserve acknowledgement for taking part in my mentoring: Anders M., most obviously for being a co-author of Paper III, but also for sharing his experiences as a senior PhD student, always enjoying a nice problem and being a great friend. Johan W., my co-supervisor, for valuable advice and for always having good input for seemingly every math problem. Marija, for providing rock solid support and valuable guidance when I needed it. Olle H., for discussions and advice about all sorts of subjects and for valuing my input on important questions.

In addition to mentors, PhD students need fast friends. I'd like to acknowledge: Linnea and Hanna, for sharing laughter, crying, panic, food, dancing, deep facts of life, early mornings, late nights, more food, exhaustion and, through it all, friendship. Jiacheng, for fun, adventure, friendship and trust. Damiano, for being by my side since the beginning.

MV has been a great department, and there are many people here who should be acknowledged for helping me in one way or the other. They have provided motivation, senior advice, friendship, encouragement, help with the practicalities of completing a PhD, and much more. Some of them are, in no particular order: Åse, Johannes, the rest of the Doktorand Band, Helga, Kristian, Frida, Carl-Joar, Christoffer, Olle E, Oskar A, Malin PF, Niklas A, Tom, Marvin, Anna J, Elin, Jeff, Aila, Genkai, Mohammad, Ulla, Umberto, Moritz, Serik, Martin R, Fia, Lotta, Marie, Jeanette and Jovan.

I would also like to acknowledge some people outside academia who nevertheless have helped me get where I am: Karolina and Sandra, for everyday life, comfort in times of despair, love, adventure, validation, fun and understanding. Adam, Erik, Leo and Tommy, my co-jesters, for countless wonderful adventures, and for being the best friends anyone could wish for. Tomas, for giving me a home when I needed one and for proving life can be simple.

Finally, I'd like to acknowledge my family members: Ingrid and Jan, my parents, for raising me to be curious and creative, and not to lose hope. Agne, my grandfather, for always listening and always believing in me. Signe, for being my sister and letting me be her bror.

I am thankful to you all.

Contents

Ancillary files containing the computer code used to run the simulations in Paper II can be found at `https://arxiv.org/abs/2107.12906`

1 Introduction

Scientists use mathematical models to study all kinds of phenomena. Sometimes when a mathematical model is devised, it turns out to be interesting beyond what it can teach us about the phenomenon it was originally meant to model. Famous examples where mathematicians have taken great interest in a model for its own sake, rather than because of some application, include various percolation models, the Ising model and the logistic equation from population dynamics.

A less famous example, which will be the subject of this thesis, is an agent based model of opinion dynamics published by Hegselmann and Krause [6] in 2002. In their model, each agent holds an opinion which is repeatedly updated to conform to those of the other agents. The point is that the agents use what the authors refer to as *bounded confidence*: Any opinion too different from your own is disregarded, so the confidence you put in other people's opinions is bounded.

Their original paper contains several variations, but the definition of the Hegselmann-Krause (HK) model that will be used throughout this thesis, and which is one of the most common, is the following.[1]

Definition 1.0.1. *Let n agents be indexed 1, 2, ..., n, and let each agent i hold the opinion $f(i) \in \mathbb{R}$. The function f is referred to as an* opinion profile *and is without loss of generality assumed to be non-decreasing.*

The opinion profile f is then updated *using the* updating operator U*, which is defined as follows.*

$$Uf(i) = \frac{1}{\#(\mathcal{N}_i(f))} \sum_{j \in \mathcal{N}_i} f(j) \tag{1.1}$$

[1]There are a few different conventions regarding notation, which is reflected in the papers of this thesis. The notation used above agrees with that in Paper II, the most recent paper, but the basic model is defined the same in all three papers.

where $\mathcal{N}_i(f) = \{j \in \{1, 2, ..., n\} : |f(i) - f(j)| \le r\}$ *and* $r = 1$.

We will refer to the evolution of a profile under repeated applications of U as the HK-dynamics.

In words, each update U moves each agent to the average taken over all agents within unit distance. The reason for fixing $r = 1$ is that any result we can show for another value of r would hold for this normalised version if profiles are re-scaled by a factor r. Given an agent i, the agents in $\mathcal{N}_i(f)$ are called the *neighbours of* i, and it's sometimes useful to think of the agents as nodes in a *neighbour graph* with edges between neighbouring agents. For instance one can fairly easily verify that if the neighbour graph at any point disconnects the two parts can never reconnect.

There are many variations of the HK-model. Examples include, but are not limited to the following.

- A version commonly referred to as the *heterogeneous HK-model*, where agents have individual confidence bounds r_i. This was discussed already in [6] and was further investigated in [11].
- A continuous-time version, where agents continuously change their opinion towards the average of their neighbours.[4]
- A version where neighbouring opinions are weighted, depending on how far away they are. This was discussed in [8].
- A version where some agents do not follow the dynamics, but can freely choose their opinion in every step. This gives rise to a control problem of how to best direct the so called *tactical agents* to achieve various goals.[7]

For the rest of this thesis we will, however, contend ourselves with the unadorned model given in Definition 1.0.1. There are at least two reasons for this. First, this is the version on which the largest amount of mathematical work has been done. Perhaps with the exception of the continuous-time variant, the versions listed above are not as well understood as what I here call the classical model. Second, the complexity of the classical model and the wealth of phenomena that emerge from its definition are surely enough to fill several theses on their own.

As stated above, the HK-model has received considerable interest, and at the time of writing Google Scholar lists about 3000 citations. Most of these are from the engineering fields, but in the twenty years since it's introduction a fair body of mathematical results about the model have accumulated. For some intuition as to why this might be interesting for a mathematician, consider the following.

On the one hand, the opinions after the update are linear combinations of those beforehand, so, encoding the opinions as a vector x, computing the updated opinions is as easy as multiplying x by some matrix A (from the left, say). Examining (1.1), we also see that A is row stochastic, and the dynamics may thus be described by a growing product of stochastic matrices. This is nice. There is theory for those.

On the other hand, the matrix A depends on the vector x, and often changes between updates. This is not nice. Say we would write the matrix that corresponds to the operator U, and updates the vector x, as $A(x)$. We then find that the corresponding matrix for U^2 to be $A(A(x)x)A(x)$, the one corresponding to U^3 to be $A(A(A(x)x)A(x)x)A(A(x)x)A(x)$ and so on, with exponentially longer expressions for each step. For some explicit opinion vector x and a modest number of updates this can be done with a computer, but for general x this is harder to handle.

Put more directly: U and its powers are *piecewise linear operators*. Another example of piecewise linear operators, which has received a tremendous amount of attention in recent years, are *ReLU neural networks*. As is the case with the neural networks, as the dimension of the input (and output) vectors grow, analysis of the system becomes highly nontrivial. There turns out to be a balance though: The rule in (1.1) gives rise to surprisingly complex dynamics, but is simple enough to admit some analysis. This makes the dynamics interesting from a mathematical point of view.

One important set of questions about the HK-dynamics, which all of the papers in this thesis are concerned with in one way or another, is that of describing its asymptotic behaviour. There are at least two limits to consider, corresponding to letting the the number of time steps or the number of agents go to infinity, respectively.

The former of these amounts to defining the operator $U^\infty = \lim_{t\to\infty} U^t$ and studying its properties. Already in [6] it was shown that not only does the sequence converge when the operator is defined for any finite number n of agents, but that for each profile f there exists a T such that $U^{T+t}f = U^T f$ for any positive t. The lowest such T is called the *freezing time* of the profile f, and the profile $U^T f$ is said to be *frozen*. It is fairly straightforward to show that a frozen profile must consist of a set of *clusters*, i.e. sets of agents that share the same opinion, and that the opinions of clusters must be separated by more than r. When all agents in a frozen profile comprise a single cluster we say that the profile has *reached a consensus*, and if they are split up into more than one cluster we say that it has *fragmented*. Characterising what frozen profiles result under different initial conditions and what their freezing times are constitute a majority of the work done on the HK-model to date. Papers I and II are both concerned with determining what different kinds of profiles will look like after they freeze, and Paper III establishes this kind of convergence in a version of the model where opinions are not real numbers but points on a circle.

The latter form of convergence, where the number of agents goes to infinity, has also been studied and is, in somewhat different ways, the subject of Papers I and II.

One important tool in studying this kind of limit is *the continuous agent HK-model*, introduced in [2]. It generalises the traditional model by allowing for an uncountable number of agents, indexed by the unit interval $[0,\,1]$ instead of by a set of natural numbers.

Definition 1.0.2. *Let the set of agents be indexed by the interval* $[0,\,1]$ *and let each agent* $\alpha \in [0,\,1]$ *hold the opinion* $f(\alpha) \in \mathbb{R}$*, where* f *is a non-decreasing function. The function* f *is referred to as a* (continuous-agent) opinion profile. *(Note that* f *itself is not assumed to be a continuous function.)*

The updating operator U*, when applied to an opinion profile, is defined as follows.*

$$Uf(\alpha) = \frac{1}{|\mathcal{N}_\alpha(f)|} \int_{\mathcal{N}_\alpha(f)} f(\beta)\, d\beta \tag{1.2}$$

where $\mathcal{N}_\alpha(f) = \{\beta \in [0,\,1] : |f(\alpha) - f(\beta)| \leq 1\}$.

This continuous agent model is truly a generalisation of the traditional model, in the sense that the latter may be emulated within the former by dividing the unit interval into n segments of equal lengths and considering profiles that are constant on each interval. It's easy to check that the resulting dynamics are totally equivalent to (1.1).

As for more general profiles in the continuous agent model, they can, however, be much harder to analyse. Even if we start with a function as simple as a linear one, each update breaks it up into piecewise smooth pieces. Writing up expressions for each piece in terms of elementary functions becomes impossible after only a few updates, and the number of pieces grows exponentially with the number of updates. As opposed to the discrete case, there is no guarantee a continuous profile will ever freeze, and in several interesting cases we can actually show that it does not.

The difficulties just laid out buy us something though. Two key advantages of working with the continuous agent model are

- it makes limits of profiles with an increasing number of agents easy to define
- any distribution of agents can be approximated by a series of step functions to arbitrary precision.

The continuous agent model also allows us to work with (random) distributions of agents directly, without taking the route through looking at particular samples. Informally, the density of agents or the "popularity" of a certain opinion then relates to the inverse of the derivative of the profile, so that the steeper the profile is at an agent the less popular is its opinion. In the extreme case, atoms in the distribution correspond to constant segments in the profile, while discontinuities in the profile describe gaps in the support of the distribution.

In many applications we are interested in distributions without atoms but with connected support. Following [2] we refer to such distributions as *regular*, and they turn out to have nice properties.

Definition 1.0.3. *A non-decreasing continuous agent profile f is said to be* regular *if there exist strictly positive real numbers m and M, such that*

$$m \leq \frac{|f(\alpha) - f(\beta)|}{|\alpha - \beta|} \leq M$$

for any distinct $\alpha,\ \beta \in [0,\ 1]$.

The term $(m,\ M)$-regular is used when there are specific numbers m and M which satisfy the above inequality.

The easiest way to think about this definition is to take $(m,\ M)$-regularity to mean that the profile f is a continuous function, and that the derivative f', wherever it is defined, is contained in the interval $[m,\ M]$.

It turns out that regularity is often preserved under updates.

Theorem 1.0.4. *([2]) Let f be a regular profile such that $f(1) - f(0) \geq 2$. Then Uf is regular.*

In many cases of interest, it turns out that $f(1) - f(0) < 2$ must lead to a consensus. Hence, regular profiles are quite nice to work with since the theorem implies that, if the range of opinions remains above 2, we can assume $U^t f$ is regular for all t. This is explored further in Paper II.

This is the background we need to discuss some important results and open problems, and to give short summaries of the papers. We will do this in the following sections, but before doing so we'll take detour by a softer question.

1.1 What does the HK-model model?

I want to take a short moment to discuss what kind of process the HK-model could be considered a model *of*.

In their original paper [6], Hegselmann and Krause ask the reader to imagine a group of experts tasked to merge their respective estimates of the world population at some particular time in the future. There are two aspects of this picture that might mislead one's intuition when thinking about the HK-model. First, experts who make estimates typically do so, at least in part, based on data and knowledge. Thus, while the opinions of their peers might influence their own, there are also external influences. Second, while the experts in the scenario are tasked to merge their estimates, each expert has an interest in keeping the result as close as possible to that of their own original estimate, rather than to keep in line with the others.

A first observation about the agents in the HK-model is that they, in contrast to the above scenario, don't base their updated opinions on anything except the current opinions of the other agents. In that sense, the agents are *perfect conformists*.

A second observation is that the agents have perfect knowledge of each other's position. The "opinions" in question must thus be openly shared by each agent in an honest way.

Thus, I here suggest it might make more sense to think of agents conforming their *expressions* rather than their actual opinions.

To illustrate what this means, we consider a real life field that better corresponds to the premise of the HK-model, namely that of fashion. Many aspects of how people dress are completely arbitrary, precisely controllable and visible to others. Imagine wearing a tie and meeting with a new group for the first time, only to find everyone else wears their ties slightly shorter than yours. Being a conformist, you "correct" your tie before the next meeting. But only up to a point: if most of the others wear their ties 10 cm long, or, for that matter, down to their knees, you might simply disregard these people as weirdos, and limit your conforming to reasonable people.

Though the choice of clothing, say, is subject to certain practical constraints, and though real humans are not perfect conformists, I thus argue that the HK-model better describes a population of conformists than a population of compromisers.

2 Important results and open problems

While some basic properties of the HK-model were presented in the previous section, this section aims to give a sense of the kinds of questions about the model that have attracted interest.

2.1 Maximal freezing time

Recall that, given a profile f, the freezing time of f is defined as the smallest natural number T such that $U^{T+t}f = U^T f$ for all t. The existence of this number T was proven already in [6] for all discrete profiles, and it even turns out that it can be bounded uniformly for all profiles with the same number n of agents. Let $F(n)$ denote this bound. The following natural question follows:

Question 2.1.1. *How quickly does $F(n)$ grow with n?*

Upper bounds on the growth of $F(n)$ have been found by several authors. In 2007 it was shown in [10] that $F(n) = O(n^5)$, and in 2013 the current record was proven independently in [12] and [1], using different techniques:

Theorem 2.1.2. $F(n) = O(n^3)$.

In trying to find a lower bound on $F(n)$, a natural first attempt is to simply spread n opinions out as much as possible, while keeping the neighbour graph connected. This approach was taken at least as early as in [10], and is achieved by placing the agents at n consecutive integers. The first time we update this profile, all but the two agents at the very ends will find themselves in a perfectly symmetric neighbourhood and will thus not move. At the second update, the penultimate agent at either side finds exactly one neighbour has moved, and will thus move as well while the remaining $n-4$ agents in the middle stay put. This continues for $\lfloor \frac{n}{2} \rfloor$ updates as the "wave" of movement travels from the edges to the centre, and we conclude $F(n) \geq \lfloor \frac{n}{2} \rfloor$, hence, in particular, that $F(n) = \Omega(n)$. In Paper I, we analyse this and other similar, equally spaced "chains" and determine precisely their evolution. We thereby improved the constant $\frac{1}{2}$ and, for a few months in 2014, the best known lower bound was $F(n) \geq n$.

Later in 2014, Hegarty and I proved the following bound.[14]

Theorem 2.1.3. $F(n) = \Omega(n^2)$.

The slow convergence is achieved by augmenting the "chain" of n agents outlined above with two large clumps of n agents each at distance $\frac{1}{n}$ from either end, and thereby obtaining something resembling a "dumbbell". The agents in the clumps only have one neighbour from the rest of the chain, and with one single neighbour at distance $\frac{1}{n}$ and with $n+1$ in total they will move about $\frac{1}{n^2}$ in the first update. It turns out this structure is stable enough to derive the above result.

Finding the real growth rate of $F(n)$ is still an open problem. After an extensive but fruitless search for a construction that takes even longer to freeze, I dare state the following conjecture.

Conjecture 2.1.4. $F(n) = \Theta(n^2)$.

Before closing this subsection, it should be noted that the freezing rates given by $F(n)$ are upper bounds only. The constructions given are specifically constructed to be slow and are sensitive to the exact placements of the agents. For comparison, just drawing agents independently at random from a large interval seems to result in freezing times growing like $O(\log(\log(n))$, though this has not been investigated in much depth.

2.2 Convergence of profiles in the continuous-agent model

In contrast to the classical HK-model, even proving convergence of its continuous-agent counterpart remains an open problem in general. Though there are some partial results, most of them published in [2] and [3], the following statement remains a conjecture.

Conjecture 2.2.1. *Let f be a continuous-agent profile. Then the pointwise limit $U^\infty f = \lim_{t\to\infty} U^t f$ exists.*

One fact which contributes significantly to the difficulty in proving Conjecture 2.2.1 is the following observation, made in [2].

Theorem 2.2.2. *Let f be a regular continuous-agent profile such that $U^t f(1) - U^t f(0) > 2$ for every time t. Then $U^t f \neq U^{t+1} f$ for every t. In other words, f never freezes.*

This basically means that the only hope we have for a continuous-agent profile to freeze in finite time is for its span of opinions to shrink below 2, which in practice often equates to reaching a consensus. To date, very few continuous continuous-agent profiles are known to converge to a fragmentation. The only known class of examples which has been proven to do so were published in [5] in 2014.

Theorem 2.2.3. *There exists a regular profile f such that*

- *$U^\infty f$ exists,*
- *$U^\infty f$ is not constant.*

In fact, we have $U^\infty f(1) - U^\infty f(0) > 2$.

The result closest to proving Conjecture 2.2.1 was presented already in [3]. To state it formally, we need to introduce an alternative notion of convergence.

Definition 2.2.4. *For a profile f and $\varepsilon > 0$, we define the the open ball*

$$B_\mu(f, \varepsilon) = \{g \in \mathcal{O} : \; m(\{\alpha \in [0, 1] : |g(\alpha) - f(\alpha)| \geq \varepsilon\}) < \varepsilon\}, \tag{2.1}$$

where m denotes Lebesgue measure and $\mathcal{O}$ denotes the set of opinion profiles. We write $\to_\mu$ to denote convergence in the topology induced by these balls.

In words, two profiles are considered close if the set of agents whose opinions differ by more than a tiny amount is tiny.

Theorem 2.2.5. *([3]) Denote by $\mathcal{F}$ the set of frozen profiles where the distance between any two clusters is larger than 1 and let f be a profile.*

Then $U^t f - U^{t+1} f \to_\mu 0$ and $U^t f \to_\mu \mathcal{F}$.

Informally, though we cannot prove that the sequence $U^t f$ converges pointwise, it converges to the set of frozen profiles.

Despite the lack of a proof, all numerical and heuristic evidence so far supports Conjecture 2.2.1, and there is, to my knowledge, nothing that suggests it would be false. Any example of a continuous-agent profile that fails to converge pointwise to some limiting profile would be truly interesting!

2.3 Stability of frozen profiles

Though we cannot be certain the continuous-agent HK-model converges for general initial profiles, in [2] the authors manage to show that *if* f is regular, *if* $U^\infty f$ exists and *if* $U^\infty f(1) - U^\infty f(0) > 2$, then $U^\infty f$ must be what they call *stable*.

I will refer to their notion of stability as *BHT stability*.

Definition 2.3.1. *Let f be a frozen continuous-agent profile. We say that f is* BHT-stable *if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that the following holds:*

Take the profile f, add a cluster of weight at most δ and renormalise all weights to give a profile g. Then $U^\infty g$ consists of the same number of clusters as f, and their weights and positions are all within ε of those of f.

In addition to the just given definition, I here suggest a stronger version of stability.

Definition 2.3.2. *Let f be a frozen continuous-agent profile. f is said to be* strongly stable *if*

$$g_k \to_\mu f \implies U^\infty g_k \to_\mu f \tag{2.2}$$

for any sequence $(g_k)_{k=1}^\infty$ of profiles.

In words, a frozen profile is BHT stable if you cannot qualitatively change the profile by inserting a single additional cluster of arbitrarily small weight and again update until the profile freezes. In contrast, a profile is strongly stable if we could not qualitatively change it by adding *any* configuration of agents in the spaces between clusters, provided their total mass is small. It is clear that strong stability implies BHT stability.

An example of a profile that is not BHT stable is given by

$$f(\alpha) = \begin{cases} -1 & \text{if } 0 \leq \alpha \leq \frac{1}{2}, \\ 1 & \text{if } \frac{1}{2} < \alpha \leq 1. \end{cases} \tag{2.3}$$

f is clearly frozen. However, if we inserted a cluster of weight ε with opinion 0 this changes. Applying (1.2), we see that the agents with opinion 1 would move to $\frac{\frac{1}{2}}{\frac{1}{2}+\varepsilon} = \frac{1}{1+2\varepsilon}$, a change of about 2ε. The central cluster will remain at its equilibrium. After somewhere around $\frac{1}{\varepsilon}$ additional updates, the two large clusters will come within $\frac{1}{2}$ of the middle and become neighbours, and the next update will bring the profile to a consensus. If the two clusters would instead be separated by some distance larger than 2, any additional cluster will be in sight of at most one of the large clusters, and may only affect it once before being absorbed into it.

There is a simple test to determine whether a frozen profile is BHT stable or not, given in [2]. Essentially, a frozen profile is stable if no two of its clusters lie too close to each other.

Theorem 2.3.3. *Let f be a frozen continuous-agent profile. Then the following two statements are equivalent.*

- *f is BHT stable.*
- *Choose any two clusters of f, denote their opinions by a and b and let w_a and w_b denote their weights. Then $|a-b| > 1 + \max\left(\frac{w_a}{w_b}, \frac{w_b}{w_a}\right)$.*

Assume now f is a regular profile. By Theorem 1.0.4, updates $U^t f$ remain regular as long as the range of opinions stays above 2. So intuitively, either $U^\infty f$ is a consensus or there must always be at least some agents holding opinions between any forming clusters. It thus seems intuitive that failing to be BHT stable would disqualify any profile from being an asymptotic limit of regular profiles, and the authors of [3] indeed go on to prove the following.

Theorem 2.3.4. *If $U^\infty f$ exists with $U^\infty f(1) - U^\infty f(0) > 2$, it must be BHT stable.*

This is as far as we get, though. If we are trying to prove convergence, the implication points the wrong way: it doesn't let us deduce that being sufficiently close to a BHT stable frozen profile will guarantee that we will converge towards something like it.

For this we would need something like the strong stability in Definition 2.3.2. Almost by definition, we have the following theorem.

Theorem 2.3.5. *Let f be a regular profile. Then for every $\varepsilon > 0$ there exists $\delta > 0$ such that, if $g \in B_\mu(f, \delta)$ is a regular profile then $U^\infty g \in B_\mu(f, \varepsilon)$.*

There is no example in the literature of a profile shown to be BHT stable but not strongly stable, and despite an extensive search I have not managed to produce any such examples myself. I thus pose the following conjecture.

Conjecture 2.3.6. *A frozen profile f is strongly stable if and only if it is BHT stable.*

Together with Theorem 2.2.5 and Theorem 2.3.4, proving this would prove Conjecture 2.2.1 for regular profiles.

2.4 Fragmentation of linear profiles

One question which has received much attention is the following.

Question 2.4.1. *If $f_L(\alpha) = L\alpha$ is a linear continuous-agent profile with opinions spanning $[0, L]$, does $U^\infty f_L$ exist and, if so, what does it look like?*

Even before the introduction of the continuous-agent model, several papers considered various discrete approximations of linear profiles in the form of *equidistant profiles*, notably [6] and [9]. It seems from simulations of equidistant profiles that linear profiles with a large L tend to fragment somewhat regularly into a sequence of no more than $\frac{L}{2}$ clusters,[1] but despite being the subject of several publications nothing is proven about the fate of linear profiles, except for the partial results for a few small values of L given in Paper II.

One fairly modest rigorous formulation of a precise conjecture in line with our current understanding of simulations made so far is the following.

Conjecture 2.4.2. *Let f_L be the linear profile with opinions from 0 to L. Then $U^\infty f_L$ exists and all but $O(1)$ of the distances between consecutive clusters in $U^\infty f_L$ are larger than 2.*

The motivation for the caveat allowing some distances smaller than 2 is that such small distances are observed in simulations at the centre of the profile,

[1] In the literature one encounters the phrase "$2r$-conjecture" in reference to this phenomenon. As far as we know, no precise formulation of an actual conjecture appears in the literature.

especially when L is small. These seem to be a consequence of symmetry, and how the two waves[2] of information from the edges collide.

We are still far from proving anything like Conjecture 2.4.2 though, and perhaps the clearest way to illustrate precisely how far is to note that the following problem is still wide open.

Conjecture 2.4.3. *If L is large enough, $U^\infty f^L$ is not a consensus.*

A first step might be to try and prove the following even weaker conjecture.

Conjecture 2.4.4. *There is some L such that $U^\infty f^L$ is not a consensus.*

It seems we need some new idea, or more likely *ideas*, before Question 2.4.1 can be answered with mathematical certainty.

[2] As described in Subsection 2.1

3 Summaries of papers

3.1 Paper I

The research leading up to Paper I started with the following observation.

Observation 3.1.1. When n HK-agents are initiated with the integer opinions 1, 2, ..., n, they seem to evolve in a very regular manner: After every 5 steps, 3 agents disconnect from the rest at either side and reach local consensus, in the way shown in Figure 3.1. This seems to go on until the profile freezes.

In the paper, we determine the the precise evolution of the neighbour graph of such profiles.

Theorem 3.1.2. *Let $n \geq 2$ be an integer, and write $n = 6k + l$ where $0 \leq l \leq 5$. Let $\xi_{1,n}$ denote the identity function on $1, 2, ..., n$, i.e. a profile with n agents holding consecutive integer opinions. If we compute the sequence of updates $U^t \xi_{1,n}$ the following occurs:*

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

(ii) the final, frozen profile, 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 profile will freeze at time $t = 5k + \epsilon(l)$, where

$$\epsilon(l) = \begin{cases} l - 1, & \text{if } l \in \{2, 3\}, \\ l, & \text{if } l = 1, \\ l + 1, & \text{if } l \in \{0, 4, 5\}. \end{cases} \tag{3.1}$$

The first basic idea behind the proof is to shift focus from the actual opinions

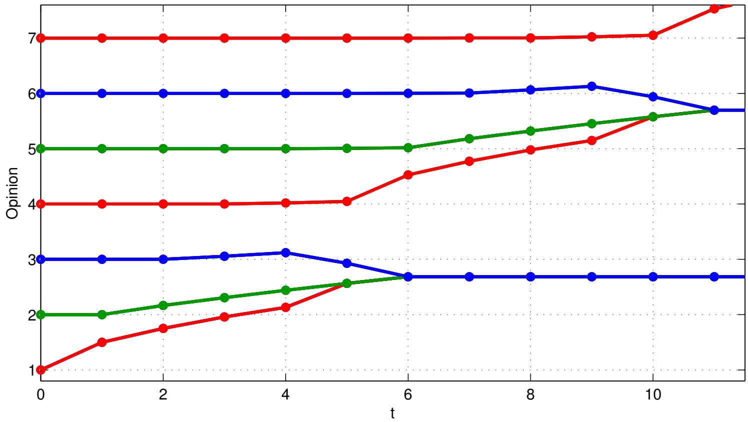

Figure 3.1: The first steps in the evolution of a large number of agents initiated with consecutive integer opinions.

of the agents to the gaps between them. We then note that we can describe the updates in terms of multiplication of vectors of gaps with state-dependent matrices, in the same way we described in the introduction for vectors of opinions.

The second key idea is to combine several of these matrices with a shift, to obtain a single operator that updates the profile until the first disconnection occurs and then removes the disconnected agents. The point is that the "edge" of what is left after the disconnection looks approximately like that of the previous profile. We show that the above combined operator indeed has an attractive fixed point, and that the edges remaining after each disconnection must converge to look the same.

Apart from being interesting in itself, Theorem 3.1.2 relates to the following basic result using integer valued profiles, which was mentioned in Subsection 2.1 and first observed in [10].

Observation 3.1.3. In a profile with initially integer opinions 1, 2, ..., n, agents t, $t+1$, ..., $n+1-t$ will keep their initial opinions for at least $t-1$ updates. This gives a linear lower bound $\lfloor \frac{n}{2} \rfloor$ for the slowest freezing time for a profile of n agents.

When Paper I was written, this was the state of the art for lower bounds on the maximum freezing time, but Theorem 3.1.2 improves the bound from $\frac{n}{2}$ to $\frac{5n}{6} + O(1)$.

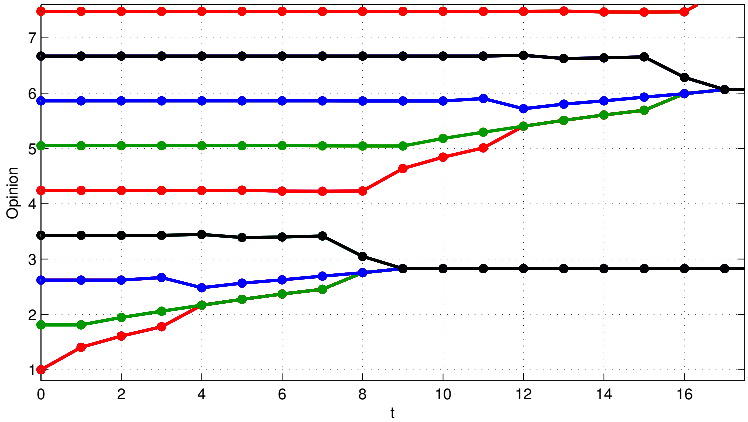

Figure 3.2: The first steps in the evolution of a large number of agents initiated with opinions at integer multiples of 0.81.

We then go on to consider more general equidistant profiles, where the initial opinions are consecutive integer multiples of some $d \in (0, 1]$. In particular, we determine the precise evolution of the neighbour graph up until the first time it disconnects for every $d \in (\frac{1}{2}, 1]$. We find that not every d results in periodic behaviour, but that every value of d that we examine seems to eventually settle into periodic evolution. We then go on to show the following.

Theorem 3.1.4. *For $n \in \mathbb{N}$ and $d \in \mathbb{R}_{>0}$, define $\xi_{d,n} : \{1, 2, ..., n\} \to \mathbb{R}$ with $\xi_{d,n}(i) = di$. $\xi_{d,n}$ is thus a profile with n agents holding opinions that are integer multiples of d.*

Then, there exists a non-empty open subinterval $I \subseteq (\frac{4}{5}, \frac{31104}{36979}]$ such that for any $d \in I$, as we compute the sequence of updates $U^t\xi_{d,n}$ the following occurs:

(i) After every eighth time step, a group of four agents disconnects from either end of the neighbour graph.

(ii) There are positive constants C, C' such that the evolution described in (i) continues until there are at most C agents left in the middle. These will collapse to a cluster after at most C' further time steps.

This further raises the lower bound on the maximum freezing time of a profile of n agents to $n + O(1)$.[1]

[1]This lower bound was not long lived, and we soon improved it to $O(n^2)$ in [14].

The poses paper a number of questions about equidistant profiles, all of which remain open.

Paper I also contains a brief discussion about the relationship between random profiles, equidistant profiles and linear continuous profiles. The ideas outlined in this particular discussion could be considered the basic idea behind Paper II.

3.2 Paper II

This paper also concerns equidistant profiles, but rather than fixing an inter-agent distance d and determining what happens as we let the span of opinions L go to infinity, we fix L and let $d \to 0$. I develop theory building on that in [3] to relate the updates of a continuous-agent linear f_L to those of fine equidistant discrete approximations $f_{n,L}$ of f_L, and also to random profiles f_L^n where opinions are drawn uniformly and independently at random from the interval $[0,\, L]$. In particular I show that

$$\begin{array}{rcrcr} U^\infty f_L \text{ is a consensus} & \Longleftrightarrow & U^\infty f_{n,L} \text{ is a consensus} & \Longleftrightarrow & U^\infty f_L^n \text{ is a consensus} \\ & & \text{for all } n \text{ sufficiently} & & \text{asymptotically} \\ & & \text{large} & & \text{almost surely.} \end{array} \tag{3.2}$$

I then apply this theory to address the following observation of Jan Lorenz, paraphrased from [9].

Observation 3.2.1. Fix a large integer m, let $f_{n,L}$ denote a discrete profile with $n = \lceil Lm \rceil$ agents equidistantly spanning the interval $[0,\, L]$, and consider the profiles $U^\infty f_{n,L}$ for increasing values of L. Surprisingly, simulations suggest the number of clusters in $U^\infty f_{n,L}$ does not increase monotonically in L. In particular, there seems to be some $L_1 \approx 5.23$, $L_2 \approx 5.67$ and $L_3 \approx 6.84$ such that

- $U^\infty f_{n,L}$ is a consensus for $L < L_1$.
- $U^\infty f_{n,L}$ consists of 3 clusters for $L_1 < L < L_2$.
- $U^\infty f_{n,L}$ is a consensus for $L_2 < L < L_3$.

In the words of Lorenz: As L increases, consensus strikes back!

Although the observation still is not fully explained, I manage to obtain the following partial result.

Theorem 3.2.2. *Denote by $C_{Eq}(L, n)$ the final number of clusters in $U^\infty f_{n,L}$, and let*

$$C_{Eq}(L) = \lim_{n\to\infty} C_{Eq}(L,\, n),$$

whenever the limit exists, i.e. when $C_{Eq}(L,\, n)$ is constant for all sufficiently large n. Then

1. $\{L > 0 : C_{Eq}(L) = 1\}$ *is an open set.*
2. *If $L'_1 = 5.2$, $L'_2 = L'_3 = 6$, there exists some number T such that, if $L \in [0,\, L'_1] \cup [L'_2,\, L'_3]$ then, for all sufficiently large n, $U^T f_{n,L}$ is a consensus.*

   Hence, there exist numbers $5.2 < L_1 < L_2 < 6 < L_3$ such that if $L \in [0,\, L_1) \cup (L_2,\, L_3)$ then $C_{Eq}(L) = 1$.

By the equivalences (3.2) this implies that, for the same values L as in Theorem 3.2.2, consensus must follow for the linear profile f_L and asymptotically almost surely for uniformly random profiles f_L^n.

The proof can be reduced to a series of steps.

First, I choose an L and compute $U^T f_{n,L}$ for some T and some large but finite n.

Second, I consider an arbitrary *refinement* of $f_{n,L}$, that is a profile obtained by starting with $f_{n,L}$ and then inserting some number k of additional agents between each pair of consecutive agents in $f_{n,L}$. For each agent i, I bound how much $U^T f_{n,L}(i)$ can deviate from the $(i + (i-1)k)$'th opinion of any refinement. In other words: If we refine a profile, update it T times and then take away the inserted agents, how much could the result differ from the T'th update of the original profile? There are n left and right bounds, one pair for each i, and they are all independent of k. As the whole process is carried out for thousands of L's and n gets as large as 300 000 for some of them, this requires a computer program.

Third, I show that $U^T f_{n,L}$ has some kind of structure, different ones for the intervals $[0,\, L'_1]$ and $[L'_2,\, L'_3]$, which guarantees $U^\infty f_{n,L}$ is a consensus.

Fourth, I show that any profile within the bounds found in the second step of $U^T f_{n,L}$ must have the same structure, and thus reach a consensus as well.

Letting the resolution k of the refinements go to infinity, I can then use the equivalences (3.2) to prove the theorem.

3.3 Paper III

While Papers I and II study the asymptotic behaviour of large profiles with many agents, Paper III considers a more basic question: If opinions sit in some other set than $\mathbb{R}$, does a limiting profile $\lim_{t\to\infty} U^t f$ even exist for any given profile f? What about the case where opinions are points on a circle?

In the discrete case with n agents with real-valued opinions, as mentioned in Subsection 2.1, convergence was settled already in the original paper [6] by Hegselmann and Krause: Any such profile must reach a limiting profile in finite time.

In other settings the question can be more subtle. For instance, though we sort of sweep the question under the rug in Paper II, as was noted in Subsection 2.2 we don't actually know whether the continuous-agent model always converges.

Let p denote the perimeter of a circle. In this paper we prove the following theorem.

Theorem 3.3.1. *If $p > 2$, the limit $U^\infty f$ exists for any discrete profile f with opinions on the circle with perimeter p.*

As an important part of proving this theorem we consider the *energy* $E(f)$ of a profile f, defined as

$$E(f) = \sum_{i=1}^{n} \sum_{j=1}^{n} \min \left\{1,\, \delta(f(i),\, f(j))^2\right\}, \tag{3.3}$$

where δ gives the shortest distance between the opinions of agents i and j, taken along the circle.

The point is that the energy is always positive and $E(Uf) < E(f)$ for any profile f that isn't frozen. The sequence of energies must thus converge to some limit. One can then show that every time the neighbour graph changes the resulting decrease in energy can be bounded away from 0, and one thus deduces that the graph can only change finitely many times. Once the final graph is reached, the updates can be described in terms of multiplication with a single matrix, and convergence follows by standard linear algebra.

The notion of energy, and the technique of finding a lower bound on the rate at which it decreases, is well known and has been employed by various other authors in Euclidean space. Our use of energy closely follows [1], and the earliest example we could find of a similar technique is in [13]. One notable feature distinguishing our proof from those in earlier works is that all previous proofs somewhere seem to use the fact that, in Euclidean space, one can always find an agent that is most extreme in any given direction. This doesn't hold on the circle, which required us to modify the approach.

Finally, it is also worth noting that the convergence for this altered model is not as well behaved as that of the classical model: Not only doesn't every profile have a finite freezing time, even the time needed for the neighbour graph to settle down cannot be bounded in terms of the number of agents used. The paper ends with some discussion of this and related facts.

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 **54** (2009), No. 11, 2586–2597.

[4] V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis, *Continuous-time average-preserving opinion dynamics with opinion-dependent communications*, SIAM J. Control Optim. **48** (2010), No. 8, 5214–5240.

[5] P. Hegarty and E. Wedin, *The Hegselmann-Krause dynamics for equally spaced agents*, J. Difference Equ. Appl. **22** (2016), No. 11, 1621–1645.

[6] R. Hegselmann and U. Krause, *Opinion dynamics and bounded confidence: models, analysis and simulations*, Journal of Artificial Societies and Social Simulation **5** (2002), No. 3.

[7] R. Hegselmann, S. König, S. Kurz, C. Niemann and J. Rambau, *Optimal Opinion Control: The Campaign Problem*, Journal of Artificial Societies and Social Simulation **18** (2015), No. 3, 1–18.

[8] J. Hendrickx, *Order preservation in a generalized version of Krause's opinion dynamics model*, Physica A: Statistical Mechanics and its Applications **387** (2008), No. 11, 5255–5262

[9] J. Lorenz, *Consensus strikes back in the Hegselmann-Krause model of continuous opinion dynamics under bounded confidence*, Journal of Artificial Societies and Social Simulation **9** (2009), No. 1, 105–120.

[10] 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** (2007), No. 12, 2214–2226.

[11] A. Mirtabatabaei and F. Bullo, *Opinion Dynamics in Heterogeneous Networks: Convergence Conjectures and Theorems*, SIAM J. Control Optim. **50** (2012), No. 5, 2763–2785.

[12] S. Mohajer and B. Touri, *On convergence rate of scalar Hegselmann-Krause dynamics*, American Control Conference 2013, Washington DC, pp. 206–210.

[13] M. Roozbehani, A. Megretski and E. Frazzoli, *Lyapunov analysis of quadratically symmetric neighborhood consensus algorithms*, Proceedings of the 47th IEEE Conference on Decision and Control (CDC 2008), pp. 2252–2257.

[14] E. Wedin and P. Hegarty, *A quadratic lower bound for the convergence rate in the one-dimensional Hegselmann-Krause bounded confidence dynamics*, Discrete Comput. Geom. **53** (2015), No. 2, 478–486.

[15] E. Wedin and P. Hegarty, *The Hegselmann-Krause dynamics for the continuous-agent model and a regular opinion function do not always lead to consensus*, IEEE Trans. Automat. Control **60** (2015), No. 9, 2416–2421.