CLUSTERS, WAVES, AND FORCE CHAINS IN FIRE-ANT COLLECTIVES:
EMERGENT BEHAVIOR IN OUT-OF-EQUILIBRIUM PARTICULATE SYSTEMS

A Dissertation
Presented to
The Academic Faculty

By

Caleb J. Anderson

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Physics

Georgia Institute of Technology

July 2021

Copyright © Caleb J. Anderson 2021
CLUSTERS, WAVES, AND FORCE CHAINS IN FIRE-ANT COLLECTIVES: EMERGENT BEHAVIOR IN OUT-OF-EQUILIBRIUM PARTICULATE SYSTEMS

Approved by:

Dr. Alberto Fernández-Nieves
Advisor
School of Physics
Georgia Institute of Technology

Dr. Guillermo Goldsztein
School of Mathematics
Georgia Institute of Technology

Dr. Roman Grigoriev
School of Physics
Georgia Institute of Technology

Dr. Michael Schatz
School of Physics
Georgia Institute of Technology

Dr. Peter Yunker
School of Physics
Georgia Institute of Technology

Date Approved: July 13, 2021
Go to the ant, thou sluggard; consider her ways, and be wise:
Which having no guide, overseer, or ruler,
Provideth her meat in the summer, and gathereth her food in the harvest.

Proverbs 6:6-8
To all the friends I met at Georgia Tech.

You made my time here pass too swiftly!
ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my advisor, Alberto Fernandez-Nieves, for your contagious enthusiasm, love of science, guidance, and direction. When my interest sometimes waned, a discussion with you often got me excited about a project all over again.

I’m grateful for my lab mates, especially those who were senior to me, Perry Ellis, John Hyatt, Alex Fragkopoulos, Winnie Chang, and Mike Tennenbaum. You were my role models in the group and taught me what it was to be a researcher.

I’d want to thank the research assistants that worked with me, Julio Urquidi and Pryor Gibson, for your help on your special projects and for the excitement you brought into the lab.

To the Gulliver lab at ESPCI, Olivier Dauchot, Guillaume Durey, and Guillaume Briand, thank you for quickly making me feel at home and part of your group. It was a treat to see how another group was organized and you left a great impression on me.

To my wonderful wife, Maggie, thank you for your constant support throughout my graduate work and for helping to edit and polish this thesis.

Finally, I am grateful to my sources of funding: the NSF FLAMEL program, the Chateaubriand program, and the School of Physics. Without them, I would not have been able to devote my time to research.
TABLE OF CONTENTS

Acknowledgments ........................................................................ v

List of Figures ............................................................................ x

Summary .................................................................................... xxvi

Chapter 1: Literature Review ....................................................... 1

1.1 Collective-Motion ................................................................. 1

1.1.1 The Vicsek Model ............................................................. 1

1.1.2 The Toner-Tu Equations .................................................... 3

1.1.3 Physical Examples ........................................................... 7

1.2 Fluctuations in Active Matter ............................................... 9

1.2.1 Equilibrium Number Fluctuations .................................... 9

1.2.2 Giant Number Fluctuations .............................................. 11

1.2.3 Sound Waves in Active Matter ........................................ 15

1.3 The Nature of the Order-Disorder Phase Transition ............... 17

1.3.1 Debate ........................................................................ 17

1.3.2 Polar Bands ................................................................. 22

1.4 The Search for State Variables ............................................. 25

1.4.1 An Effective Temperature .............................................. 25


1.4.2 Effective Temperature in Active Matter ........................................... 26
1.4.3 Explicit Violations of Detailed Balance ............................................. 32

1.5 Motility-Induced Phase Separation .................................................... 33
  1.5.1 Motility-Induced Potentials ......................................................... 33
  1.5.2 Phase Separation .................................................................... 36

1.6 Previous Physics Work on Ants .......................................................... 38

1.7 Conclusion ......................................................................................... 42

**Chapter 2: Ant Behavior and Interactions** ............................................. 44

2.1 Harvesting and Storing Ants ............................................................... 44

2.2 Ants at Low Density ........................................................................ 48
  2.2.1 Tracking .................................................................................. 49
  2.2.2 Speed Distributions ................................................................. 53
  2.2.3 Motility Induced Attraction ....................................................... 54
  2.2.4 Potential of Mean Force ............................................................ 56

2.3 Clustering .......................................................................................... 60
  2.3.1 Structure ................................................................................. 62
  2.3.2 Cluster Size Distributions ......................................................... 64

2.4 Ants at High Density ......................................................................... 66
  2.4.1 Density Autocorrelations ........................................................... 66
  2.4.2 Activity Cycles ....................................................................... 70

2.5 Conclusions ....................................................................................... 74

**Chapter 3: Ant Waves** ........................................................................ 76
Chapter 4: Active Janssen Effect

4.1 Review of Janssen’s Experiment

4.1.1 Fluid Columns

4.1.2 Granular Systems

4.2 Scaling Down

4.2.1 Set Up

4.2.2 Overshoot

4.2.3 Time Dependence
Vita ................................................................. 183
LIST OF FIGURES

1.1 a) One of Vicsek et al.’s simulations only a few steps after initialization. The particles have random orientations. b) For simulations with low density and low noise, the particles form separate ordered clusters that move together. c) For simulations with high density and high noise, the particles move in random directions with strong correlations. d) Finally, for simulations with high density and low noise, the rotational symmetry of the system is broken and the system enters a collective-motion phase. This figure is reproduced from Reference [17].

1.2 a) Desert locusts demonstrating collective-motion in two dimensions. This contrasts the result of the Mermin-Wagner theorem for equilibrium matter. Photo reproduced from Reference [35]. b) Sixfinger threadfins demonstrating collective-motion. Ward et al. [32] showed that fish use quorum sensing to weigh the motion of all of their visible neighbors. Photograph from the National Oceanic and Atmospheric Administration’s photo library.

1.3 a) An illustration of a Nambu-Goldstone mode. Because all orientations are allowed ground states and all interactions are local, arbitrary rotations become energetically cheaper as $\Delta \vec{r}$ is increased. b) An illustration of a density fluctuation caused by a Nambu-Goldstone mode in a self-propelled system. Consider several equal spaced regions given by the bottom circles, each with a net velocity given by the black arrows. The net velocities have a soft Nambu-Goldstone mode allowing for rotations. After some time the particles from the original regions find themselves in the regions shown by the top circles. Notice that the density fluctuation caused by this mode is primarily in the axis perpendicular to the direction of the collective-motion.
1.4  a) The positions and orientations of active disks during an experiment using polar vibrated disks. The boundary is flower shaped to avoid particles accumulating in corners, and data is only taken from the center of the flower. The color of the disks represents the strength of the local order, with darker red disks more highly ordered. The inset shows a diagram of a single disk. The disk has two asymmetric feet, causing it to travel in the direction of its smaller foot when it is shaken vertically. b) As the acceleration (scaled by gravity), $a_0$, of the shaking increases, the order parameter, $\Psi$, begins to disappear. The increase in acceleration corresponds to increasing the amplitude of the shaking which increases the rotational noise of the particles and decreases the collective-motion. c) The measured number fluctuations in measured regions in an experiment with $a_0 = 2.8$. The results show extremely strong number fluctuations. This figure is adapted from Reference [19].

1.5  (a-c) The power spectrum measured in a system of colloidal rollers undergoing collective-motion for three different propagation directions with respect to the flow direction, $\theta = \pi/4, \pi/2, \text{and} \pi/8$, respectively. The speed of sound is given by the slope of the dispersion relations at $q = 0, \omega = 0$. d-f) The angular dependence of the speed of sound for area fractions $\phi_0 = 0.11, \phi_0 = 0.18, \text{and} \phi_0 = 0.24$ respectively. The points represent their experimental measurements, while the solid lines show the fit to the theoretical expectation, Equation 1.5. This figure is adapted from Reference [47].

1.6  a) Vicsek et al.’s original 1995 results comparing the order parameter, $\Psi$ to the noise as the noise approaches the critical noise level, with density held constant. The power law scaling is indicative of a continuous phase transition. b) Vicsek’s measurements of the order parameter as the density is varied and the noise is held constant near the phase transition. This too shows power law behavior and indicates a continuous phase transition. This figure is adapted from Reference [17].

1.7  a) Grégoire and Chaté’s simulation results. This diagram shows the probability distribution function (PDF) of the global order parameter over time, $\phi_t$, for noise levels close to the phase transition. The two peaks show the phase transition to be discontinuous with the order parameter switching back and forth from disordered to ordered over time, with the ordered state becoming more and more common as the phase transition is approached. b) Their simulations showed that order parameter became discontinuous with noise only for large enough system sizes than those used by Vicsek et al. c) The order parameter as a function of time near the phase transition exhibits the switching suggested by the PDFs. This figure is adapted from Reference [24].
1.8  a) From Vicsek et al.’s 2007 paper: probability distribution functions (PDFs) of instantaneous measurements of the global order for varying noise levels near the phase transition. The single peak in the PDFs that shifts with noise level is indicative of a continuous phase transition. b) The traveling waves that Vicsek et al. considered a computational artifact. The waves always traveled along one of the primary axes of the simulated cell. In the figure, the cell is repeated to show its periodic boundary conditions. This figure is adapted from Reference [50].

1.9  a) An example of a polar band in simulation. b) Polar bands in Chaté et al.’s simulations. The color represents the local density, averaged along the y-direction. c) The measured density (left axis) and polar order parameter (right) averaged across the y-direction of a simulation with polar bands. d) The difference in orientation between the local and global order as a function of local density during a simulation with low noise. Points represent measurements and the black line shows average difference as a function of density. There are no traveling bands in this simulation. e) The difference in orientation between the local and global order as a function of local density during a simulation with moderate noise. The dependence of average order on density leads to phase separation. These figures have been modified from Reference [24].

1.10 a) An example of hysteresis around the phase transition. The blue curve shows the average global velocity of the system as its density is slowly increased, and the green curve shows the average global velocity as the density is slowly decreased. Notice that the density at which the first band forms and the density at which the last band dissolves are not the same. Figure reproduced from Reference [54]. b) A generic phase diagram for the Vicsek model when varying density and noise. The coexistence region lies between the binodals ($B_{gas}$ and $B_{liq}$). Between the spinodals ($S_{gas}$ and $S_{liq}$), any fluctuation from a homogeneous state results in phase separation. Figure reproduced from Reference [51].

1.11 a) A Scanning Electron Microscope image of the Janus particles used for Palcci et al.’s sedimentation experiments demonstrating that only one hemisphere of each particle is coated with platinum. b) The mean squared displacement of the particles with (red) and without (blue) fuel. At long times, the particles diffuse. c) The density as a function of height in sedimentation experiments with different fuel levels. The sedimentation length of the particles increases with increasing fuel. d) The gravitational length from the sedimentation length is directly proportional to the effective diffusion constant for varying fuel levels, suggesting the usefulness of an effective temperature. This figure is adapted from Reference [61].
1.12 a) Ginot et al.’s setup for 2D sedimentation. The Janus particles they use are much more dense than water, so they settle to the bottom of the cell. Tilting the cell on a slight incline allows them to measure sedimentation under a very small gravitational potential. The scalebar is 20 µm. b) The adhesive parameter’s dependence on the measured ideal gas limit effective temperature. The adhesive parameter measures how strongly colloidal particles stick to each other and can be used to define the first two virial coefficients in an equation of state for passive colloids. Ginot et al. report that this seems to hold true for their active colloids as well. This figure is adapted from Reference [66], Creative Commons.

1.13 a) A simulation explicitly demonstrating Solon et al.’s results on pressure in active systems. The self-propulsion speed of any particles that find themselves on the right are higher than those found on the left, and there is no physical barrier separating the two sides. b) By varying the levels of self-propulsion in the two halves of the cell, they demonstrate that the system fails to equalize pressure. $v_1$ and $P_1 = \rho v_1^2$ are the propulsion speed and measured pressure in the left half of the box, while $v_1$ and $P_2$ are the corresponding values in the right half of the box. The measured pressure inside each box follows the ideal gas equation of state, where the effective temperature has been measured from the diffusion constant of the particles. c) Instead of equalizing pressure between the two halves, the system evolves to equalize $\rho v$. This Figure is adapted from Reference [54].

1.14 MIPS in Linden et al.’s Janus particles. The images are photos of the particles after a) 0.02s, b) 0.4s, c) 2s, d) 5s, e) 20s, f) 68s. The scale bar is 100µm. This figure is reproduced from Reference [75].

1.15 a) A creep experiment on an interconnected aggregation of fire ants showing the strain on the top plate $\gamma$ as a function of time. When the ants are active, they rotate the plate. b) The normal force measured during the same experiment. When the ants are active they push out against the plate. c) A frequency sweep, in which the top plate is oscillated through a range of frequencies and the ants’ elastic, $G'$, and viscous, $G''$, moduli are detected. In this sweep the ants are active and this activity makes $G'$ (dark squares) and $G''$ (light squares) roughly equal over the range of measured frequencies. d) Three frequency sweeps of the ants while they are in their inactive state. In this case the elastic modulus (filled symbols) is much higher than the viscous moduli (empty symbols) for all measured frequency, indicating that the ants behave as an elastic solid. This figure is modified from Reference [13].

1.16 The response of model ants to a shear demonstrated by the distributions, $\phi$ of the lengths of the links between the ants, $r$. The shear initially results in a stretching of some of the ants’ connections. This causes a short term elastic response. However, given enough time, the activity of the ants results in them breaking connections and forming new connections to their current nearest neighbors, dissipating the stored elastic energy. This Figure is reproduced from Reference [81].
1.17 a) Space-time surface of experimental observations of ants confined to a tunnel. $\tau$ is the relaxation time and corresponds to the temporal length of a traffic jam. b) One of the results from Gravish, et al.’s simulations of ant traffic with varying interaction times. Shorter social interactions, $T_{int}$, lead to a fragile glass transition, while longer interactions lead to a strong Arrhenius-type glass transition. Experimentally measured interactions are of an order that leads to a fragile transition. This figure is adapted from Reference [82].

2.1 a) A fire ant mound in Kennesaw, Georgia. It has been disturbed with a shovel to examine the aggression of the ants’ response, which we take to signal the health of the colony. b) The same region of soil after the mound and surrounding soil have been collected along with the ants. c) The ants and soil in a five gallon bucket 24 hours after the ants were collected. d) Several buckets that together hold the soil and ants from a single colony and a system of irrigation tubing that slowly drips water. The ants eventually form a raft on the water’s surface. e) The ants from the previous panels after being removed from the buckets and placed into an open Tupperware bin.

2.2 a) The length distribution of the worker ants we collect. b) A single sterile female fire ant. The red line represents the measurement we take to determine the length of the ant. The black circle represents the effective area of the ant, $A_{ant}$, which we use to calculate packing fractions.

2.3 a) An image of the empty experimental set up we use to track ants in 2D. We place the circular cell flat and light it from below through a sheet of light-diffuser. b) An image of $N = 40$ ants in the cell taken from above. c) The same image after background reduction and thresholding. Connected regions of pixels are taken to be a single ant. d) The tracked positions of the ants. The short black bars represent the center of mass positions and orientations of the ants shown in the previous images, while the colored dots show the trajectories of each of the ants over the last 20 frames.

2.4 a) The probability distribution of the ants’ speeds in the $D = 9 \text{ cm}$ cells for varying numbers of ants. b) The probability distribution for the length of time an ant remains stationary before moving again when it is alone in a cell. The black line is a power law with a critical exponent of 2.8. c) The probability distribution for the length of time an ant remains stationary for $N > 1$. The black line is a power law with an exponent of 2.1.

2.5 a) The average speed of an ant as a function of its distance from a second ant. b) The effective motility induced potential calculated from the average speed. c) The average alignment of the ants as a function of distance. Ants are only aligned as very short distances (when they are likely to be stationary) or in the rare circumstance that they are on exactly opposite sides of the circular cell.
2.6 a) An illustration of the variables used to calculate the pair correlation of the ants. 
b) The measured pair correlation of the ants. The single peak becomes less prominent as the density is increased. 
c) The effective potential of mean force for the ants, as calculated from their pair correlations. 
d) A comparison of our two different measures of an effective potential for \( N = 3 \) ants in the cell.  

2.7 a-c) A comparison of the shapes of \( U_{eff,g} \) and \( U_{eff,v} \) for experiments using \( N = 3, 5, \) and 40 ants, respectively. Notice that the attractive part of the potentials have the same shape. 

2.8 a) A frame during an experiment with \( N = 40 \) ants, in which the stationary ants are highlighted in red and marked with their label from the tracking algorithm. 
b) A dendrogram showing which ants are the most closely clustered. Dividing the ants into clusters is done by checking which ants are closer to a cluster than \( r_{thresh} \). The labels in this figure are the same as in the previous one. 
b) The result of the clustering algorithm. Three clusters were found that contained more than \( N_c = 1 \) ant. 
d-e) The same process for an image from an experiment with \( N = 100 \) ants, showing detection of a cluster with \( N_c = 22 \). The labels have been left off for clarity. 

2.9 The average radius of gyration for clusters made of different numbers of stationary ants. Small clusters scale like a line because they are more likely to form on the walls. Large clusters appear to fill space. 

2.10 a) Log-log plot of observed cluster size distributions. For \( N = 100 \) ants, there is apparent power law behavior at small cluster sizes. 
b) Log-linear plot of observed cluster size distributions. For \( N = 100 \) ants, there is an exponential tail for large cluster sizes. 

2.11 False color image of ants with a packing fraction \( \phi_{eff} = 3.6 \) in a \( D = 4.5 \ cm \) cell. Ants shown in black do not move within the next 30 s. 

2.12 a) An image of ants in a \( D = 4.5 \ cm \) cell with a global packing fraction \( \phi_{eff} = 2.5 \). 
b) The coarse grained optical density corresponding to the previous image. Yellower squares correspond to higher optical density. 
c) The density autocorrelation function for \( \phi_{eff} = 2.5 \), illustrating the two inverse timescales. \( \tau_1 \) in blue corresponds to the walking movement of the ants. \( \tau_2 \) in red corresponds to formation and break up of clusters. 
d) The two timescales as a function of the global packing fraction. Both timescales generally increase as the density is increased, indicating that the ants’ dynamics slow down. 

2.13 a) A false color image in which black pixels represent ants that are stationary for \( 2\tau_1 = 16.8 \ s \) and magenta pixels represent ants that moved within \( 2\tau_1 \). 
b) Another false color image created the same way but starting \( 2\tau_2 = 12.3 \ min \) later.
2.14 a-c) Three images, each taken 4.4 min apart \( \sim 119 \) min into the experiment, outside of an activity cycle. The global packing fraction \( \phi = 3.6 \). d-f) Three images, each taken 8.3 s apart \( \sim 150 \) min into the experiment, during an activity cycle. g) The optical area fraction (gray) and moving ratio (red) as function of time with a packing fraction of \( \phi = 3.6 \). The strong peak in the moving ratio corresponds to an activity cycle. The black vertical lines correspond to the image sequences shown above. h) The presence (blue) or lack (red) of activity cycles as a function of the packing fraction in the experiment. Some points are offset vertically for clarity.

3.1 (a) The large cell used to confine the ants. (b) A dense mass of ants confined to the cell shortly after the ants are knocked down to the base of the cell. (c) The same mass of ants fully expanded. (d) A schematic of the experimental set up. A large fluorescent lamp behind the cell serves as a back-light. (e) The calibration curve used to calculate the density of ants in the cell from the transmission. (f) The space-time surface calculated for the same experiment as images b, c, and d of this figure. Colors closer to red indicate higher density.

3.2 (a) A three dimensional view of the space time surface for the wave highlighted in Figure 3.1f. The surface has been flipped and rotated to highlight the constant speed and amplitude of the wave. (b) Three real space images of the same wave taken 20 seconds apart. The scale bar 5 cm and the arrows have been added to draw attention to the wave.

3.3 a) The vertical speed, \( v_{0,y} \) of three ants as a function of time, while a wave passes through their region. The ants activate to move in the same direction as the wave. b) Three images taken 44 s apart during a wave with three ants highlighted by arrows that point from their abdomen to their mandibles. The vertical speed was calculated by tracking the position of the center of the arrows. Scale bar is 2 cm.

3.4 (a) A close up image of ants near the horizontal center of a large cell. The white arrows point from the abdomen to the mandibles of the ant and are used to determine the ants’ positions and orientations. The scale bar is 5 mm. (b) A small region of interest in which the orientations of ants have been measured over the course of a wave. The region of interest has been separated into three bins, and the polar histograms to the sides show the orientational probability distribution functions for each bin. The bottom two bins show the orientational order brought on by the wave. The scale bar is 1 cm. (c) An illustration of images shifted into the wave frame. The red line indicates the center of the density wave, \( y' = 0 \), and the blue box represents the area of one of the vertical 2 cm bins used to average the ants density and orientation. (d) The density (blue) and average alignment (red) in each of the bins as a function of the bin’s position relative to the center to the wave, \( y' = 0 \). \( y' < 0 \) corresponds to regions through which the wave has already passed, and \( y' > 0 \) corresponds to regions with ants that have not yet felt the effects of the wave.
3.5 Three frames from our simulation showing a traveling density wave. 

3.6 An illustration of the terms used in the solution to our analytical model. a) In the reference state, \( y = x \) and the free surface lies at \( M/\rho_c \). b) At \( t = 0 \), all material particles are in the uncompressed state (blue rectangle), where \( y_x = \rho_c/\rho_d \). Therefore, \( y(x_g) = x_g = 0 \). c) At some arbitrary time later, the free surface has drifted downwards a distance \(-gt\) and some material particles near the base of the column are compressed (red rectangle). Notice that \( x_g(t_1) \) now refers to a different material particle than \( x_g(0) \). d) Another arbitrary time, \( t_2 > t_1 \). Even more of the column is compressed.

3.7 The three key ingredients of the continuum model correspond to three density regimes in the over-damped PDE: Uncompressed ants (blue springs) drift down at constant speed due to gravity. Compressed ants (green springs) are subjected to competing gravitational and decompressing forces. For active ants (red springs), gravity and compression effects are ignored and ants travel upwards at a set velocity.

3.8 Analytical solutions to the PDE for four different times, where \( t^* \) is the time that the wave reaches the interface between the compressed and uncompressed regions and we have chose \( \rho_a = 5/4, \rho_d = 4/5, \) and \( \rho_c = 1 \). The curves have been offset vertically relative to the \( t = 0 \) curve for clarity. The first curve is immediately after the density at the bottom of the column reaches the critical density \( \rho_a \). The second shows the wave during the expansion phase in the compressed region. The final two show the wave traveling at a constant speed through the uniform density background.

3.9 a) Space-time diagram of waves in our simulated ant column. b) A representative sample of the speeds and amplitudes of the waves in our simulations. c) The speed of the waves in our analytical model, given that \( v_0 \) is constant.

3.10 a) An image of a wave propagating through the column of ants. b) Another image of the same wave 40 s later. c) The result of blurring and subtracting the two intensities, then converting to the transmission difference into a density difference. The \( \times \) indicates the location with the peak change in density between these two frames, \( \Delta \rho_{\text{max}} \). d) The peak change in density between frames as a function of time for the last 50 seconds before the wave reaches the free surface. The solid line indicates the average amplitude of the wave, and the dotted lines indicate the error. e) A measurement of the speed of the wave. The red line is a linear fit of the position of the peak of the wave. Its slope is used as the speed of the wave, \( s \).

3.11 a) The measured speed of the waves in experiments, \( s \), as a function of their amplitude, \( \langle \Delta \rho_{\text{max}} \rangle \). b) The speed of the waves in simulations, in which \( v_0 \propto \rho_{\text{local}} \). c) Speed of the wave in the analytical model as a function of \( \rho_a \), given that \( v_0 \propto \rho_a \).
4.1  a) Janssen’s original setup. The base of his column is connected to a balance that he loaded with varying mass (marked G), and the column walls are connected to the floor with springs (marked S). b) An illustration of a horizontal slice in a cylinder that is in mechanical equilibrium. Possible forces on the slice in include gravity, compressive stresses, and frictional forces with the walls. c) Navy line: The theoretical prediction for the apparent mass detected by a scale for a fluid in Janssen’s setup. Navy and red points: Two trials from Janssen’s original experiments using a square column with side length \( l = 60 \text{ cm} \). Black line: The theoretical prediction assuming a homogeneous isotropic elastic material at the Coulomb limit with its walls.

4.2  a) A composite image showing a 2D granular system under ambient light and the same system through crosspolarizers, highlighting the force chains. This figure is modified from Reference [101]. b) Force chains transferring a purely vertical surface stress to the walls. This figure is reproduced from Reference [102].

4.3  a) The probability distribution function (PDF) of the single particle mass, \( m_0 \), of the plastic beads we use in our passive Janssen experiments, measured from \( N = 200 \) beads. b) The PDF of the diameter, \( a \), of the plastic beads. c) To measure the coefficient of friction between the beads and the cylinders, we tilt a column containing a tripod (shown in the inset) of particles until it first slips and use the angle of the cylinder with respect to the horizontal to calculate the coefficient of friction. Inset: a top down view of the tripod of particles we use to make these measurements.

4.4  a) An experiment using a thin cylinder and a clamp that holds the cylinder away from the scale. b) A closer look at the base of the column showing that the walls of the cylinder are not in contact with the scale. c) An alternate set up used for large diameter tubes. Wooden beams support the weight of the open cylinder, and a platform inside the cylinder supports the beads and puts them in contact with a scale. d) Three trials in which the mass was measured by the scale, \( m_{\text{app}} \), for \( D/a = 29.0 \) using clamps to hold the cylinder off the scale. The cylinder is not held tightly enough and eventually it slips so that some of its mass is supported by the scale. The \( m_{\text{app}} \) never saturates, regardless of \( m_{\text{add}} \). e) Three trials with \( D/a = 29.0 \) using wooden beams to support the cylinder. In all three trials, \( m_{\text{app}} \) eventually saturates as expected. The black line shows the results of Equation 4.5 with \( M \) set as the average maximum \( m_{\text{app}} \) from the three trials.

4.5  a) An overshoot in which \( m_{\text{app}} > m_{\text{add}} \) is visible in the results for cylinders with \( D/a = 8.2 \). b) The overshoot for \( D/a = 3.7 \) varies from experiment to experiment. c) A set up in which the base of the column is supported by a wooden beam and the cylinder is in contact with the scale. d) \( \Delta m_c \) for the column (cyan) and the corresponding calculated apparent mass of the grains, \( m_{\text{app}} \) (red) for \( D/a = 3.7 \). Experiments directly measuring \( m_{\text{app}} \) are shown as empty points. e) Height of the overshoot as a function of \( D/a \). The line shows \( D^{-1} \).
4.6  a) A projection of the force chains in Mahajan et al.’s simulations into 2D showing the strongest 5% of the force chains present in the simulation. The force chains generally point downwards from the wall. Notice the filling height is low compared to the radius of the cylinder. b) Force chains later in the simulation, after enough particles have been added to fill the cylinder to more than double the height presented. Now the force chains are mostly supportive. c) An example in 2D of a secondary rearrangement that results in compressive frictional forces. d) A fit of Mahajan et al.’s theory against our experimental data for $D/a = 8.3$. This figure is modified from Reference [105].

4.7  The height of the column during the maximum overshoot of $m_{\text{app}}$ in our experiments with spherical plastic beads for varying $D/a$.

4.8  a) Time dependence of $m_{\text{app}}$ in an experiment with $D/a = 29.0$. The red curve is $m_{\text{app}}$, and the blue curve is $m_{\text{add}}$. b) The same time dependence in an experiment with $D/a = 3.7$ during the beginning of the overshoot. c) Time dependence of the mass of the cylinder relative to its initial mass, $\Delta m_c$.

4.9  The $m_{\text{add}}$ (navy) and $m_{\text{app}}$ (red) of glass beads in a column with $D/a = 14.8$. Every time new beads were added, the column was shaken. After each shaking event, $m_{\text{app}}$ returns to a level consistent with an overshoot.

4.10 a) The measured size distributions of our garbanzo beans along their three principle axes. b) The mass distribution of the our garbanzo beans. c) The size distribution of our mung beans along two principle axes. d) The mass distribution of our mung beans.

4.11 a) An example of an overshoot in three trials using mung beans in a column with $D/a = 5.1$ (green points) and an overshoot in three trials using garbanzo beans with $D/a = 5.1$ (tan points). b) The magnitude of the overshoot as a function of $D/a$ for three types of particles: spherical plastic beads (navy), mung beans (green), and garbanzo beans (tan). The points represent the average, and the error bars represent the maximum and minimum values measured in different trials. c) The height at which we recorded the peak magnitude of the overshoot for all three types of particles.

4.12 a) A cylinder modified by wrapping the base in thin plastic. b) Ants confined to the modified cylinder. c) Ants in contact with a scale during an experiment. d) A sample of the measured $m_{\text{add}}$ (blue) and $m_{\text{app}}$ (red) during an experiment with the ants. e) A plot of $m_{\text{app}}$ vs $m_{\text{add}}$ for several trials of ants in a $D = 22.2 mm$ open cylinder. The color of the points for each trial corresponds to how long we had already held the colony in captivity. f) The same plot as in the previous figure, but this time the trials are color coded so that all trials of a given color used ants from the same colony.
4.13 a) The change in $m_{\text{app}}$ during the 60 s immediately following the addition of more particles for an experiment with passive spherical grains and $D/a = 29.0$. The apparent mass decays over time. b) The same measurement for an experiment with passive spherical grains and $D/a = 3.7$. c) The change in $m_{\text{app}}$ for the ants over 60 s after adding more mass, after $m_{\text{app}}$ was averaged over 1.33 s. d) The apparent mass of the ants when 17.5 g of ants are added all at one time at $t = 0$. The apparent mass continues to increase until it approaches the apparent mass for the same mass added in an experiment with incremental chunks.

4.14 a) The fluctuations in $m_{\text{app}}$ as a function of time. Notice that the fluctuations become larger as more ants are added to the column. b) The root-mean-square fluctuations of the apparent mass during each interval between adding chunks of ants across three trials using the same colony. c) Averaging across all three experiments, we measure the average fluctuations for various $m_{\text{app}}$. The black line represents $m_{\text{app}}^{0.5}$.

4.15 Average autocorrelations $\delta m$ for various measurements of $m_{\text{app}}$ (colored points). The autocorrelation times do not seem to depend on $m_{\text{app}}$. The black points are autocorrelation measurements averaged over all added chunks.

5.1 a) Side view of one of the polar grains. When shaken the grain self propels in the direction $\vec{n}$. b) Bottom view of the polar grain showing that the back foot is much wider than the front foot. c) An image of the grains at low density after image processing and tracking. The grains have been highlighted with white circles, red arrows indicate the grains’ orientations, and colored lines indicate the trajectories of the particles. Parts of this figure are adapted from References [19, 46].

5.2 a) The mean squared rotation of the particles as a function of time in experiments with $a_0/g = 2.4$ (solid blue curve) and $a_0/g = 2.6$ (solid red curve). The dashed lines are the best fits of a line to the first ten points in each experiment and are used to calculate $D_\theta$. b) Measurements of $D_\theta$ for varying $a_0/g$.

5.3 The trajectories of a simulated grain after a perturbation has disaligned its orientation (black arrows) from its velocity (colored arrows).

5.4 a) Polar grains confined to our ring geometry at a global area fraction of $\phi_0 = 0.17$. The Cartesian coordinate system we use for some calculations is indicated in red. b) Polar grains at a global area fraction of $\phi_0 = 0.39$. The polar coordinate system we use is in other calculations is indicated in red.
5.5 a) The global order parameter $\Psi$ over a range of densities when the amplitude of the bottom plate’s vibration is $a_0/g = 2.2$. The order parameter decreases smoothly for decreasing $\phi_0$ but never reaches $\Psi = 0$. b) The global order parameter over a range of $D_\theta$ with $\phi_0 = 0.4$. c) The probability distributions of measurements of $\Psi$ during experiments with varying $\phi_0$ d) The PDFs of $\Psi$ during experiments with varying $D_\theta$.

5.6 a) A space-time plot of the area fraction, $\phi(\theta, t)$, showing dense bands traveling in the $-\theta$ direction in a trial with $\phi = .39$ and $a_0/g = 2.2$. b) The space-time plot of $u(\theta, t)$ for the same trial. To the eye, the regions with the lowest $u(\theta, t)$ correspond to more rarefied regions in (a).

5.7 a) Dispersion relation for $\phi_0 = 0.39$. The slope of the peaks through $q = 0, \omega = 0$, demonstrated by the red line, corresponds to the typical propagation speed of the traveling bands. b) The average global drift speed, $u_0$, of each trial in the density sweep (blue circles) and the speed of the propagating fronts, $c_0$, in those same trials (red triangles). c) The drift speed and the wave speed during one of the noise sweeps.

5.8 a) The average polarity of the grains, $\langle p \rangle$, as a function of the local area fraction, $\phi$. The error bars represent one standard deviation from the mean and show the typical spread of the distribution. b) The average squared speed of grains $\langle v^2 \rangle$ as a function of $\phi$. The error bars are one standard deviation from the mean. c) A heat map showing our measured values of $u(\theta, t)$ and $\phi(\theta, t)$ over the course of the density sweep. d) A zoomed in view of a region in the space-time plot in Figure 5.6a, showing a jam and the resulting band.

5.9 The average speed, $\langle u(\phi) \rangle$, as a function of local area fraction (left axis) compared to the probability distribution functions for the local area fraction for 4 different trials with global packing fractions (right axis): $\phi_0 = 0.17$ (blue), $\phi_0 = 0.39$ (green), $\phi_0 = 0.53$ (yellow), $\phi_0 = 0.66$ (red). The spread in density shows how trials with global packing fraction below $\phi_0 = 0.5$ can still exhibit traveling waves with moving at the maximum allowed speed.

5.10 (a) The same heat map as was shown in Figure 5.8 with an arrow indicating that fluctuations that lead to local packing fraction of $\phi(\theta, t) > 0.6$ tend to self reinforcing, pushing the density even higher. (b) The stationary jam from the $\phi = .66$ trial in the density sweep. The direction of the collective motion is counter-clockwise.
5.11 a) The individual balls in our chain marked with blue to yellow from one end of the chain to the other. This example shows the first image in trial with a polymer with \( N = 90 \) monomers in an active bath. The flower shaped cell prevents active particles from accumulating on the walls. b) The last image from the same trial, showing that none of our monomers have switched labels. c) A side view of one of our passive grains, showing only a single isotropic foot. d) A bottom view of a passive grain. The inside of the circular foot is concave so only its edges come into contact with the glass plate. e) A sample of the trajectories of some passive particles showing that they undergo Brownian motion. Parts of this figure are reproduced from Reference [19].

5.12 a) The probability distribution functions for \( R_g/(N\sigma) \) for three different lengths of chains in a passive bath. The color represents the number of monomers in the chain. In these units, higher \( N \) would correspond to lower \( \kappa \) in equilibrium systems, which results in lower radii of gyration. b) The probability distribution functions for \( R_g/(N\sigma) \) for five different lengths of chains c) \( \langle R_g \rangle/\sigma \) for passive (blue) and active (red) baths. The black line represents \( R_0^{0.75} \), corresponding to a self-avoiding random walk.

5.13 a) A chain with \( N = 45 \) in a nearly fully extended configuration. The lines represent the principle axes of the configuration and their length is \( 2\sqrt{\lambda} \). b) The chain in a more tightly confined configuration. c) The chain in a hairpin configuration.

5.14 a-c) The measured prevalence of conformations with various \( R_g \) and \( A^2 \) for polymers with \( N = 32, 45, \) and \( 78 \) monomers, respectively, immersed in a passive bath. d-f) The prevalence of conformations for polymers with \( N = 32, 45, \) and \( 78 \) monomers, respectively, immersed in an active bath.

5.15 a) An example of a chain in a passive bath depleting to the boundaries. Once the chain has depleted, it never returns to the bulk. b) An example of a chain coiling in a passive bath. This is also a steady state as the grains can't exert any force that would uncoil it.
SUMMARY

Active matter consists of large numbers of particles that can each convert stored or ambient energy into work. Some progress has been made studying synthetic active matter within the last two decades, but the most familiar active systems are biological. In fact, all organisms and groups of organisms can be viewed as types of active matter. They burn stored chemical energy to enact motion on many different scales, from churning micron-scale microtubules and motor protein combinations [1, 2], to flocking birds [3, 4] or herding sheep [5]. Despite the difference in scales and particle interactions, we now know that active matter systems display some universal behaviors.

The constant energy input and dissipation from the active particles drives and holds active matter far from equilibrium. This makes active systems distinct from many other common out-of-equilibrium systems, which tend towards equilibrium and only remain in other instances for relatively short amounts of time.

Equilibrium systems are governed by the principle of detailed balance, which states that every microscopic process must be in balance with its own reverse process. For example, consider an equilibrium system consisting of water in a metal cup. At any instant, there is a probability $p$ that a water molecule will collide with the cup’s wall and transfer some of its kinetic energy into a vibration in the cup’s wall. Detailed balance states that a reverse process must exist, in which a vibration in the cup’s wall excites a water molecule, increasing the water molecule’s kinetic energy, and further that, at any instant, the probability of this process occurring is also exactly $p$. Ultimately, detailed balance stems from the time reversal symmetry of microscopic processes [6] and it is a necessary prerequisite for equilibrium [7, 8].

Active systems are out of equilibrium because there is a net flow of energy in violation of detailed balance. For example, one of the birds in a flock may burn chemical energy to propel itself forward by flying, but there is no reverse process in which the bird harnesses flows of air and converts it back into chemical energy. The energy is dissipated from the system through a separate process, usually friction.
Noteworthily, full understanding of active matter may eventually prove invaluable because of its prevalence in our own biology [9], for example in the migration of metastasizing cancer cells [10] or in the process of protein folding in the cytoplasms of cells [11]. Eventually, active materials may also become commercially useful. In the material science paradigm, the properties of a material are completely determined by its components and structure [12]. Active matter could one day offer a new twist on this paradigm by creating materials whose properties and performance depend also on their activity level [13]. As just one example, consider that an active material that could switch between being rigidly solid to flowing, based on its internal energy instead of on ambient temperature.

Active matter has been shown to host a wide range of interesting new physics, including global order-disorder transitions in 2D [14], which are impossible in equilibrium systems due to the Mermin-Wagner Theorem [15], mesoscale pattern formation [1], and giant number fluctuations [16]. However, the details of how the microscopic manifestations of the particles’ activity impacts the macroscopic properties of the collective are still largely unknown. This thesis reviews our efforts to understand the link between microscopics and macroscopics in active matter, focusing on our chosen system of fire ants.

We begin with a literature review in Chapter 2 that briefly introduces much of the new physics that has been discovered in active systems over the last two decades. This includes the previously mentioned phase transition, giant number fluctuations, types of traveling waves, and the search for effective state variables.

Next, we discuss the behavior of ants individually and in small groups in Chapter 3. This chapter includes a section on our methods for capturing and manipulating colonies of fire ants for use in experiments. We then discuss some of our experiments in which we directly track ants in two dimensions to study their dynamics and interactions, especially the details of how they switch from active to inactive and their tendency to form dense clusters.

In Chapter 4, we scale up the two dimensional experiments and report the surprise observation of solitary waves in both activity and density. Our measurements show that these waves travel
with a constant speed and that they behave nonlinearly. We explain the existence of these waves using a mixture of the existing theory for active systems and our results from studying the ants’ interactions. We then verify that these conditions are enough to produce nonlinear waves with simulations and an analytical model.

Chapter 5 explicitly compares the behavior of ants, an active granular system, to traditional granular materials using Janssen’s experiment. We first report some of our surprising results from scaling down the classic Janssen’s experiment to the length scale necessary to compare with ants. We then show that the ants can support force chains like a traditional granular system. This happens in conjunction with having comparable viscous and elastic responses throughout a wide range of timescales. It seems that the ants’ activity does not automatically wash out their granular nature, but it does cause the strength of supportive force chains to fluctuate as they form, break, and reform.

Our studies on another interesting active system, composed of vibrated disks, are contained in Chapter 6. These disks have very different interaction rules compared to the ants, which allows us to again examine the role of the microscopics in the shaping of macroscopic active behavior. We use the disks to examine the nature of the collective motion phase transition and discuss the discovery of density waves in their collective motion phase. Lastly, we show some preliminary data on the behavior of flexible polymers in an active bath that demonstrates their tendency to preferentially fold into hairpin configurations.

Chapters 7 and 8 discuss, respectively, possible follow up work to continue the direction of this thesis and some closing thoughts and conclusions.
CHAPTER 1
LITERATURE REVIEW

1.1 Collective-Motion

1.1.1 The Vicsek Model

In 1995, while studying flocking in birds, T. Vicsek, A. Czirok, E. Ben-Jacob, I. Cohen, and O. Shochet published what has become a keystone paper for active matter [17]. In this paper, they proposed a discrete particle-based model to understand the rise of self-ordered collective-motion from individually active particles. This model consisted of a large number of interacting particles that each move at a constant speed, $v_0$. The only interaction between particles causes them to reorient to noisily align with their nearest neighbors. This simple model can account for a wide variety of real world behavior, including birds that use their vision to align with their nearest neighbors to prevent collisions [18] and vibrated polar grains that only align through physical collisions [19]. Altogether, these rules can be summed up in the following equations:

\[
\begin{align*}
\vec{x}(t + \Delta t) &= \vec{x}(t) + \vec{v}(t)\Delta t \\
\vec{v}(t) &= v_0 \cos \theta(t)\hat{x} + v_0 \sin \theta(t)\hat{y} \\
\theta(t) &= \langle \theta(t) \rangle_r + \Delta \theta
\end{align*}
\]

where $\vec{x}$, $\vec{v}$, and $\theta$ are the particle’s position, velocity, and orientation angle, respectively. In addition, $\langle \theta(t) \rangle_r$ denotes the average orientation of all particles within a specified radius, $r$, from the updating particle, and $\Delta \theta$ is a uniformly distributed random perturbation, $-\eta \leq \Delta \theta \leq \eta$.

For high $\Delta \theta$ and high density, $\rho$, the particles start moving without correlations (see Fig. 1.1a) but then quickly develop strong orientational correlations (see Fig. 1.1c). The Vicsek Model bears a strong resemblance to the famous Ising model, which is classically used to model ferromag-
netism [20], albeit with the spins distributed at random instead of on a lattice. In the Ising Model, symmetry breaking leads to the rise a non-zero net magnetization. It may then come as no surprise that under high enough choices for $\rho$ and low enough choices for $\Delta\theta$, the Vicsek model results in spontaneous symmetry breaking from a disordered phase (see Fig. 1.1b) to a phase with non-zero net transport, (see Fig. 1.1d), which we will hereafter refer to as the “collective-motion” phase. In the literature, this ordered phase is also frequently called the polar-liquid, flocking, or swarming phase.

![Figure 1.1](image-url)

Figure 1.1: a) One of Vicsek et al.’s simulations only a few steps after initialization. The particles have random orientations. b) For simulations with low density and low noise, the particles form separate ordered clusters that move together. c) For simulations with high density and high noise, the particles move in random directions with strong correlations. d) Finally, for simulations with high density and low noise, the rotational symmetry of the system is broken and the system enters a collective-motion phase. This figure is reproduced from Reference [17].

Interestingly, even if the “collective-motion” phase may be expected intuitively, this result is at odds with equilibrium behavior. In equilibrium systems, the Mermin-Wagner theorem imposes a severe constraint on symmetry breaking in $d \leq 2$ dimensions. The theorem states that local
interactions cannot lead to long-range order in $d \leq 2$ dimensional space in an equilibrium system with finite noise [15, 21]. Consistent with this, the ferromagnetic XY model, which is the Ising model generalized to allow the spins to point in any direction, does not lead to symmetry breaking in 2D. Instead, the 2D XY model leads to the the formation of magnetic domains that never span the full system [22]. Clearly, active systems break rules that equilibrium systems must follow, despite sharing similarities, like exhibiting phase transitions.

The symmetry breaking from a phase with zero net momentum to the collective-motion phase with non-zero net momentum also highlights the lack of momentum conservation in the model. Two particles that approach each other and align typically increase the net momentum of the system, which may then relax as the particles separate and experience rotational diffusion. This is possible because the particles self propel against a momentum sink, such as the earth, that is not included in the model.

Despite its simplicity, the Vicsek model and its variations remain relevant today as the basis for most active matter simulations. Some of the popular variations include adding pairwise attractions and repulsions [23], modifying the noise term [24], using only nematic alignment [25], or coupling the particles to Stokes’ equations to simulate swimming particles [26].

1.1.2 The Toner-Tu Equations

Inspired by Vicsek et al.’s simulation work, Toner and Tu formulated the first hydrodynamic (continuum) model of active matter [27]. They considered a coarse grained active fluid and analyzed the degrees of freedom to construct the equivalent of the Navier-Stokes’ equations for such a system. The resulting equations of motion are:

$$\partial_t \vec{v} + \lambda_1 (\vec{v} \cdot \vec{\nabla}) \vec{v} + \lambda_2 (\vec{\nabla} \cdot \vec{v}) \vec{v} + \lambda_3 (\vec{\nabla} |\vec{v}|^2) =$$

$$(\alpha - \beta |\vec{v}|^2) \vec{v} + D_B \vec{\nabla} (\vec{\nabla} \cdot \vec{v}) + D_t \nabla^2 \vec{v} + D_2 (\vec{v} \cdot \vec{\nabla})^2 \vec{v} - \vec{\nabla} P + \vec{f} \quad \text{(1.1)}$$
\[ \partial_t \rho - \nabla \cdot (\bar{v} \rho) = 0 \] (1.2)

Equation 1.2 is the familiar continuity equation for mass conservation. Equation 1.1 is the active analogue to the Navier-Stokes momentum equation, which governs classical, momentum conserving, simple fluids, and that can be written, assuming incompressibility, as:

\[ \rho \left( \partial_t \bar{v} + (\bar{v} \cdot \nabla) \bar{v} \right) = \eta \nabla^2 \bar{v} - \nabla P, \] (1.3)

where \( \eta \) is the viscosity and \( P \) is the pressure. The left side of Equation 1.3 is the density multiplied by the total derivative of the velocity, \( \frac{d\bar{v}}{dt} \), representing the acceleration of a fluid parcel. The right side terms are all related to contact forces acting on the fluid parcel. Altogether then, the Navier-Stokes equation is Newton’s second law applied to every fluid parcel.

We will see that each of the new terms in equation 1.1, compared to 1.3, arise because momentum is not conserved in an active fluid. Enforcing momentum conservation requires an equation to be invariant under a Galilean transformation; that is:

\[ \bar{v}' = \bar{v}(\bar{r})' - \bar{u}_0 \]
\[ \bar{r}' = \bar{r} - \bar{u}_0 t \]

where \( \bar{u}_0 \) is an arbitrary constant velocity field. The effect of this requirement is that there is no “preferred reference frame” for a Galilean-invariant equation; all inertial frames are equivalent and the system obeys physics equations no matter the relative velocity of the reference frame.

In contrast, the Toner-Tu equations only hold for the particular reference frame in which the momentum sink is stationary because the active fluid parcels behavior depends not only on interactions with other fluid parcels, but also on their velocity, \( \bar{v} \), with respect to the momentum sink. For a flock of birds, the momentum sink is the air, so the preferred reference frame of the Toner-Tu equations is the one in which the wind speed is 0. For ants swarming on a table, the momentum
sink is the table, so the preferred reference frame is the one in which the table is stationary.

The vector $\mathbf{v}$ is the coarse grained velocity field, which represents the average velocity of all particles in a small region. If the system is completely disordered, the coarse grained velocity vanishes, $\mathbf{v} = 0$. On the left side of Equation 1.1, the $\lambda_1(\mathbf{v} \cdot \nabla)\mathbf{v}$ term is the analogue of the convection acceleration term in the Navier-Stokes equation and the other two $\lambda$ terms are completely new “convection terms.” Consider briefly the left hand side of equation 1.1 under a Galilean transformation:

$$L.H.S. = \partial_t \mathbf{v}'(\mathbf{r}') - (\mathbf{u}_0 \cdot \nabla)\mathbf{v}(\mathbf{r}') + \lambda_1 \left( \mathbf{v}'(\mathbf{r}') + \mathbf{u}_0 \right) \cdot \nabla \mathbf{v}'(\mathbf{r}')$$
$$+ \lambda_2 \left( \nabla \cdot \mathbf{v}'(\mathbf{r}') \right) \left( \mathbf{v}'(\mathbf{r}') + \mathbf{u}_0 \right) + \lambda_3 |\nabla| \mathbf{v}'(\mathbf{r}') + |\mathbf{u}_0|^2 \quad (1.4)$$

This side cannot be reduced to the same form it had before the Galilean transformation unless $\lambda_1 = 1$ and $\lambda_2 = \lambda_3 = 0$, so that the left side of equation 1.1 would be the total derivative, $\frac{d\mathbf{v}}{dt}$, as in the Navier-Stokes equation. In active fluids, alignment interactions typically lead to $\lambda_1 < 1$, which has the effect of a fluid region accelerating faster than in the momentum conserving case as it approaches regions with a higher flow speed because the underlying particles in the fluid region begin to align with those in the faster flowing region. The other two $\lambda$ terms have a similar effect but are usually less important; they are thus often neglected when applying the Toner-Tu equations.

The first term on the right hand side of Equation 1.1 allows an ordered phase to have a non-zero velocity, resulting in collective-motion. The parameter $\beta > 0$, by definition. When the parameter $\alpha > 0$, the $(\alpha - \beta |\mathbf{v}|^2)\mathbf{v}$ term is the acceleration due a sombrero-shaped potential,

$$U(\mathbf{v}) = -\frac{1}{4} \beta |\mathbf{v}|^4 - \frac{1}{2} \alpha |\mathbf{v}|^2,$$

which has an unstable equilibrium point at $\mathbf{v} = 0$ and infinitely many stable equilibrium points at $|\mathbf{v}| = \sqrt{\alpha / \beta}$. This makes $|\mathbf{v}| = 0$ an unstable solution to the Toner-Tu equations when $\alpha > 0$, resulting in the fact that any fluctuation drives the local velocity field away from $|\mathbf{v}| = 0$ and
towards \( |\vec{v}| = \sqrt{\alpha/\beta} \). In general, \( \alpha \) and \( \beta \) are allowed to be density dependent, so that increasing the density can increase \( \alpha \), eventually causing \( \alpha > 0 \); at this point, spontaneous symmetry breaking occurs to start collective-motion. The presence of this term which depends directly on \( \vec{v} \) without any derivatives cannot remain unchanged under a Galilean transformation, so this term also violates momentum conservation and could not exist in any momentum conserving fluid.

Intuitively, this term makes it clear why the preferred reference frame for the Toner-Tu equations is the one in which the momentum sink is stationary. Consider two flocks of birds: one has \( |\vec{v}| = 0 \) in a frame with no wind and the other has \( |\vec{v}| = 0 \) in a frame with wind velocity \( \vec{u}_0 \). The flock in the frame with no wind must be disordered because the individual birds are continually self-propelling. On the other hand, the flock in the frame with wind has broken symmetry; the birds must be aligned so that the birds are flying against the wind. These two flocks have different symmetries and are therefore not equivalent.

The \( D_B, D_t \) and \( D_2 \) terms are all viscous terms that allow for fluctuations to propagate due to coupling between nearby fluid parcels. The \( D_t \) term is an analogue to the familiar viscous diffusion term that causes fluctuations in the velocity field to spread out isotropically. The other two viscous terms allow for fluctuations in the velocity field to spread out anisotropically depending on whether the fluctuations are in the component of the velocity parallel to the collective-motion direction. In particular, \( D_B \) sets the stress that causes fluctuations in this component of the velocity to spread out in all directions and \( D_2 \) sets another stress that allows fluctuations in the collective-motion direction to spread even faster in that same preferred direction. Once again, the \( D_B \) and \( D_2 \) terms would not be allowed in an equation governing a momentum conserving fluid because they cannot remain unchanged under Galilean invariance.

Finally, \( \vec{f} \) represents a Gaussian white noise and \( -\vec{\nabla} P \) is the internal stress from pressure, which is function of \( \rho \) only.

Equation 1.1 was derived by considering a generic fluid and including every possible combination of vectors and differential operators that yields a vector and truncating at derivatives of the second order.
The Toner-Tu equations are equally valid in the ordered, collective-motion phase and in the disordered phase, which has no net transport. Knowing the coefficients for a system allows one to predict the system’s behavior. However, as in the Navier-Stokes equation, the details of the inter-particle interactions are hidden in the coefficients.

The Toner-Tu model is considered extremely successful for describing the phase transition reported by Vicsek et al. in two dimensions and showing that the lack of Galilean invariance can lead to long range order in 2D. Careful analysis using dynamical renormalization group theory in References [14, 28] shows that, even in 2D, fluctuations in the velocity fields are washed out as the system size is scaled up. In contrast to the 2D XY model, initially distant coarse grained regions can eventually interact if their velocity is different. When a fluctuation in velocity arises in the collective-motion phase, the fluctuating region is then brought into closer contact with its neighboring regions, resulting in an increased density that leads to an increased alignment that suppresses the fluctuation.

1.1.3 Physical Examples

Collective-motion is ubiquitous across the animal kingdom, driven by a variety of effective aligning potentials. For example, it has been shown that desert locusts marching at high density will break symmetry and march from fear of cannibalism [29–31]. They are more vulnerable to predation from the side, and therefore naturally align along an axis, as seen in Figure 1.2a. Fish have adapted to use quorum sensing to choose their orientation in order to avoid following a neighbor with insufficient information [32], as illustrated in Figure 1.2b. Meanwhile, pigeons flock by following a social hierarchy so that there is some delay between the movement of the lead pigeons and the bulk movement of the flock [4].

Even bacteria have evolved to take advantage of collective-motion despite their simple interactions. Wu et al. have demonstrated that Myxobacteria colonies typically grow as the Myxobacteria spread out to search for new food sources. However, as the Myxobacteria move away from the colony, they periodically reverse course to ensure collisions with their neighbors. Because the bac-
teria are elliptical, hard body repulsions can be enough to create an aligning potential. It seems the purpose of this energetically expensive motion is to increase collisions and order the colony to allow for collective-motion [33].

Similarly to the bacteria, the first synthetic active matter system, vibrated polar grains, took advantage of hard body repulsions between particles with anisotropic shapes to create aligning potentials [34].

Figure 1.2: a) Desert locusts demonstrating collective-motion in two dimensions. This contrasts the result of the Mermin-Wagner theorem for equilibrium matter. Photo reproduced from Reference [35]. b) Sixfinger threadfins demonstrating collective-motion. Ward et al. [32] showed that fish use quorum sensing to weigh the motion of all of their visible neighbors. Photograph from the National Oceanic and Atmospheric Administration’s photo library.

However, the effective aligning potentials between animals or between colliding grains can be complicated, so it is useful to examine synthetic systems where the aligning potentials are better understood. Colloidal rollers are one such synthetic system that shows a transition to collective-motion with increasing density [36, 37]. This system is formed by submerging insulating colloids
in a conducting fluid and applying a uniform electric field. This induces an effect called Quincke rotation, in which the colloids break symmetry and begin to rotate at a constant rate around an arbitrary axis orthogonal to the electric field [38]. If the colloids are denser than the fluid surrounding them, they sink and the Quincke rotation causes the particles to self-propel by rolling along the floor of their container. Once moving on the base of the container, the colloidal rollers align and slightly repel each other through electric and hydrodynamic dipole interactions [39]. The collective-motion demonstrated by these rollers shows conclusively that the phase transition in 2D can arise from local alignment interactions, instead of relying on subtle long-range biological communication between agents.

1.2 Fluctuations in Active Matter

1.2.1 Equilibrium Number Fluctuations

In equilibrium matter, the grand canonical ensemble is used to calculate the probability that a system will be in a given state while it can trade both particles and energy with a reservoir [6]. The probability of having a particular number of particles, \( N^* \), and energy, \( E^* \), is

\[
P(N^*, E^*) = \frac{\exp[\beta \mu N^* - \beta E^*]}{\Xi}
\]

where \( \beta = 1/kT \), \( \mu \) is the chemical potential, and

\[
\Xi = \sum_{r,k} \exp[\beta \mu N_r - \beta E_k]
\]

is the partition function. The summation in the partition function runs over all possible numbers of particles and all possible quanta of energy that could be found in the system.

The expectation value for the number of particles in the system is

\[
\langle N \rangle = \frac{\sum_{r,k} N \exp[\beta \mu N_r - \beta E_k]}{\Xi}
\]
Taking the derivative of this expectation value with respect to the chemical potential results in a calculation of the mean squared fluctuations

\[ kT \left( \frac{\partial \langle N \rangle}{\partial \mu} \right)_{T,V} = \langle N^2 \rangle - \langle N \rangle^2 = \langle (\Delta N)^2 \rangle \]

Using the specific volume, \( v = V/\langle N \rangle \), and the Gibbs-Duhem equation, \( d\mu = v \, dP - s \, dT \), the relative mean squared fluctuations are then

\[ \frac{\langle (\Delta N)^2 \rangle}{\langle N \rangle^2} = -\frac{kT}{v} \left( \frac{\partial v}{\partial P} \right)_{T,V} = -\frac{kT}{V} \left( \frac{1}{v} \right) \left( \frac{\partial v}{\partial P} \right)_{T} \]

Note that \( \frac{1}{v} \left( \frac{\partial v}{\partial P} \right)_{T} \) is a material constant of the system, the isothermal compressibility, \( \kappa_T \). So,

\[ \langle (\Delta N)^2 \rangle = -\frac{kT}{V} \kappa_T \langle N \rangle^2 \]

Since \( T \) and \( \kappa_T \) are intensive and \( V \propto \langle N \rangle \), the mean square fluctuations scale with \( \langle N \rangle \) as \( \langle (\Delta N)^2 \rangle \propto \langle N \rangle \) so that the scaling for the relative root mean fluctuations becomes \( \sqrt{\langle (\Delta N)^2 \rangle}/\langle N \rangle \propto \langle N \rangle^{-0.5} \). The result is that they become negligible in the thermodynamic limit. This is a general feature of equilibrium systems, with the exception of equilibrium systems near a phase transition, since in this situation the compressibility diverges, causing the fluctuations to also diverge.

Fundamentally, the \( \sqrt{\langle (\Delta N)^2 \rangle}/\langle N \rangle \propto \langle N \rangle^{-0.5} \) scaling of equilibrium fluctuations is in agreement with the Central Limit Theorem, which governs all systems of independent, identically distributed (i.i.d.), random variables [40]. Consider a sample of \( n \) i.i.d. random variables, \( X_1, X_2, ..., X_n \), each drawn from a distribution of possible values, \( x_1, x_2, ... x_k \), that occur with probabilities \( p_1, p_2, ... p_k \).

If all values of \( x_i \) are finite, the distribution has an expected value, \( E[X] \equiv \sum_{i=1}^{k} x_i p_i = \mu \), and the distribution has a finite variance, \( E[(X - \mu)^2] \equiv \sum_{i=1}^{k} (x_i - \mu)^2 p_i = \sigma^2 \), the Strong Law of Large Numbers dictates that the sample average, \( \bar{X}_n \) converges almost surely to \( \mu \) when \( n \) is very large:

\[ \bar{X}_n \equiv \frac{1}{n} \sum_{i=1}^{n} (X_1 + X_2 + ... + X_n) \xrightarrow{a.s.} \mu \]
In other words, the probability $Pr(\lim_{N \to \infty} \bar{X}_n = \mu) = 1$.

The central limit theorem describes the size of the expected deviations of $\bar{X}_n$ from $\mu$ as $n$ is increased and relates them to the normal distribution:

$$\lim_{n \to \infty} Pr\left(\sqrt{n} (\bar{X}_n - \mu) \leq z \right) = \Phi \left( \frac{z}{\sigma} \right),$$

where $\Phi$ is the cumulative Gaussian probability distribution and $z$ is a free variable. One result of this theorem is that $Pr|\bar{X}_n - \mu| \propto 1/\sqrt{n}$ for averages of large samples of i.i.d. random variables.

This powerful theorem has broad applications in Physics. For example, consider a single particle executing a random walk in 1 dimension. The steps can each be considered i.i.d random variables, and as such the probability of finding the particle a distance $\Delta r$ away from its starting point after $n$ steps converges to a Gaussian distribution and $\langle \Delta r \rangle \propto \sqrt{n}$. The central limit theorem also plays an important role in error analysis, in which we use that the average of $n$ independent measurements with random errors approaches the expected value, and we can use the rate of the approach from the central limit theorem to define the standard error: $s = \sigma_x / \sqrt{n}$, where $\sigma_x$ is the standard deviation of the measurements, which we use because $\sigma$ isn’t observable.

As long as the positions of particles in an equilibrium system are uncorrelated, the number of particles in a given region can also be considered the sum of i.i.d random variables and the relative fluctuations scale as $\sqrt{\langle (\Delta N)^2 \rangle / \langle N \rangle} \propto \langle N \rangle^{-0.5}$. In physical systems, there is usually some correlation in particle positions, due to particle interactions. However, as long as the region we are examining has dimensions much longer than the particle’s correlation length, the Central Limit Theorem holds. As a result, equilibrium fluctuations are sometimes said to follow “law of large numbers scaling” [41].

1.2.2 Giant Number Fluctuations

In contrast to equilibrium systems, one of the results of symmetry breaking in the Toner-Tu equations is the production of giant number fluctuations in active matter [14, 27, 42]. Physically, this
can be thought of in terms of the natural current in the broken symmetry case of Equation 1.1. There is a soft Nambu-Goldstone mode related to the continuous symmetry breaking, such that fluctuations in particle orientation are strong. To understand the existence of this mode, consider again the ferromagnetic XY model and let one mesoscopic region located at \( \vec{r}_1 \) have a net magnetization, \( M\hat{n}_1 \), representing a ground state. Further, let another region at \( \vec{r}_2 \) have a different net magnetization, \( M\hat{n}_2 \), also representing a ground state. Remember that the continuous symmetry means that all rotations of \( M\hat{n}_1 \) are allowed ground states, so we can consider a continuous rotation to take place between them, as sketched in Figure 1.3a. As long as all of the particle interactions are local, if \( |\Delta \vec{r}| \to \infty \) the energetic cost of any rotation \( E_r \to 0 \), so that as long as finite fluctuations are allowed, long wavelength rotations in the local alignment direction should be expected.

Figure 1.3: a) An illustration of a Nambu-Goldstone mode. Because all orientations are allowed ground states and all interactions are local, arbitrary rotations become energetically cheaper as \( \Delta \vec{r} \) is increased. b) An illustration of a density fluctuation caused by a Nambu-Goldstone mode in a self-propelled system. Consider several equal spaced regions given by the bottom circles, each with a net velocity given by the black arrows. The net velocities have a soft Nambu-Goldstone mode allowing for rotations. After some time the particles from the original regions find themselves in the regions shown by the top circles. Notice that the density fluctuation caused by this mode is primarily in the axis perpendicular to the direction of the collective-motion.

In the collective-motion phase of an active fluid, Goldstone mode fluctuations in the velocity
field lead to large fluctuations in particle number because particles experience a local current set by the velocity field. Due to geometry, this effect is most pronounced on the axis perpendicular to the collective-motion. Consider several equally spaced regions in an active fluid, represented by the bottom circles in Figure 1.3b, that are mostly aligned except for the slight rotation caused by the Goldstone mode so that the net velocity in each region points in the direction of the black arrows. After some time, each region has moved in the direction of their velocities to the positions of the top circles, with the net result that the density fluctuations are strongest along the horizontal direction. Notice that these type of fluctuations can occur on arbitrarily long length-scales. In fact, it can be shown that the density fluctuations in the collective-motion phase are scale-free, which means there is no region we could choose to measure the number fluctuations that would be larger than the correlation length. This decouples number fluctuations in the collective-motion phase from the Central Limit Theorem [41, 43].

Carefully solving for fluctuations in the density field of the Toner-Tu equations leads to fluctuations that are on the order of $\langle (\Delta N)^2 \rangle \propto \langle N^{1.6} \rangle \gg \langle N \rangle$ [42, 44]. This means that strong fluctuations should be expected even in relatively large active systems, compared to similarly-sized equilibrium systems.

This effect was first seen experimentally in collections of shaken granular rods lying in 2D but shaken in the third dimension. Blair et al. demonstrated that, under suitable conditions, symmetric rods could self-organize into moving ordered domains of nearly vertical rods [34]. Close inspection showed that the motion is driven by a broken symmetry due to the rods tilting slightly away from the vertical. Each cycle of the vibration launches the tilted rods into the air. Then the rods land on their low end, which is not under their center of mass, so friction with the bottom plate pushes the grain in the direction of its tilt. This causes the rod to move in the direction of its tilt at every step, turning it into an active particle. The shaken grains can form local currents that lead to predicted giant number fluctuations [16, 45].
Figure 1.4: a) The positions and orientations of active disks during an experiment using polar vibrated disks. The boundary is flower shaped to avoid particles accumulating in corners, and data is only taken from the center of the flower. The color of the disks represents the strength of the local order, with darker red disks more highly ordered. The inset shows a diagram of a single disk. The disk has two asymmetric feet, causing it to travel in the direction of its smaller foot when it is shaken vertically. b) As the acceleration (scaled by gravity), $a_0$, of the shaking increases, the order parameter, $\Psi$, begins to disappear. The increase in acceleration corresponds to increasing the amplitude of the shaking which increases the rotational noise of the particles and decreases the collective-motion. c) The measured number fluctuations in measured regions in an experiment with $a_0 = 2.8$. The results show extremely strong number fluctuations. This figure is adapted from Reference [19].

This surprising method of creating active particles inspired the creation of circular particles with asymmetrical feet [19], like the one shown in the inset of Figure 1.4a. Vibrating these particles between two plates results in directed motion through the same mechanism as the rods discussed above, but the shape of the vibrated disk particles allows for simpler isotropic hard-body interactions [46]. Despite the isotropic shape of the particles, the particles experience an aligning interaction, brought about by a torque the particles feel when collisions force them to move in a direction different from their orientation. Figure 1.4a shows instantaneous positions and orientations of all of the particles during one of Designee et al.’s experiments. Notice that in the center of the cell, where wall effects are not as important, dense clusters of particles are aligned.

The result of this alignment is that experiments with active disks demonstrate the expected onset of collective-motion when the noise level in the system is low. The onset of collective-motion can be seen in the plot of probability distribution functions (PDFs) of instantaneous measurements...
of the order parameter, $\Psi$, near the center of the cell in Figure 1.4b. The order parameter, $\Psi$, varies from 0 when the particles are completely unaligned to 1 when the particles are perfectly aligned. As the maximum acceleration of the bottom plate, $a_0$ (scaled by gravity), is decreased, the orientational noise of the particles, $\eta$, also decreases proportionately and measured values of $\Psi$ increase.

The giant number fluctuations are measured by picking regions of different sizes in the cell, counting the number of particles in the regions for each instant in time and then calculating their variance $\langle (\Delta N)^2 \rangle$ and average $\langle N \rangle$. These two measures scale $\langle (\Delta N)^2 \rangle = \langle N \rangle^{1.45}$ for their experiments, as shown in Figure 1.4c. Recall that the number fluctuations scale as $\langle (\Delta N)^2 \rangle = \langle N \rangle$ for an equilibrium system, so the fluctuations measured by Deseigne et al. are indeed giant in comparison.

### 1.2.3 Sound Waves in Active Matter

The Toner-Tu equations were also immediately recognized to be capable of propagating small perturbations as sound waves [42], although the lack of Galilean invariance causes the speed of sound to vary with its orientation relative to the direction of collective-motion. These waves can be solved for by linearizing the Toner-Tu equations around the average density and the velocity of the collective-motion, $\rho_0$ and $u_0$, respectively. We can take $\rho(\vec{r}, t) = \rho_0 + \delta \rho(\vec{r}, t)$ and $\vec{v}(\vec{r}, t) = u_0 \hat{x} + \delta \vec{v}(\vec{r}, t)$ and look for plane wave solutions, $\delta \rho(\vec{r}, t) = A \exp(iwt - iq \cdot \vec{r})$. The speed of sound is then found to be

$$c(\theta) = \frac{1}{2}(1 + \lambda_1)u_0 \cos(\theta) \pm \frac{1}{2} \sqrt{(\lambda_1 - 1)^2 u_0^2 \cos^2(\theta) + 4\sigma \rho_0 \sin^2(\theta)},$$

(1.5)

where the pressure has been assumed to be linearly related to the density, $P = \sigma \rho$, with $\sigma$ a constant of proportionality. This shape of this theoretical curve can be seen for three different densities in the solid lines of Figure 1.5d-f. There are two speeds for each measured orientation because waves travel differently depending on whether they are traveling with or against the collective-motion. For
example, in the direction of collective-motion, $\theta = 0$, the fast propagation speed is $u_0$, the speed of the collective motion, and the slow propagation speed is $\lambda_1 u_0$, which is less than $u_0$, since these waves “try” to propagate in the reverse direction relative to direction of the collective-motion, but are nevertheless convected along in the direction of the flow. Waves traveling in any other direction have two components: one component that is convection in the direction of collective motion and an orthogonal component with a speed set by the compressibility of the fluid, $c_0 = \sqrt{\sigma \rho_0}$, as it would be in a momentum-conserving fluid. The direction that allows for the fastest propagation of waves is the one in which the wave “tries” to propagate completely in the direction orthogonal to the collective-motion while being simultaneously convected by the collective-motion.

Figure 1.5: (a-c) The power spectrum measured in a system of colloidal rollers undergoing collective-motion for three different propagation directions with respect to the flow direction, $\theta = \pi/4$, $\pi/2$, and $\pi/8$, respectively. The speed of sound is given by the slope of the dispersion relations at $q = 0$, $\omega = 0$. d-f) The angular dependence of the speed of sound for area fractions $\phi_0 = 0.11$, $\phi_0 = 0.18$, and $\phi_0 = 0.24$, respectively. The points represent their experimental measurements, while the solid lines show the fit to the theoretical expectation, Equation 1.5. This figure is adapted from Reference [47]

In spite of the above theoretical prediction, it took more than two decades for experimental measurements to become precise enough to detect these active sound waves. In 2018, D. Geyer, A. Morin, and D. Bartolo finally reported the first experimental observation and measurement of sound waves in the collective-motion phase of a system of colloidal rollers by computing their dispersion relation [47]. They tracked the particles in a system of colloidal rollers and then took
coarse graining averages to obtain density and velocity fields. Then they took dimensional Fourier transforms of the fluctuations of fields in time and space to obtain power spectra for their trials. By choosing axes inclined from the direction set by the average velocity by an angle, \( \theta \), to use in their calculation of the power spectra, Geyer et al. could detect the power spectrum for various propagation directions, like those in Figures 1.5a-c, which show the power spectra in a single trial with propagation directions \( \theta = \pi/4, \pi/2, \) and \( \pi/8 \), respectively. These sound waves were measured by taking the slope of the power spectra through the origin which represents the long wavelength limit that corresponds to sound modes. Some of their measurements of the speed of sound as a function of \( \theta \) for varying densities are plotted as points in Figures 1.5d-f, along with lines that show fits to Equation 1.5. Figures 1.5d-f show the speeds of sound for trials with packing fractions \( \phi_0 = 0.11, 0.18, \) and \( 0.24 \), respectively. Notice that the speed of sound in their trials increases with density as expected, especially in the directions with substantial components in the direction transverse to the collective motion.

Their results finally confirmed that sound waves propagate in every direction in active fluids and that their speed depends on the angle between their propagation direction and the direction of the collective-motion. By detecting these sound waves over a range of densities, Geyer et al. were able to measure nearly all of the material constants in the Toner-Tu equation for their colloidal roller system.

In the future, similar methods may be used to determine the material constants of different organic systems. Already, despite their complicated interactions, human crowds have been shown to propagate some kind of linear waves, with speeds that are reproducible across different crowds around the world [48], suggesting that human crowds also have measurable material constants.

1.3 The Nature of the Order-Disorder Phase Transition

1.3.1 Debate

When Vicsek et al. first published their results showing their model’s collective-motion phase transition, they argued that this transition was continuous [17]. They saw that the global order
parameter, which they defined as \( \Psi = |\langle v \rangle| \), scaled with the noise of the system as \( \Psi \propto |\eta_c(\rho) - \eta|^\beta \), where \( \beta \) is a constant critical exponent and \( \eta_c(\rho) \) is the critical noise, above which \( \Psi = 0 \) and below which \( \Psi \neq 0 \) for a constant \( \rho \). Notice that the particle speed is set to \( v_0 = 1 \) in their simulations, so the maximum order parameter is \( \Psi = 1 \). Likewise, they reported that \( \Psi \) scaled with the density of the system as \( \Psi \sim (\rho - \rho_c(\eta))^\delta \), where \( \rho_c(\eta) \) is a critical density that depends on noise and \( \delta \) is a different critical exponent. In both cases they found that the range of densities and noises for which this scaling held increased as they increased the system size, as shown in Figure 1.6a-b, presenting strong evidence of a continuous phase transition.

![Figure 1.6: a) Vicsek et al.'s original 1995 results comparing the order parameter, \( \Psi \) to the noise as the noise approaches the critical noise level, with density held constant. The power law scaling is indicative of a continuous phase transition. b) Vicsek's measurements of the order parameter as the density is varied and the noise is held constant near the phase transition. This too shows power law behavior and indicates a continuous phase transition. This figure is adapted from Reference [17].](image)

However, almost a decade later, G. Grégoire and H. Chaté published results challenging Vicsek et al.’s original result [49]. They reported that the phase transition was clearly discontinuous in simulations, and that this should become apparent for system sizes only slightly larger than Vicsek originally published. Figure 1.7a shows a probability distribution function (PDF) for instantaneous measurements of the order parameter, \( \Psi \), in different choices of noise, \( \eta \), in their simulations. They define
\[
\Psi(t) = \frac{1}{v_0} |\langle \vec{v}_i(t) \rangle_i|
\]
which is equivalent to the definition by Vicsek et al. In this case, $v_0$ is the speed of the particles in the simulations and $\vec{v}_i(t)$ is the velocity of the $i$th particle.

The two peaks in the PDF of $\Psi$ for each noise level near the transition are an elegant indication of a discontinuous phase transition because they show that the system switches back and forth between two phases. At noise levels below the critical point, the system spends most of its time in the state with finite order. As the critical point is reached, the system begins to spend more time in the globally disordered state, corresponding to values near $\Psi = 0$, but the system can still sometimes be found in an ordered state. An example of this discontinuous switching of the order parameter in time is shown in Figure 1.7c which shows instantaneous measurements of $\Psi$ over time. When the system size is increased, the time autocorrelation of the order parameter increases, indicating that switching between phases becomes less common. In the infinite size limit, it would be expected that the autocorrelation times become infinite and the transition becomes discontinuous. Figure 1.7b shows the average order parameter of simulations as a function of noise levels for different system sizes, $L$. The discontinuity in the curve for $L = 256$ is more evidence that the collective motion phase transition becomes discontinuous for large enough systems.
Figure 1.7: a) Grégoire and Chaté's simulation results. This diagram shows the probability distribution function (PDF) of the global order parameter over time, $\phi_t$, for noise levels close to the phase transition. The two peaks show the phase transition to be discontinuous with the order parameter switching back and forth from disordered to ordered over time, with the ordered state becoming more and more common as the phase transition is approached. b) Their simulations showed that order parameter became discontinuous with noise only for large enough system sizes than those used by Vicsek. et al. c) The order parameter as a function of time near the phase transition exhibits the switching suggested by the PDFs. This figure is adapted from Reference [24].

Vicsek et al. then published new results that argued the phase transition did still seem continuous for the original conditions while using the same measures proposed by Grégoire and Chaté, and that it only began to look discontinuous at very high choices of $v_0$ [50]. Figure 1.6a shows PDF’s of Vicsek et al.’s simulations, using a larger simulation than they had 12 years earlier and low particle velocities. In this case, the single peak for each noise level indicates a continuous phase transition, since there is no sudden switching back and forth between two distinct phases near the transition.

Vicsek et al. argued that the high velocity limit was unphysical, since it effectively meant that the particles were rarely updating their velocities to align with their neighbors. Further, they showed that under high velocity conditions, in which the transition looks discontinuous, there are
traveling fronts of aligned particles passing through the disordered background. An example of one of these fronts is shown in 1.8b. Vicsek et al. noticed that the fronts were always found traveling along one of the primary axes of the simulation cell, even though the cells had periodic boundaries. Because of this biasing, Vicsek et al. believed these waves to be a computational artifact and proposed that this artifact was responsible for the discontinuous behavior reported by Grégoire and Chaté.

Figure 1.8: a) From Vicsek. et al.’s 2007 paper: probability distribution functions (PDFs) of instantaneous measurements of the global order for varying noise levels near the phase transition. The single peak in the PDFs that shifts with noise level is indicative of a continuous phase transition. b) The traveling waves that Vicsek et al. considered a computational artifact. The waves always traveled along one of the primary axes of the simulated cell. In the figure, the cell is repeated to show its periodic boundary conditions. This figure is adapted from Reference [50].

In yet another rebuttal, Chaté et al. published results again for the same conditions and made sure their simulation was following the exact same rules as Vicsek et al.’s, asserting that the phase transition was discontinuous for any particle velocity [24]. This has been verified separately, and today it is generally accepted that the phase transition is in fact discontinuous [51, 52]. We emphasize that one should be careful when reading the early literature on this phase transition, understanding that this question was not resolved until about one decade ago.
1.3.2 Polar Bands

The waves reported by Vicsek at al. in their rebuttal of Grégoire and Chaté were not an artifact of the simulations. Chaté et al. went on to describe these waves as “polar bands”, in which the ordered collective-motion phase intrudes into the disordered phase on the leading edge of the wave and dissipates back into the disordered phase on the trailing edge [24]. A close up of one of these waves in simulation is shown in Figure 1.9a.

![Figure 1.9: a) An example of a polar band in simulation. b) Polar bands in Chaté et al.’s simulations. The color represents the local density, averaged along the y-direction. c) The measured density (left axis) and polar order parameter (right) averaged across the y-direction of a simulation with polar bands. d) The difference in orientation between the local and global order as a function of local density during a simulation with low noise. Points represent measurements and the black line shows average difference as a function of density. There are no traveling bands in this simulation. e) The difference in orientation between the local and global order as a function of local density during a simulation with moderate noise. The dependence of average order on density leads to phase separation. These figures have been modified from Reference [24].](image)

From a uniform initial disordered state with suitable choices of $\rho_0$ and $\eta$, the Vicsek model leads to spontaneous phase separation, as shown in the space time plot of the density over the course of one of Chaté and Grégoire’s simulations in Figure 1.9b. Here the $x$-axis is time, the $y$-axis is space along the long dimension of a rectangular simulation cell, and lighter colors represent higher density. Notice the quick formation of traveling bands, which are visible as straight lines of
high density, that then travel at a constant velocity, independent of their amplitude.

Figure 1.9c shows a snapshot of the density and local order during a simulation like the one in 1.9a. In this case, Grégoire and Chaté have binned the particles by the position along the long axis of the cell and then measured the density and calculated $\Psi$ separately for each bin, effectively averaging the density and order over the direction perpendicular to the propagation of the bands. The left $y$-axis shows the density of each bin, and the right $y$-axis shows $\Psi$ for each bin. Clearly the traveling waves are characterized by not just much higher density than the background, but also by high order, demonstrating the coexistence of low density, disordered regions and high density, ordered regions.

The waves can be thought of as a result of a “local equation of state,” in which the local density is related to the local order. To examine this, Chaté and Grégiore divided the system up into regions much smaller than the width of a traveling band and calculated the density, $\rho_l$, and average orientation, $\theta_{\text{local}}$, for each of these small regions. Then they compared $\theta_{\text{local}}$ to $\theta_{\text{global}}$, the average orientation of all the particles in the system, which corresponds to the global broken symmetry.

For one choice of $\rho_0$ and low $\eta$ that places the system in the homogeneous collective-motion phase, there is little difference in orientation between high density and low density regions, as seen in Figure 1.9d. The low density regions are also aligned in the direction of the broken symmetry. However, if the noise level is increased to a level that puts the system in the coexistence region for the same $\rho_0$, low density regions become disordered, as indicated by the plateau below $\rho_l = 2$ in Figure 1.9e, and the high density regions remain ordered. This relationship between density and order is thought of as a local equation of state.

In another comparison to discontinuous equilibrium phase transitions, the Vicsek model shows hysteresis around the phase transition. In simulations with a constant $\eta$, slowly increasing the density of the system results first in small dense clusters that move together in random directions, like those shown in Figure 1.1b. Increasing the density further leads to these small clusters merging into a single band that spans the shortest axis of the simulation cell [53]. This merging breaks the rotational symmetry of the system and causes a step in the order parameter of the system, as shown...
in the blue curve of Figure 1.10a. If the density is increased still further, the traveling bands multiply and eventually fuse, and the system transitions into the homogeneous collective-motion phase.

![Figure 1.10: a) An example of hysteresis around the phase transition. The blue curve shows the average global velocity of the system as its density is slowly increased, and the green curve shows the average global velocity as the density is slowly decreased. Notice that the density at which the first band forms and the density at which the last band dissolves are not the same. Figure reproduced from Reference [54]. b) A generic phase diagram for the Vicsek model when varying density and noise. The coexistence region lies between the binodals ($B_{gas}$ and $B_{liq}$). Between the spinodals ($S_{gas}$ and $S_{liq}$), any fluctuation from a homogeneous state results in phase separation. Figure reproduced from Reference [51].](image)

The reverse process, gradually decreasing $\rho_0$ while holding $\eta$ constant, results first in the collective-motion phase breaking into bands and then eventually the bands dissolving into individual clusters. However, the critical densities for these transitions are now lower. For example, consider the green curve in Figure 1.10a in which the density is gradually lowered and the critical density for the last band to dissolve is less than it was for the first band to form in the blue curve.

This hysteresis is reminiscent of discontinuous phase transitions in equilibrium systems, such as the liquid-gas transition, which prompts us to think of this hysteresis as a result of binodal and spinodal curves and create a phase diagram, like the one in Figure 1.10b.

Although they were at first controversial, polar bands have been verified in hydrodynamic simulations based on the Toner-Tu equations [55] and observed in experiments with colloidal rollers [36]. Finally, they have been described with an elegant solution to the Toner-Tu equations by Caussin et al. [56]. Now, traveling bands are considered to be strong evidence of a discontinuous collective-motion phase transition [51].
1.4 The Search for State Variables

1.4.1 An Effective Temperature

Effective temperatures have previously been used to describe the behavior of a variety of systems, such as glasses [57], foams [58], and driven grains [59]. In 2004, Ohja et al. tested how useful an effective temperature could be by measuring the statistics of a ping pong ball buffeted by a uniform updraft of air [60] and comparing its behavior to equilibrium systems, which satisfy the fluctuation-dissipation theorem, which amounts to having $D = \mu k_B T$, where $D$ is the diffusion constant and $\mu$ is the particle’s mobility. The only requirement for the fluctuation-dissipation theorem is that a system obeys detailed balance, which is the requirement in an equilibrium system that every process is in balance with its reverse process [6]. In this case, diffusion is the result of a process whereby thermal fluctuations impart kinetic energy to a larger particle. Detailed balance requires that this process be in balance with its reverse process, which is the friction (dissipation) in the system that converts the kinetic energy back into thermal fluctuations. The relationship between these two processes is provided by the thermal energy $k_B T$.

The updraft caused turbulence around a ping pong ball resting on a screen, which in turn buffeted the ball around such that it executed a random walk. The speed distribution of the ball fit the Maxwell-Boltzmann distribution, which describes the velocity distribution of particles in a classical ideal gas as a function of only the particles’ mass and the temperature. Because Ojha et al. knew the mass of the ball, they could easily extract an effective temperature, which they found to be $10^{15}$ times larger than room temperature.

They then independently measured the diffusion constant by tracking the ping pong ball’s position and the drag coefficient, $\mu$, of the rolling ball by periodically tilting the system and looking for resonance. In the end, the effective temperature they found from the ball’s velocity distribution was the same as the effective temperature needed to fulfill the fluctuation-dissipation theorem for their system. Their results showed their system obeyed the fluctuation-dissipation theorem, which meant that their effective temperature was useful in describing the dynamics of their system and
that their system could be mapped onto an equilibrium system.

1.4.2 Effective Temperature in Active Matter

Because, like the turbulence in the previous section, activity seems to play a role similar to an effective temperature, some groups have pushed to define effective temperatures for their active systems. Palacci et al. were the first to use sedimentation to probe for an effective temperature in active particles [61].

They used a relatively new class of colloidal particles, known as Janus particles, that react with solvents asymmetrically [62]. This asymmetry allows Janus particles to deplete to interfaces and act as strong emulsifiers or function as a mechanism for self-assembly if one side of the particle has low affinity for the solvent. However the most interesting potential of these particles for this review is the potential to turn them into active particles by coating one side with platinum, like the particles in Figure 1.11a, and then submerging them in hydrogen peroxide. The platinum catalyzes a reaction with the hydrogen peroxide that produces \( \text{O}_2 \) bubbles behind the particles propelling them away from their platinum side. Because of the Janus particles’ scale, thermal diffusion controls their rotation, leading to wandering, self propelled particles [63]. Spherical Janus particles can also conspicuously lack any kind of effective aligning potential, allowing them to sidestep the collective-motion that characterizes much of active matter.

Palacci et al. calculated an effective temperature for the system of Janus particles by measuring their diffusion constant from their particles mean squared displacement as a function of time

\[
\langle (\Delta r^2) \rangle = D_{\text{eff}} t
\]

in very dilute environments, for varying fuel levels. For high fuel levels, the Janus particles were ballistic on very short timescales and diffusive at long times. However, the effective diffusion constant of the particles with a high fuel level was found to be much larger than the diffusion constant of the passive particles, as shown in Figure 1.11b. This figure shows the mean squared
displacement of the particles with fuel (red) and without fuel (blue) as a function of time. At long times, the mean squared displacement of the particles in both experiments is diffusive and the diffusion constant can be measured from the slope of the line. Higher fuel levels correspond to higher activity, so clearly, the activity plays a similar role to temperature in making the particles diffuse faster.

By comparing the diffusion constant to the drag forces in the medium, Palacci et al. were able to calculate an effective temperature under the assumption that the fluctuation-dissipation theorem held for their system, just as it had for Ohja et al.’s ping pong ball. Because the Janus particles were self propelled, their velocity distribution no longer follows the Maxwell-Boltzmann distribution, but Palacci et al. wondered if the effective temperature might still be useful in an equation of state that related the pressure to the density and temperature.
Figure 1.11: a) A Scanning Electron Microscope image of the Janus particles used for Palcci et al.’s sedimentation experiments demonstrating that only one hemisphere of each particle is coated with platinum. b) The mean squared displacement of the particles with (red) and without (blue) fuel. At long times, the particles diffuse. c) The density as a function of height in sedimentation experiments with different fuel levels. The sedimentation length of the particles increases with increasing fuel. d) The gravitational length from the sedimentation length is directly proportional to the effective diffusion constant for varying fuel levels, suggesting the usefulness of an effective temperature. This figure is adapted from Reference [61].

For an ideal gas, the equation of state relates the state variables of pressure, $P$, density, and temperature.

$$P = \rho k_B T$$

This allows us to straightforwardly calculate the density gradient that results from an applied gravitational field by balancing the stress on a horizontal slice of the gas with the assumption of a constant temperature:

$$dP = -\rho mg dz$$

where $m$ is the mass of a single particle of the gas, $z$ is the vertical position of the slice, and $g$ is
the acceleration due to gravity. Then we use the equation of state to rewrite the left hand side.

\[ k_B T d\rho = -\rho mg dz, \]

to find that

\[ \rho(z) = \rho_0 \exp \left(-\frac{z}{l_g}\right) \]

where \( l_g = kT/(mg) \) is called the sedimentation length scale.

Palacci et al. measured density of their active Janus particles under the influence of gravity and found that the density was an exponential function of height, with a sedimentation length controlled by their previously calculated temperature, as shown in Figure 1.11b. Particles with more fuel had larger \( D_{\text{eff}} \) and larger \( l_g \). In fact they found the ratio of \( l_g/D_{\text{eff}} = 1/(\mu mg) \), so that the temperature found by each of their two methods agreed (see Figure 1.11c). Based on this result, they proposed that the diffusion rate could be used to define an effective temperature for the active Janus particles and that this temperature could be used as a state variable.

The universality of Palacci et al.’s results, however, has been challenged by Enculescu and Stark [64] and by Szamel [65], who together claim that Palacci et al.’s results are only valid in a narrow class of cases. Enculesco and Stark show that the measured equation of state only holds for relatively low activity levels. At moderately high activity levels, an effective adhesion begins to kick in and decouple diffusion from the gravitational length, so that the effective temperature calculated by the two methods disagree. Szamel reports that the effective temperature derived from diffusion can also be decoupled from the stationary state by changing the confining potential from linear in \( z \), such as that due to gravity, to harmonic. If an effective temperature is only valid for a narrow class of potentials, then it cannot be a state variable.
Ginot et al. performed experiments on active Janus particles in 2D that extend Palacci’s results beyond the ideal gas limit [66]. They performed sedimentation experiments at an incline so that the effects of gravity were weaker and used much higher densities and fuel levels than Palacci et al. as shown in 1.12a. Through their sedimentation experiments, they argued that an effective temperature was still a useful tool at moderately high densities if one properly accounts for the activity-dependent adhesion that results from the balance between self-propulsion and rotational diffusion, as shown in Figure 1.12b. This adhesion could be used to define the first two virial coefficients in an equation of state and maintain the usefulness of the effective temperature.

Theoretical and simulation work continue to cast serious doubt on the interpretation of these experimental results. For example, Solon et al. have proposed that even pressure cannot be used as a state variable for most active particles, calling into doubt all state equations derived from sedimentation [54] and the whole idea of a single definition of an effective temperature. In their simulations of active particles without an aligning potential, they gave particles different self-propulsion speeds based on whether their instantaneous position was on either of two sides of a simulation cell. Experimentally, this situation could be realized by Janus particles with light...
induced propulsion if the light is brighter in part of the system. This resulted in a different diffusion constant and therefore different effective temperature between the two halves. They further found that this rule naturally caused more particles to accumulate on the side of the cell where they self-propel at a lower speed, as shown in Figure 1.13a.

![Figure 1.13](image)

**Figure 1.13:** a) A simulation explicitly demonstrating Solon et al.’s results on pressure in active systems. The self-propulsion speed of any particles that find themselves on the right are higher than those found on the left, and there is no physical barrier separating the two sides. b) By varying the levels of self-propulsion in the two halves of the cell, they demonstrate that the system fails to equalize pressure. \( v_1 \) and \( P_1 = \rho v_1^2 \) are the propulsion speed and measured pressure in the left half of the box, while \( v_1 \) and \( P_2 \) are the corresponding values in the right half of the box. The measured pressure inside each box follows the ideal gas equation of state, where the effective temperature has been measured from the diffusion constant of the particles. c) Instead of equalizing pressure between the two halves, the system evolves to equalize \( \rho v \). This Figure is adapted from Reference [54].

However the two halves of the cell failed to reach mechanical equilibrium by equalizing pressure, \( P \propto \rho_0 v^2 \), as shown Figure 1.13b. At high ratios of speeds, \( v_2/v_1 \), where \( v_1 \) and \( v_2 \) are the self-propulsion speeds of the particles in the left and right sides of the cell respectively, the pressures in the two halves of the of the cell became increasingly different. Instead, the particles tended to equalize the quantity \( \rho v \), as shown in the inset of Figure 1.13c; we will return to this
1.4.3 Explicit Violations of Detailed Balance

All of the previous studies have relied on the fluctuation-dissipation theorem to define an effective temperature by measuring diffusion constants. However, when measuring a diffusion constant for active particles or for tracer particles embedded in a bath of active particles, one must be careful to take the long time limit and also to avoid the collective-motion phase. For example, Xiao-Lun Wu and Albert Libchaber found that collisions with bacteria caused the inactive particles to undergo super-diffusion on short timescales and that the crossover time between super-diffusion and normal diffusion occurred at longer times for higher concentrations of bacteria [67]. Super-diffusion occurs because there are strong spatial and temporal correlations in the forces that inactive particles feel due to flow fields generated by the active bacteria. These strong correlations decouple the motion of the tracer particles from the fluctuation-dissipation relation, so it is not clear that long-time diffusion constants should be an effective measure of an effective temperature.

Violations of the fluctuation-dissipation theorem remind us that active systems do not obey the principle of detailed balance, but the lack of detailed balance in active systems can led to other striking results. In an equilibrium system, detailed balance prevents “perpetual motion machines of the second kind,” which are machines that extract heat from a source and convert it completely into another type of energy. Sometimes, these hypothetical machines are called “Maxwell’s demons.” Ultimately, these machines are expected to fail because detailed balance requires the machine to return all energy it gains from the fluid’s thermal fluctuations back to the fluid in the form of heat. However, it is sometimes puzzling to understand the microscopics of why these machines don’t work. One machine, proposed by Richard Feynman, consists of asymmetric gears embedded in a fluid, so the collisions between the fluid particles and the gear are biased to rotate the gear in one direction, [68]. Simulation and theory work by Broeck et al. shows the details of how these kinds of gears maintain detailed balance when they are only in contact with equilibrium baths [69]. In other words, there is no way to rectify the random motion of fluctuations to get unidirectional
However, Sokolov et al. have shown that gears with asymmetric teeth, like those proposed by Feynman, can rectify fluctuations with the result of unidirectional motion when they are embedded in a fluid film along with bacteria [70], even when the bacteria are not in the collective-motion phase.

1.5 Motility-Induced Phase Separation

1.5.1 Motility-Induced Potentials

The finding by Solon et al. [54] that systems with varying levels of self-propulsion equalize $\rho v$ can be explained by considering the master equation for the probability, $P(\vec{r}, \hat{u})$, of a particle to have a certain position, $\vec{r}$, and orientation, $\hat{u}$, given a spatially dependent speed, $v(\vec{r})$. This equation reads:

$$\dot{P}(\vec{r}, \hat{u}) = -\nabla \cdot [v(\vec{r})\hat{u}P(\vec{r}, \hat{u})] + \Theta[P(\vec{r}, \hat{u})]$$

(1.6)

where $\Theta[P(\vec{r}, \hat{u})]$ is the average rate of change in orientation of the particle at a given $\vec{r}$ and $\hat{u}$; for isotropic systems without any alignment $\Theta[P(\vec{r}, \hat{u})] = 0$. Equation 1.6 can be viewed as the continuity equation for $P(\vec{r}, \hat{u})$. Alternatively, we can consider the two ways $P(\vec{r}, \hat{u})$ can change. There may be a net probability current that would allow more particles to flow into that position due to the directed motion of the particles in the surrounding regions or there may be particles already in that location with different orientations that can rotate to increase $P(\vec{r}, \hat{u})$. The steady state solution of Equation 1.6 without alignment is found by eliminating the divergence of $[v(\vec{r})\hat{u}P(\vec{r}, \hat{u})]$. Because active systems without alignment are typically isotropic, the steady state is $[v(\vec{r})\hat{u}P(\vec{r}, \hat{u})] = \text{const}$. In an equilibrium system, $v(\vec{r})$ is a random variable whose distribution depends on the temperature and not on $\vec{r}$. In this case, the trivial steady state for Equation 1.6 is $P(\vec{r}, \hat{u}) = \text{const}$. A particle is equally likely to be found anywhere in the system, so the density of the system is uniform as expected.

However, if we let $v(\vec{r})$ vary inside the system, as in Solon et al.’s simulations, the steady state
of Equation 1.6 is \( P(\vec{r}, \hat{u}) = C_0 / v(\vec{r}) \), where \( C_0 \) is a constant. The average density at \( \vec{r} \) is directly related to the probability of finding a particle at \( \vec{r} \) through the number of particles in the system:

\[
\rho(\vec{r}) = N_{tot} \int_0^{2\pi} P(\vec{r}, \hat{u}) d\theta \propto \frac{N_{tot}}{v(\vec{r})}
\]

This means the average density is higher where the particle velocity is lower so that \( \rho(\vec{r}) v(\vec{r}) \propto N_{tot} \) is constant over the whole system as shown in Figure 1.13c.

To first order, this motility-induced spatially-dependent density in this active system can be modeled as an equilibrium system with an extra energetic potential related to the motility. We can see this by considering the continuity equation for the probability distribution for the location of a single particle to be at \( \vec{r} \):

\[
\dot{P}(\vec{r}) = -\vec{\nabla} \cdot \vec{j}(\vec{r})
\]

where \( \vec{j} \) is the probability current brought about by diffusion and the coarse grained drift velocity, \( \vec{V}(\vec{r}) \), of the particle caused by an external potential:

\[
\vec{j} = -D(\vec{r}) \vec{\nabla} P(\vec{r}) + \vec{V} P(\vec{r})
\]

For passive particles, a uniform steady state, \( \vec{j}(\vec{r}) = 0 \), is reached when \( P(\vec{r}) = \text{const} \) and \( \vec{V}(\vec{r}) = 0 \). For our motility-induced, non-uniform, probability distribution, \( P(\vec{r}) = C_0 / v(\vec{r}) \), we can still find a steady state, but in this case, we need to assume an effective driving force.

The particle’s diffusion coefficient is no longer constant over space but is instead set by the particle’s propulsion speed and orientational relaxation time, \( \tau \).

\[
D(\vec{r}) = \frac{v(\vec{r})^2 \tau}{d(d-1)}
\]

where \( d \) is the dimensionality of our system.
Setting Equation 1.8 to zero results in

\[
\frac{\vec{V}(\vec{r})}{D(\vec{r})} = -\vec{\nabla} P(\vec{r}) = -v(\vec{r}) \vec{\nabla} \left( \frac{1}{v(\vec{r})} \right) = -\vec{\nabla} \ln(v(\vec{r})) \tag{1.10}
\]

Now, for our model equilibrium system, we recall that the fluctuation-dissipation theorem relates the diffusion coefficient to the particle mobility:

\[
D(\vec{r}) = \mu(\vec{r}) k_B T = k_B T \left| \frac{\vec{V}(\vec{r})}{\vec{F}(\vec{r})} \right| \tag{1.11}
\]

where we have used the equation for the required force, \( \vec{F}(\vec{r}) \) required to induce a drift velocity \( \vec{V}(\vec{r}) \): \( \vec{F} = \mu \vec{V} \).

Substituting Equation 1.11 into Equation 1.10 results shows

\[
\vec{F}(\vec{r}) = -k_B T \vec{\nabla} \ln(v(\vec{r})) \tag{1.12}
\]

We then see that the spatial variation in motility in our active system has the same effect as an external potential in an equilibrium system, \( \ln(v(\vec{r})) = \beta U_{\text{eff}}(\vec{r}) \), and that the probability of finding our particle at a specific location is

\[
P(\vec{r}) = \frac{C_0}{v(\vec{r})} = C_0 \exp(-\beta U_{\text{eff}}(\vec{r})) \tag{1.13}
\]

This relation can be generalized and is still valid to leading order when the functional dependence of \( v \) on \( \vec{r} \) is through the density, \( v(\rho(\vec{r})) \) [71]. If \( v(\rho(\vec{r})) \) is a monotonically increasing function of \( \rho(\vec{r}) \), then the particles experience an effective repulsion. On the other hand, if \( v(\rho(\vec{r})) \) decreases with increasing \( \rho(\vec{r}) \), the particles experience an effective attraction. This explains why Ginot et al. found the adhesive parameter useful in writing an equation of state in Reference [66].
1.5.2 Phase Separation

However, if \( v(\rho(\vec{r})) \) decreases too strongly with \( \rho(\vec{r}) \), our comparison with an effective potential breaks down because we never reach a steady state.

To see this, consider a system that starts out with a uniform particle speed, \( v_0 \). By our previous analysis, this results in a uniform density, \( \rho_0 = C_0/v_0 \). However, now let there be a small density fluctuation so that the density is \( \rho(\vec{r}) = \rho_0 + \delta \rho_1(\vec{r}) \). This results in a spatially dependent speed,

\[
v(\rho_0 + \delta \rho_1(\vec{r})) \approx v_0 + \frac{dv}{d\rho} \bigg|_{\rho_0} \delta \rho_1,
\]

which we expect to result in a new steady state density:

\[
\rho_0 + \delta \rho_2 = \frac{C_0}{v_0 + \frac{dv}{d\rho} \bigg|_{\rho_0} \delta \rho_1} \approx \frac{C_0}{v_0} \left( 1 - \frac{1}{v_0} \frac{dv}{d\rho} \bigg|_{\rho_0} \delta \rho_1 \right) = \rho_0 - \frac{\rho_0}{v_0} \frac{dv}{d\rho} \bigg|_{\rho_0} \delta \rho_1
\]

so

\[
\delta \rho_2 = -\frac{\rho_0}{v_0} \frac{dv}{d\rho} \bigg|_{\rho_0} \delta \rho_1
\]

If the resulting “steady state” seems to have larger density variations than the original fluctuations, \( \delta \rho_2 > \delta \rho_1 \), there is no steady state, and the system is unstable; fluctuations will continue to grow with time.

This occurs when

\[
\frac{dv}{d\rho} \bigg|_{\rho_0} < -\frac{v_0}{\rho_0}
\]

The positive feedback loop results in a few regions in the system becoming denser and denser, depleting the surrounding areas. This continues until the system reaches a steady state with two coexisting phases, a gas phase and a dense, sometimes crystalline, phase in a process known as Motility Induced Phase Separation (MIPS) [71]. Like the aligning potential in the Vicsek Model, the density-dependent speed in MIPS may a rise from a variety of mechanisms, such as fuel depletion when too many active particles are close together or chemotaxis in bacteria that slow down
when they come in proximity of each other to form a colony [72].

MIPS can even arise in active matter made of particles that only interact via collisions. Examples of this come from simulations and experiments using hard active spheres, like Janus particles [73, 74]. In this case, we can understand MIPS as a kinetic phase separation. Consider two Janus particles that collide head on. Their self propulsions act as an attracting force and a hard body repulsion repels them, resulting in the pair remaining stationary. They will remain in contact until rotational diffusion allows them to propel past each other. The overall result is an effective adhesion, as seen by Ginot et al. [66]. If their rotational diffusion is too slow, more spheres may join the collision before the first two separate, seeding a droplet of jammed active spheres.

Figure 1.14: MIPS in Linden et al.’s Janus particles. The images are photos of the particles after a) 0.02s, b) 0.4s, c) 2s, d) 5s, e) 20s, f) 68s. The scale bar is 100µm. This figure is reproduced from Reference [75].

Because it has been so difficult to create active particles with effectively zero alignment, experimental evidence of MIPS is very recent. In 2019, Linden et. al. finally reported arrested MIPS in their experimental system of Janus particles [75]. Their Janus particles had a huge persistence length of about 4000σ, where σ is the radius of a single particle. With this persistence length, the usually small hydrodynamic aligning potentials are negligible, and the speed of the
particles is slowed only by hard body interactions. This results in a nearly ideal case of MIPS, that is only halted when the slight aligning interactions of the Janus particles cause stresses in the crystal droplets that eventually fragment the droplets. Figure 1.14 shows the progression of their system from just after the activity of the particles is turned on. In Figure 1.14a-e, the clusters of hexatic crystals grow extremely quickly. After the time shown in Figure 1.14e, the system reaches a steady state in which the growth of the crystals is held in check by fragmentation due to the aligning interactions.

1.6 Previous Physics Work on Ants

*Solenopsis invictae*, commonly referred to as “fire ants”, are one of the several species of ants known to self-assemble into useful structures in nature [76]. Native to the Amazon rain forest, they form living rafts to survive the frequent floods. They can also arrange themselves into towers to scale surfaces that their oily clawed feet cannot grip and bridges to shorten foraging routes. Close observation by Mlot et al. [77] has demonstrated that these structures are formed by the ants latching onto each other with their mandibles and claws and that these connections are continuously broken and reformed, allowing individual ants to traverse the system. This tendency of fire ants to grab onto each other yields an interesting physical system - an active viscoelastic material.
Tennenbaum et al. have found that dense aggregations of fire ants, like the rafts that the ants naturally build, undergo “activity cycles” [13]. The ants spend long periods of time with many ants stationary, but then the group spontaneously activates to a state in which all of the ants are moving. Tennenbaum et al. loaded the ants into a rheometer, a machine that can precisely measure the strain and strain rates of a material under an applied stress, and observed the effects of these cycles.

The presence of the activity cycles in the ants could be clearly seen in the strain of the ants with no applied stress and simultaneously in the normal force the ants exerted on the upper tool of the rheometer. Figure 1.15a shows the strain of the ants over several hours without any applied stress. For much of the experiment, the ants hardly move the plate of the rheometer, but about once per hour, the ants activate and spontaneously rotate the plate. These spikes in strain rate correspond...
exactly to an increase in the normal force the ants exert, as shown in Figure 1.15b, so that active ants seem to exert a larger pressure on their surroundings.

Tennenbaum et al. also showed that the effects of the activity cycles had to be due to the ants’ ability to rearrange because packing the ants too tightly washed the active effects away [78]. They modeled the activity cycles as a catalytic reaction, where the rate of activation of new ants, \( \frac{dN_a}{dt} \), was proportional to the number of already active ants in the chamber, \( N_a \) and the amount of time the ants had been active, \( t_a \)

\[
\frac{dN_a}{dt} = k_a N_a t_a
\]

Frequency sweep measurements differed between activity peaks and inactive periods, indicating that the material properties of the active ants were distinct from those of the inactive ants. Figure 1.15c shows the elastic (dark squares) and viscous (light squares) moduli, \( G' \) and \( G'' \) respectively, as a function of frequency of the ants during an activity cycle. The elastic modulus, \( G' \), measures the amount of energy stored during each cycle, and the viscous modulus, \( G'' \), measures the amount of energy dissipated in each cycle. While active, the ants have almost exactly the same elastic and viscous moduli over a large range of frequencies, as seen in Figure 1.15c, suggesting that the active ants lack a characteristic relaxation timescale. In an equilibrium system, the linear responses \( G' \) and \( G'' \) are not independent, but are instead related to each other through the Kramers-Kronig relation. In the case in which \( G' \) and \( G'' \) are roughly equal, the Kramers-Kronig relation dictates that they must each scale as \( G' = G'' \propto \omega^{0.5} \) [79, 80]. Remarkably, this is close to the scaling measured by Tennenbaum et al. in Figure 1.15c, even though the ants are far from equilibrium so the Kramers-Kronig relation doesn’t need to hold. On the other hand, when the ants were inactive, they behave much like an elastic solid, as evidenced by an elastic modulus that is larger than their viscous modulus across the measured frequencies in Figure 1.15d.

Vernerey et al. have explained some of the rheology results for ants by modeling the ant aggregations as an active network, in which connections between individual ants are continuously broken and re-formed [81]. On short timescales this leads to elastic behavior because the ants don’t have time to rearrange their interconnections. The connections between individuals are maintained
but stretched, yielding an elastic response. On long timescales the ants are viscous because the stressed connections between ants are broken and new connections are formed between nearby ants, as shown in Figure 1.16. They further postulated that the rate at which these connections were broken and re-formed was stress dependent and predicted a crossover at slightly higher frequencies than those measured in Figure 1.15c, in which $G'$ begins to dominate so that the ants are primarily solid-like on short timescales.

Figure 1.16: The response of model ants to a shear demonstrated by the distributions, $\phi$ of the lengths of the links between the ants, $r$. The shear initially results in a stretching of some of the ants’ connections. This causes a short term elastic response. However, given enough time, the activity of the ants results in them breaking connections and forming new connections to their current nearest neighbors, dissipating the stored elastic energy. This Figure is reproduced from Reference [81].

Separatedly, Gravish et al. have compared fire ant traffic flow in confined tunnels to a fragile glass transition [82]. The glass transition is an out-of-equilibrium transition in which the dynamics of a liquid slow down, until the timescale of rearrangements becomes comparable to the timescale associated to the cooling rate, and eventually diverges. In a fragile glass transition, this slowdown happens relatively quickly as the system approaches the glass transition temperature. The observations of Gravish et al. were comparable to a fragile glass transition.

They established two-way traffic through plastic tunnels by instigating foraging behavior and used real space imaging to track the 1D density throughout the tunnels as a function of time, as shown in Figure 1.17a. They noticed that sometimes groups of ants jammed the tunnels and that the duration of these jams increased as more ants tried to use the tunnels, until eventually a jam lasted for several minutes.
Figure 1.17: a) Space-time surface of experimental observations of ants confined to a tunnel. $\tau$ is the relaxation time and corresponds to the temporal length of a traffic jam. b) One of the results from Gravish, et al.'s simulations of ant traffic with varying interaction times. Shorter social interactions, $T_{int}$, lead to a fragile glass transition, while longer interactions lead to a strong Arrhenius-type glass transition. Experimentally measured interactions are of an order that leads to a fragile transition. This figure is adapted from Reference [82].

The ants tended to pause when they came into contact, so Gravish et al. postulated that social interactions function as an effective localizing potential that can lead to jamming in a manner similar to the glass transition. They built a simulation based on the ants’ behavior and modified the interaction times, finding that increasing the interaction times led to less and less fragile behavior so that the slowdowns were felt over a larger range of densities, as shown in Figure 1.17b. Comparing their simulation results to their experiments, Gravish et al. determined that ants have short enough interactions that tunnel jamming behaves like a fragile glass transition. This helps the ants keep their tunnels clear of jams for larger ranges of density.

1.7 Conclusion

All in all, there has been significant progress in the field of active matter over the last couple of decades. We have explored the Vicsek Model and the Toner-Tu equations, which are together the most important models of active matter. In discussing how active matter systems can support fluctuations that differ from the Central Limit Theorem and how they can support global symmetry breaking in 2D, we saw some of the important differences between active and equilibrium systems.

We then covered the different types of phase transitions that can arise based on the type of
inter-particle interactions. Aligning interactions tend to lead to a phase transition between a disordered state and collective-motion and motility interactions in the absence of alignment can lead to motility-induced phase separation. We believe the field of active matter is in need of new experimental systems to test the generality of some of its conclusions because most of the important experimental results come from a few synthetic systems, which work best under relatively dilute conditions. We would like to systematically study fire ants as self-propelled particles to see if their behavior mirrors that of active systems with simpler interactions. Although interesting work has already been done on fire ants, much of this work has focused on ants as a material and not as a system of self-propelled particles. We hope that the following chapters will push the field forward by testing the generality of models of active matter.
CHAPTER 2
ANT BEHAVIOR AND INTERACTIONS

2.1 Harvesting and Storing Ants

Fire ants were introduced to the United States in Mobile, Alabama, in the 1930s. Proving to be an invasive species, fire ants have spread quickly and are now common across all of the southeastern United States [83, 84]. In their native habitat in South America, fire ant colonies are routinely threatened by floods, which can wash away their mounds. In response, they have evolved to cling to each other and form living rafts to escape rising water [77].

We collect our fire ants from the wild in a vacant lot in Kennesaw, GA, (34°01’10.7”N 84°31’36.6”W) between March and October. During the winter months, the ants move deeper underground to avoid the cold, making them difficult to locate. Even when ants can be located in the winter months, we find that their behavior is significantly different from ants collected in warmer months. For example, ants collected in the winter fail to expand in the column experiments discussed in Chapter 4.

When we collect wild ants, we search for large above ground mounds that are aggressively defended, which gives us an idea of the size and health of the colony. Experience has shown that it is easier to find large mounds and judge the size of colonies on warm days 24 to 72 hours after heavy rainfall. After rain, the ants repair their tunnels, resulting in fresh mounds. We have also learned that we are likely to find the largest colonies next to, but not on, low lying land where water accumulates.

To collect the ants, we slick the walls of five gallon buckets with talcum \( \text{Mg}_3\text{Si}_4\text{O}_{10}(\text{OH})_2 \) powder to prevent ants from climbing out of them and then seek out the largest colonies available. Upon finding a desirable wild colony, such as the one shown in Figure 2.1a, we collect the ants along with all of the soil of their mound and subterranean tunnel system with a shovel. We continue
collecting soil until it seems that there are few ants left in the surrounding soil, typically to a depth of about 20 cm, leaving a hole like the one shown in Figure 2.1b and filling each bucket no more than two-thirds of the way to the rim.

Once we have brought the ants to the lab, we leave the ants undisturbed in the buckets for one full day to allow the ants to form a tunnel system, as shown in Figure 2.1c. Then we use a system of irrigation tubing to drip water into the soil, as shown in Figure 2.1d, at a rate of about $1 mL/s$, flooding the buckets over the course of another two days. Dripping the water allows the ants to make their way to the surface of the soil without drowning [85]. As the water level continues to rise past the surface of the soil, the ants form rafts on the water’s surface. We remove them with a spoon and place them into large open Tupperware bins, Figure 2.1e. We slick the walls of the bins with Polytetrafluoroethylene, also known as Fluon, a copolymer which keeps the ants from climbing out of the lidless bin. These bins make for convenient storage because they allow for the easy removal of ants for experiments by keeping the ants away from soil.
Figure 2.1: a) A fire ant mound in Kennesaw, Georgia. It has been disturbed with a shovel to examine the aggression of the ants’ response, which we take to signal the health of the colony. b) The same region of soil after the mound and surrounding soil have been collected along with the ants. c) The ants and soil in a five gallon bucket 24 hours after the ants were collected. d) Several buckets that together hold the soil and ants from a single colony and a system of irrigation tubing that slowly drips water. The ants eventually form a raft on the water’s surface. e) The ants from the previous panels after being removed from the buckets and placed into an open Tupperware bin.

We have tried combining colonies that were harvested from mounds less than 5 feet apart in the wild, with the hope that we could create larger colonies that would be more useful in our experiments. To do this we have tried spooning two colonies from flooded buckets into one dry Tupperware or into one flooded bucket. In either case, after the ants from the two colonies dry off, they begin killing each other and do not stop until one of the two colonies is dead. We have repeated these attempts 5 times without ever successfully creating a larger colony.

Once in our care, the ants are given a continuous supply of high protein baby food, in our case a
smooth purée containing chicken or turkey, and water so that they never run out. We supply water with a cotton ball that is resting in water, either in a shallow petri dish of water or as a stopper on a test tube filled with water. If water is given in an open dish without a cotton ball, individual ants get trapped by the water surface tension and die, polluting the rest of the water in the dish. The ants are also given two overturned halves of a petri dish that have been blacked out, where they naturally store their eggs, winged males, and possibly queens. We do not always manage to collect the queens of a colony because they typically reside at the lowest depths of the colony in the wild. However, we do on occasion collect one or even two queens with the wild colony. Fire ants colonies are known to sometimes be polygynous, which means one colony may contain more than one queen.

For our experiments, we only use sterile female ants. We never use queens, eggs, or males, because we have noticed that including them can dramatically change the sterile females’ behavior. This is especially true for the queens, which the sterile females try hard to shelter and protect. In some cases, in which we know we will repeatedly need very large samples from a colony, we take the time to selectively kill off all of the males.

Fire ants are not polymorphic; they do not have separate worker and soldier classes. Instead the sterile females continue to grow over the course of their lives and take on different tasks [86]. Figure 2.2a shows a PDF of the size distribution of the sterile females we have collected, calculated from a sample of slightly more than 1000 ants from 10 different colonies. The average and standard deviation of our collected sterile females’ length are \( L = (3.4 \pm 0.7) \text{mm} \). We use this length to calculate an idealized circular area that represents the space a single ant takes up without interacting with its peers, \( A_{\text{ant}} = \pi L^2 / 4 \), which is illustrated in Figure 2.2. We use this to calculate the effective area fraction, \( \phi_{\text{eff}} = \rho A_{\text{ant}} \), in this thesis, where \( \rho \) is the number density of the ants. We measure the average mass of our sterile females to be \( m_{\text{ant}} = (0.8 \pm 0.1) \text{mg} \). This mass is used to calculate the number of ants we use in an experiment when counting is not feasible, typically when \( N > 150 \) ants.

When we collect ants from the bins to use them in an experiment, it is convenient to chill the
ants in a refrigerator to temporarily reduce their metabolism to ease the transfer. We have found that chilling the ants for several hours does not harm them. However, the ants behavior changes over the course of the first few weeks of captivity, even if the ants are fed and watered consistently. This can be observed in their expansion rate in column experiments and in the fact that we have never observed spontaneous waves in a colony that has been in captivity for more than three weeks. For this reason, we typically obtain new ants about once per month in the warmer months.

Figure 2.2: a) The length distribution of the worker ants we collect. b) A single sterile female fire ant. The red line represents the measurement we take to determine the length of the ant. The black circle represents the effective area of the ant, $A_{\text{ant}}$, which we use to calculate packing fractions.

### 2.2 Ants at Low Density

In Chapter 1, we reviewed the physics of idealized self-propelled particles. In the Vicsek Model, inter-particle alignment led to collective-motion and in the active Brownian particle models without alignment, motility interactions led to phase separation. We hope to understand the behavior of groups of ants within the active matter framework, so it is critical to compare individual ants’ behavior to the known active matter particle models. Ants are clearly self propelled in 2D; they burn chemical energy to power their movement relative to a surface, but it is not obvious which behaviors they may share with the idealized models, if any. Ants are social insects with complicated interactions that may remove them completely from the ideal cases.

In this chapter, we will try to quantify the ants’ social behavior by comparing their interactions to equilibrium particles and to active Brownian particles. Then we will begin to look at denser
systems of ants to search for and quantify any unique behaviors.

2.2.1 Tracking

We start by tracking small numbers of ants in 2D cells with a height of \((1.6 \pm 0.1) \text{ mm}\), which is slightly larger than the height of a large fire ant, and a diameter of \(D = (9.00 \pm 0.02) \text{ cm}\). This effectively restrains the ants to two dimensions, making it easier to observe their behavior. The cells are made by cutting a circle out of a thin sheet of acrylic and then sandwiching the acrylic between two pieces of glass. We place the acrylic on one sheet of glass, load the circular cell with a number, \(N\), of chilled ants, and then secure the second sheet of glass on top, confining the ants. Then we place these cells flat, backlight them through a sheet of light-diffuser, and image them from above, as shown in Figure 2.3a. This results in images like the one seen in Figure 2.3b, which we collect at \(3.75\text{ fps}\) and then process to find the ants’ center of mass positions, orientations, and velocities as a function of time.

Processing begins by subtracting the background from each image. The images are then thresholded to locate pixels that are part of an ant, as shown in Figure 2.3c, which allows for calculating the area, \(A\), orientation relative to a fixed frame of reference, \(\theta\), and centroid, \(\vec{r}\), of each of the connected groups of pixels that belong to an ant. We choose to calculate orientation as the orientation of the major axis of an ellipse with the same second moments of the area as our group of pixels that corresponds to the ant. Because this calculation cannot distinguish between the head and abdomen of an ant, we consider measurements of \(\theta\) and \(\theta + \pi\) equivalent.
Figure 2.3: a) An image of the empty experimental set up we use to track ants in 2D. We place the circular cell flat and light it from below through a sheet of light-diffuser. b) An image of $N = 40$ ants in the cell taken from above. c) The same image after background reduction and thresholding. Connected regions of pixels are taken to be a single ant. d) The tracked positions of the ants. The short black bars represent the center of mass positions and orientations of the ants shown in the previous images, while the colored dots show the trajectories of each of the ants over the last 20 frames.

We use this data to track the positions of individual ants by connecting the data between frames, as seen in Algorithm 1. Starting at the beginning of the image sequence, information about individual ants is sorted and assigned to tracks frame by frame by minimizing the total “cost” of adding the ants in the new image to the existing tracks. Our defined cost compares each ant’s position, orientation, and area in the new frame to the positions, orientations, and areas of the ants in the previous frame:

$$c_{ij} = C_r |\Delta \vec{r}_{ij}| + C_\theta |\Delta \theta_{ij}| + C_A \Delta \sqrt{A_{ij}}$$

where $i$ is the label of the ant most recently added to a track and $j$ is a label of one of the ants in the new frame. The constants $C_r$, $C_\theta$, and $C_A$ are set so that, on average, the distance between ants accounts for 90% of $c_{ij}$. This means that the area and orientation of the ants only come into play
when $|\Delta \vec{r}_{ij}|$ is very small, i.e. the ants are close together.

**Algorithm 1** Ant Tracking

1: **procedure** TRACK($\vec{r}_i$ positions, $\theta_i$ orientations, $A_i$ areas)
2: Assign each of the $N$ ants in the first image to different tracks, labeled $i = 1 : N$
3: for each frame from the second onward do
4: Label the ants in the new image $j = 1 : N$
5: Compare the new image’s ants to the most recent ants in each of the tracks
6: $c_{ij} = C_r |\Delta \vec{r}_{ij}| + C_\theta \Delta \theta_{ij} + C_A \Delta \sqrt{A_{ij}}$
7: for each track, $i$ do
8: new($i$) = $j'$, the ant that minimizes $c_{i,j}$
9: while An ant, $j^*$, is assigned to more than one track do
10: Find the track, $i'$, that minimizes $c_{i,j^*}$
11: $c_{i \neq i', j^*} = \infty$
12: for each track that is no longer assigned $j^*$ do
13: new($i$) = $j'$, the ant that minimizes $c_{i,j}$
14: Each track, $i$ gains the ant listed in new($i$) for this frame

Searching for the global minimum of the cost to connect all of the ants in the two frames would require us to check each of the $N!$ possible pairings of data from the old frame to data from the new frame. For $N = 2$, the algorithm would need only to calculate the two possible total costs

$$C_1 = c_{11} + c_{22}$$

$$C_2 = c_{12} + c_{21}$$

and check to see which is the global minimum cost.

$$C_0 = \min(C_1, C_2) \tag{2.1}$$

This is trivial for $N = 2$, but as $N$ increases, the runtime complexity of global minimization becomes prohibitive.

Instead, we choose to begin minimizing the cost in a “greedy” manner, one track at a time, as seen in Line 7. Each track searches for the new data that minimizes its contribution to the total cost.
of connecting the two frames. For a track \( i' \), our algorithm finds \( j' \) such that

\[ c_{i'j'} = \min_j(c_{i'j}) \]

The effect of this choice is clear when comparing the total cost of the greedy connections to the true global minimum cost in Equation 2.1. Now the assumed minimum cost is

\[ C_0 = \min(c_{11}, c_{12}) + \min(c_{21}, c_{22}) \tag{2.2} \]

Hypothetically, this can reduce the runtime complexity of the algorithm to runtime complexity \( O(N) \), instead of the \( O(N!) \approx O((N/e)^N\sqrt{N}) \) that would result from precisely minimizing the global. Runtime complexity measures the dependence of the runtime of an algorithm on a given variable, so an algorithm with \( O(N!) \) would take \( 10!/2! \approx 2 \times 10^6 \) times longer to run for \( N = 10 \) than for \( N = 2 \), while it would only take \( 10/2 = 5 \) times as long if the algorithm were order \( O(N) \) [87].

Realistically, our greedy algorithm results in some conflicts in which two different tracks select the same new set of data, \( j' = j^* \), to minimize their “local” contribution to the cost. For example, in Equation 2.2, there exists the possibility that \( c_{12} < c_{11} \) and \( c_{22} < c_{21} \), so there is a conflict over the data \( j^* = 2 \).

We solve these conflicts in Line 9. In the case of competition, in which one ant in the new frame is the most desirable addition for each of two tracks, the track with the lowest cost for that ant gets to keep it. The second track’s cost of adding that ant is set to infinity and that track searches again for the ant with the minimal cost associated to it. This process is repeated until there are no remaining conflicts. As a result, Algorithm 1 has a maximum but exceedingly unlikely complexity of \( O(N!) \). In practice, it runs close to \( O(N^2) \).

Successfully tracking the ants allows us to calculate their instantaneous velocity, so that after tracking, we know the position, orientation, and velocity of each ant for the duration of the experiment, as illustrated in Figure 2.3d. The statistics of this data allow us to quantify the behavior and
interactions of the ants.

### 2.2.2 Speed Distributions

We first turn our attention to the probability distribution for the ant’s propulsion speed. This is shown, for varying numbers of ants, in Figure 2.4a. In the idealized Vicsek Model of an active system, each of the particles self propels at a constant speed. On the contrary, an individual ant spends much of its time creeping along slowly, at only a few \( mm/s \), but also rarely chooses to dash at more than a centimeter per second. The fastest instantaneous speed we have ever recorded for an ant is a quick \( 3.2 \, cm/s \). The wide range in possible speeds could lead to interesting emergent physics, but perhaps the most important aspect of a lone ant’s speed distribution is that the ant spends \( 30\% \) of its time sitting motionless (see the arrow in Figure 2.4a).

Adding even one more ant to the cell leads to a marked change in behavior. In general, the ants’ walking speed distribution remains mostly unchanged, but the ants now spend about \( 50\% \) of their time motionless. Adding more ants to the cell after the second ant has little effect on the ants’ speed distributions. In fact, the speed distribution for \( N = 2 \) ants in the cell is more similar to the distribution for \( N = 100 \) ants in the cell than to a single ant in the cell. This highlights the social nature of the ants. A single neighbor in a large cell can significantly change an ant’s behavior.

![Figure 2.4](image)

**Figure 2.4:** a) The probability distribution of the ants’ speeds in the \( D = 9 \, cm \) cells for varying numbers of ants. b) The probability distribution for the length of time an ant remains stationary before moving again when it is alone in a cell. The black line is a power law with a critical exponent of \( 2.8 \). c) The probability distribution for the length of time an ant remains stationary for \( N > 1 \). The black line is a power law with an exponent of \( 2.1 \).
We next turn to the dynamics with which the ants switch between actively walking and remaining motionless, searching for a timescale. This is done by following an individual ant’s track and searching for periods of time when the ant is motionless, which we define as the ant maintaining instantaneous velocities of less than 1 mm/s. Changing this threshold a little bit either way does not change the results qualitatively. Each time the ant deactivates, we measure the amount of time it remains motionless until it starts moving again.

Remarkably, the statistics of the deactivating-reactivating times demonstrate power-law distributions, indicating a lack of a characteristic timescale. Figure 2.4b shows the probability of $N = 1$ ant in a cell remaining motionless for a given amount of time. The heavy tail of the power law behavior indicates that the ant can remain stationary for times much longer than the average. We can see again that the addition of one more ant to the cell changes an ant’s behavior in Figure 2.4c. In this case the power law behavior persists, but the critical exponent changes from 2.8 to 2.1, showing that the ants now spend much longer times stationary once they stop.

2.2.3 Motility Induced Attraction

The source of this behavior change is emphasized in Figure 2.5a, which shows the average speed of an ant a given distance away from another ant. To make this calculation, we start with one ant, in one frame, and calculate the center-to-center distances, $r$, from this ant to each of the other ants in the cell. Then, for each distance we measure, we add a copy of the ant’s speed to a 1 mm wide bin that includes that distance so that there are $(N - 1)$ copies of this ant’s speed in various bins. In effect we are binning the measurements of the speed by $r$ into 1 mm bins, but each measurement of $|v|$ is placed into multiple bins. We repeat this process for every ant in the frame, then for every frame in an experiment, and then for several experiments with the same number of confined ants. Finally, we take the average speed from the measurements in each bin:

$$\langle v(r) \rangle = \frac{1}{n(r)} \sum_i |v_i|$$
where the index runs over every instance in a bin and \( n(r) \) is the number of counts in the bin.

One way to understand Figure 2.5a is to consider a reference ant, Ant X, then consider that there is another ant, ant Y, a distance \( r \) away. Figure 2.5a shows the average speed at which we are likely to measure Ant Y moving. Notice that \( \langle v(r) \rangle \) saturates to different values for different \( N \) because increasing \( N \) increases the odds that ant Y is interacting with another ant besides ant X.

Figure 2.5: a) The average speed of an ant as a function of its distance from a second ant. b) The effective motility induced potential calculated from the average speed. c) The average alignment of the ants as a function of distance. Ants are only aligned as very short distances (when they are likely to be stationary) or in the rare circumstance that they are on exactly opposite sides of the circular cell.

Clearly, the ants slow down on average when they approach a neighbor. At long distances the average speed of the ants decreases with increasing \( N \) because higher \( N \) increases the probability that Ant Y may interact with another ant, independent of Ant Y’s center-to-center distance to Ant X.

We know from Section 1.5.1 that a spatial dependence in motility can be modeled to leading order as an effective potential:

\[
U_v = k_B T \ln(v(r))
\]

Here we have no measure of the ants’ temperature, but we can still look at the shape of the effective potential, \( U_{\text{eff},v} = \ln(v(r)) \), which is useful in determining the relative strength of the attraction or repulsion between the ants. These potentials are shown in Figure 2.5b. Evidently, the ants have an effective attractive potential, since \( U_{\text{eff},v} \) increases when \( r \) increases, so that we expect to find the ants close together statistically more often than we find them apart.

In keeping with a motility induced attraction, the ants at these densities show no sign of
collective-motion. Figure 2.5c shows the average alignment of the ants as a function of the distance between them. For this measurement, the pairs of ants are binned in the same manner as they are when we measure $v(r)$. We then calculate the averaged quantity:

$$\sum_i \frac{2 \cos^2(\Delta \theta_i) - 1}{n(r)}$$

which spans from 0 for completely disordered rods to 1 for perfectly aligned rods.

Note the slight decrease in $\langle 2 \cos^2(\Delta \theta) - 1 \rangle$ for intermediate distances and the steep increase near $r = 9 \text{ cm}$. Both of these effects are a result of the circular cell that we have confined the ants to. The only way for two ants to have a center-to-center distance approaching $r = 9 \text{ cm}$ is for them to find themselves on exactly opposite sides of the cell, oriented such that each of their centers are close to the wall, which makes them parallel.

Aside from the confinement effect, the ants only align at very low distances. However, comparison with Figure 2.5a shows that for these same low distances, the ants, on average are moving very slowly. Because collective-motion requires simultaneous self-propulsion and alignment, there is no collective-motion in the ants to disturb their motility induced attraction.

2.2.4 Potential of Mean Force

We can independently calculate the effective potential of mean force by measuring the pair correlation function of the ants. The pair correlation function in 2D is given by:

$$g(r) = \frac{1}{\rho} \frac{P(r)}{2\pi r}$$

where $P(r)$ is the probability of finding a particle a distance $r$ away from another particle and $\rho$ is the average number density of the system. The circumference of a circle in the denominator serves to normalize $P(r)$ by the perimeter over which a particle could possibly be found at a given distance. Whereas $P(r) \to \infty$ when $r \to \infty$, $g(r) \to 1$ as $r \to \infty$.

Our confinement geometry requires us to use a normalization factor different from $2\pi r$, which
we call the available arc length, \( s \), represented by the red arc in Figure 2.6a. For two ants, labeled X and Y, with Ant X a distance \( R \) from the center of the cell, the available arc length for Ant Y to be found a distance \( r \) away from Ant X is

\[
s(X, Y) = \begin{cases} 
2r(\pi - \arccos \left(\frac{(D/2)^2 - R^2 - r^2}{2rR}\right)) & \text{if } R + r > \frac{D}{2} \\
2\pi r & \text{if } R + r \leq \frac{D}{2}
\end{cases}
\]

The available arc length, \( s(X, Y) \), is not symmetric because \( A \) depends only on the position of Ant X, so it is calculated twice for each pair of ants. From the measured positions of the ants, we calculate the center-to-center distances from each ant to each other ant and the available arc lengths for each of those distances every frame.

Once again we sort the measured center-to-center distances, with their available arc lengths, into bins. This time the bins are only 0.14 mm, but we have enough data that there are at least \( 10^4 \) counts in each bin up to \( r = 7 \) cm for each measured number of ants, \( N \), for which we report data. The bins are smaller this time because there is less noise in the data. We can think of this clearly by considering that the probability of an ant being in a given bin is a property of the entire data set, but \( |v(r)| \) is a property only of the given bin.

Next, we calculate a weighted measure of the counts for each bin,

\[
w(r) = \sum_i s_i^{-1}
\]

where the summation runs over all of the measured counts in the bin. This calculation of \( w(r) \) normalizes each count individually by its associated available arc length. For example consider a pair of ants found at a distance apart in which only \( \pi \) radians are possible at that distance apart in the cell. Weighting that count by \( 1/s \) is equivalent to adding a phantom count to make up for the fact that this count was made despite it being half as statistically likely.
In a system without boundary concerns, \( w(r) \) would simplify to

\[
w(r) = \sum_c \frac{1}{2\pi r} = \frac{N}{2\pi r}
\]

Observations of \( w(r) / \sum_r w(r) \), where the summation runs over all bins, should approach \( P(r) / 2\pi r \) by the law of large numbers. Finally, we normalize \( w(r) \) by its average value between 4 cm and 6 cm to insure that \( g(r) \to 1 \) appropriately at long distances.

\[
g(r) = \frac{w(r)}{\langle w(r) \rangle_{4,6}}
\]

This results in our measurement of the pair correlation functions, as seen in Figure 2.6b.

![Figure 2.6](image)

**Figure 2.6:** a) An illustration of the variables used to calculate the pair correlation of the ants. b) The measured pair correlation of the ants. The single peak becomes less prominent as the density is increased. c) The effective potential of mean force for the ants, as calculated from their pair correlations. d) A comparison of our two different measures of an effective potential for \( N = 3 \) ants in the cell.

Notice that \( g(r) \) has a strong peak for \( N = 2 \) ants and that this peak becomes less prominent as \( N \) is increased. This shows that the ants have a strong preference to be a given distance apart, but that this preference becomes less important as more ants are added to the cell. The pair correlation functions can be best understood by calculating an effective potential of mean force, which is a tool used in equilibrium many-body systems, such as colloidal suspensions, which have been used to model classical atomic behavior [88]. For thermal systems, this potential is

\[
U_g = -k_B T \ln(g(r))
\]
Because we lack a measure of $k_B T$, but in analogy to an equilibrium system, we calculate

$$U_{\text{eff},g} = -\ln (g(r)).$$

Our calculated effective potentials can be seen in Figure 2.6c, which show potential wells that become shallower as $N$ is increased. A pair of ants with a separation of less than $r' \approx 3.6\text{mm}$ experience an effective repulsion and pairs of ants with a separation distance of greater than $r'$ experience an effective attraction.

The attractive regions of the effective potentials are motility induced, as shown by the agreement in the shapes of $U_{\text{eff},v}(r)$ and the attractive regime of $U_{\text{eff},g}(r)$ over this region in Figure 2.7a-c. The similarity in shape provides strong evidence that the effective potentials we calculate are physical. Comparing Figure 2.5b and Figure 2.6c helps us understand why the wells in $U_{\text{eff},g}(r)$ tend to become more shallow as the number of ants in the cell is increased. Remember that we normalized $g(r)$ so that it approached 1 at long distances, so $U_{\text{eff},g}(r)$ is the effective potential at $r$ relative to the background. Figure 2.5b shows that the effective potential of the background decreases with increasing $N$. This causes the relative difference in potential between $r'$ and the background to decrease. Physically, the decrease in magnitude of the effective forces is a result of competing attractions as one ant may feel effective forces that pull it in different directions towards several of its nearest neighbors.

Figure 2.7: a-c) A comparison of the shapes of $U_{\text{eff},g}$ and $U_{\text{eff},v}$ for experiments using $N = 3$, 5, and 40 ants, respectively. Notice that the attractive part of the potentials have the same shape.
The repulsive part $U_{eff,g}$ must be unrelated to motility because it is completely absent in $U_{eff,v}$. Instead, we believe it is a mixture between the non-isotropic excluded-area of ants and a behavioral repulsion that keeps the ants at a preferred distance from each other, in which ants are close enough together to physically communicate but far enough apart that they don’t restrict each others’ motion.

2.3 Clustering

Because of their tendency to slow down and stop for some time when they approach each other, ants confined at low density naturally form clusters, such as those highlighted ants in Figure 2.8c. These clusters are easily visible to the eye during experiments, and they continually form and disperse over the experiment’s duration. To characterize the clusters, we first filter out any ants that are moving in each frame with $v_i(t) > 1 \text{ mm/s}$ so that we are left with only “stationary” ants in each frame. For example, only stationary ants are highlighted in Figure 2.8a, demonstrating that stationary ants are likely to be found in a small cluster with other stationary ants.
Figure 2.8: a) A frame during an experiment with $N = 40$ ants, in which the stationary ants are highlighted in red and marked with their label from the tracking algorithm. b) A dendrogram showing which ants are the most closely clustered. Dividing the ants into clusters is done by checking which ants are closer to a cluster than $r_{\text{thresh}}$. The labels in this figure are the same as in the previous one. b) The result of the clustering algorithm. Three clusters were found that contained more than $N_c = 1$ ant. d-e) The same process for an image from an experiment with $N = 100$ ants, showing detection of a cluster with $N_c = 22$. The labels have been left off for clarity.

Then, in each frame, we sort the stationary ants into individual clusters by creating a dendrogram, which is a common chart used in clustering analysis [87], from the ants’ center-to-center-distances and choosing a cutoff distance that corresponds to twice the length of an ant, as shown in Figure 2.8b. The $x$-axis shows the labels associated to each stationary ant. Each ant starts out as a branch on the dendrogram, and the branch continues vertically until it reaches a $y$ value that corresponds to the distance between the starting ant and its nearest neighbor, at which point the branch merges with the branch that includes this neighbor. This merged branch, which now represents a cluster, continues vertically until a $y$ value that corresponds to the minimum distance between any ant in this cluster and another ant outside of the cluster, and the branches merge again. This continues until all of the ants are included in one branch (trunk) at the top.
Segregating the ants into individual clusters is then a matter of picking a $r_{\text{thresh}}$ and checking which ants are included in each branch. The result of choosing $r_{\text{thresh}} = 7 \, \text{mm}$ is shown in Figure 2.8c, in which the ants are color coded according to which cluster they belong to. As is apparent in Figure 2.8b, most clusters are insensitive to slight variations of $r_{\text{thresh}}$. Figures 2.8d-f show the same process for an experiment with $N = 100$ ants, but the labels hidden for clarity. Notice that about half of the ants are stationary in this frame and that the largest cluster, highlighted in cyan, is $N_c = 22$.

Once we have separated the ants into clusters, we can then analyze the clusters’ internal structure and size distributions.

### 2.3.1 Structure

To quantify the cluster structure, we first determine its radius of gyration, $R_g$, which is related to the moment of inertia, $I$, around an axis perpendicular to the cluster and passing through the cluster’s center of mass:

$$I = \sum m_i(\vec{r}_i - \vec{r}_{CM})^2 = m_{\text{tot}}R_g^2,$$

where $\vec{r}_i$ is the position of the $i$th particle and $\vec{r}_{CM}$ is the position of the center of mass of the object. This means that an arbitrary object rotates around its center of mass as if all of its mass was concentrated at point located a distance $R_g$ from the axis of rotation.

For our calculation of $R_g$ for each cluster, we treat the masses of the ants as equal and treat each ant as a point particle so that

$$R_g^2 = \frac{1}{N_c} \sum_i (\vec{r}_i - \vec{r}_{CM})^2$$

where the index runs over all $N_c$ of the ants in a cluster.

If our clusters are fractal, because we are in 2D, we expect $N_c$ to scale with $R_g$ as $N_c \propto R_g^\delta$, where $\delta \leq 2$ is the fractal dimension. The case of $\delta = 2$ corresponds to space-filling clusters; for example, the mass a uniform disk is $m = \rho \pi R^2 = \rho \pi R_g^2/2$, where $\rho$ is the density. For our
clusters, \( N_c \propto m \), so we expect space-filling clusters to scale with \( \delta = 2 \).

On the other hand, clusters formed by Diffusion-Limited-Aggregation (DLA), in which particles approach the cluster in random walks and adhere to the cluster as soon as they touch it, famously have a fractal dimension significantly less than the dimension for space-filling. In 2D, the fractal dimension of DLA clusters is 1.71 [90]. A low fractal dimension like this can lead to gelation when cluster sizes become comparable to the system size because \( R_g \propto N^{1/1.71} \) and the number of particles in the system is \( N_C \propto \rho L^2 \), where \( L \) is the system size. If ants have a low fractal dimension, like DLA, then we might expect to observe a gelation-type transition, in which clusters span the system and restrict the motion of free ants.

Figure 2.9 shows a log-log plot of our average measured radii of gyration for clusters of stationary ants in our experiments with various \( N \), up to \( N = 100 \) ants. This is the highest number of ants we have used in an experiment with confidence that our image processing and tracking allow us to successfully find the positions of stationary ants. It appears that small clusters scale with \( \delta = 1 \), like a line. We believe this behavior is due to the tendency of these clusters to form near the walls. However, when there are more stationary ants in a cluster (\( N \geq 15 \) ants), our clusters seem to fill space with \( \delta \approx 2 \).

We note that we don’t have large enough clusters to determine if the ant clusters scale with a fractal dimension consistent with an aggregation model like the Vold-Sutherland model. The Vold-
Sutherland model parallels the DLA model except that particles approach the cluster ballistically from random directions [91]. We speculate that this may be an attractive model for clusters of active particles with strong adhesion and it results in clusters that are nearly space filling, with \( \delta = 1.9 \) [92].

Nevertheless, space-filling clusters are consistent with motility-induced attraction. Recall that the runaway process of MIPS results in clusters that are maximally dense and therefore space-filling.

### 2.3.2 Cluster Size Distributions

We also characterize the size distribution of our clusters, \( P(N_c) \), by counting the number of clusters we measure of each size, \( N_c \), and then dividing by the total number of observed clusters over the course of several experiments. As in other studies of dynamic clustering in active systems [93–96], we find that the distributions scale as \( P(N) \propto N_c^{-\gamma} \) for small clusters and \( P(N) \propto \exp(-N_c/N^*) \) for large clusters, as shown in Figures 2.10a-b. However, the apparent power law behavior seems to persist for larger clusters for our ant clusters than it does in other active systems, perhaps because of the power law distribution for the amount of time an ant spends stationary.

![Figure 2.10:](image)

Figure 2.10: a) Log-log plot of observed cluster size distributions. For \( N = 100 \) ants, there is apparent power law behavior at small cluster sizes. b) Log-linear plot of observed cluster size distributions. For \( N = 100 \) ants, there is an exponential tail for large cluster sizes.

Following the theory by Ginot et al. in Reference [93], we explain the exponential behavior for the probability distribution for large clusters by considering a detailed balance between cluster
sizes after assuming that particles only aggregate or fragment by one ant at a time. The rate of change of the probability of a given cluster size is related to the aggregation and fragmentation rates of the cluster and of clusters one ant smaller or larger:

\[
\frac{d}{dt} P(N_c) = P(N_c - 1) a_{N_c-1} - P(N_c)(a_{N_c} + f_{N_c}) + P(N_c + 1) f_{N_c+1}
\]

where \(a_i\) and \(f_i\) are the rates at which a single ant joins or leaves a cluster of size \(i\), respectively. If we further use that the system seems to be in a steady state, \(\frac{d}{dt} P(N_c) = 0\), because the clusters don’t seem to continue growing in time, we can postulate a detailed balance:

\[
P(N_c) a_{N_c} = P(N_c + 1) f_{N_c+1}
\]

If after a certain cluster size, \(N_c^*\), the two rates scale the same way, \(f_{i+1}/a_i = W\) with \(W < 1\) a constant, then we are left with a recursion relationship:

\[
P(N_c) = W P(N_c - 1)
\]

\[
P(N_c) = W^2 P(N_c - 2)
\]

\[
P(N_c) = W^{N_c-N_c^*} P(N^*)
\]

Because we are in a steady state, \(N_c^*\) and \(P(N_c^*)\) are constant, which results in:

\[
P(N_c) = \left(\frac{1}{W}\right)^{N_c^*} P(N^*) \left(\frac{1}{W}\right)^{-N_c} = const * e^{-N_c}
\]

and we recover the exponential tail in Figure 2.10b.

The assumption that \(f\) and \(a\) scale the same way is reasonable for active aggregations after a certain cluster size. We have seen that the clusters tend to fill space. If only the ants on the outside edge of a cluster are free to leave the cluster, then we might expect \(f_{N_c} \propto N_c^{0.5}\), the same as the perimeter of the cluster. Likewise, if we the ants move ballistically on the length scales similar to
our cluster sizes, the probability of an ant encountering a cluster so that it is given an opportunity to join it should be proportional to the cross-sectional length of a cluster. Because our clusters fill space, this cross-sectional length is also proportional to $N_c^{0.5}$.

The apparent power law for smaller cluster sizes is not as well understood, but it must be related to cluster formation. However, we note that for small clusters, growth is far from space-filling, so it is likely that the two rates $f_{N_c}$ and $a_{N_c}$ are not proportional for small $N_c$.

2.4 Ants at High Density

2.4.1 Density Autocorrelations

To measure the behavior of ants at very high effective packing fractions, we now confine them to cells with the same height, $(1.6 \pm 0.1) \text{ mm}$, but a smaller diameter, $D = (4.50 \pm 0.02) \text{ cm}$, as shown in Figure 2.12a. As we increase the density of ants in our cells, we lose the ability to track individual ants, but consistent with Motility-Induced Phase Separation, we begin to see dynamical heterogeneities as the ants micro-phase separate into dense clusters of stationary ants and a background of active ants.

The two phases can be clearly seen in Figure 2.11, which shows a false color image from an experiment with $\phi_{eff} = 3.6$. Ants shown in black do not move for the next 30 s and ants shown in magenta move within the next 30 s. This shows that there are dense clusters are made up of stationary ants and that these clusters evolve on a slower timescale than the moving ants that make up the background.
Now we further explore the analogy to phase separation. In traditional phase coexistence, changing the volume while keeping temperature constant changes the proportion of the two phases but doesn’t change the nature of either phase. For example in a liquid gas system, this results in a vapor pressure for the gas in equilibrium with a liquid that depends only on temperature. Decreasing the volume of the system results in more of the gas condensing into a liquid but does not change the pressure of the liquid or the gas. If the analogy of a phase coexistence continues to hold for the ants, then we might expect the dynamics of the two phases to remain essentially unchanged with global density, while the proportion of the two phases changes.
Figure 2.12: a) An image of ants in a $D = 4.5$ cm cell with a global packing fraction $\phi_{\text{eff}} = 2.5$. b) The coarse grained optical density corresponding to the previous image. Yellower squares correspond to higher optical density. c) The density autocorrelation function for $\phi_{\text{eff}} = 2.5$, illustrating the two inverse timescales. $\tau_1$ in blue corresponds to the walking movement of the ants. $\tau_2$ in red corresponds to formation and break up of clusters. d) The two timescales as a function of the global packing fraction. Both timescales generally increase as the density is increased, indicating that the ants’ dynamics slow down.

To measure the dynamics of the phases at these high densities, we use local optical intensity as a stand-in for local number density. We coarse grain the cell into square bins with a side-length of 2 mm, which roughly corresponds to half the length of an ant, and measure the average optical intensity absorbed by the ants in each bin, $I_a(x, y)$, by comparing an image to an image of an empty cell. Figure 2.12b shows the result of coarse graining the absorbed intensity for the image in Figure 2.12a. The absorbed intensity, $I_a$, has been scaled so that the 0 corresponds to no intensity absorbed by the ants and 1 corresponds to a perfectly dark region, in which all of the light has been blocked by ants. Then we measure the time-autocorrelation of the fluctuations in $I_a(x, y, t)$, $C(\Delta t)$:
\[ \Delta I_a(x, y, t) = I_a(x, y, t) - \langle I_a(x, y, t) \rangle_t \]

\[ C(x, y, \Delta t) = \frac{\langle \Delta I_a(x, y, t + \Delta t) \ast \Delta I_a(x, y, t) \rangle_t}{\sqrt{\langle \Delta I_a(x, y, t)^2 \rangle_t}} \]

\[ C(\Delta t) = \langle C(x, y, \Delta t) \rangle_{x,y} \]

Figure 2.12c shows an example of the \( C(\Delta t) \) for a single three hour experiment, which shows that \( C(\Delta t) \) decays with two characteristic timescales, which we call \( \tau_1 \) and \( \tau_2 \). At short \( \Delta t \), \( C(\Delta t) \) relaxes with a fast timescale, \( \tau_1 \), that corresponds to the movement speed of the ants in the actively moving phase. Observation shows that \( 3 \text{ s} \leq \tau_1 \leq 10 \text{ s} \) is related to the moving phase. For example, Figures 2.13a-b show composite images where dark pixels are ants that remain stationary for twice the measured \( \tau_1 \) for that density, and magenta pixels show ants that move within the next \( 2\tau_1 = 16.8 \text{ s} \). Notice that the active phase has fluctuated within that time, but the clusters have remained intact—almost unchanged.

At longer times, \( C(\Delta t) \) relaxes on a very slow timescale, \( \tau_2 \), that corresponds to the formation and break up of the clusters of stationary ants, \( 2.5 \text{ min} \leq \tau_2 \leq 10 \text{ min} \). For example, Figure 2.13b was created using frames that were captured \( 2\tau_2 = 12.3 \text{ min} \) after the images in Figure 2.13a. Notice that there are once again dense clusters of stationary ants, but they are not in the same locations as the clusters in Figure 2.12b. This exemplifies the relation between \( \tau_2 \) and the cluster dynamics.
We find that as the global effective packing is increased, both of these timescales continue to grow, as can be seen in Figure 2.12d. Apparently, the ants in the actively moving phase slow down due to the increased traffic, and the clusters take longer and longer to break up. Observation shows that, at moderately high densities, the clusters tend to slowly grow or dissolve from the edges due to interactions between the stationary ants and the actively moving ants, in keeping with our assumption in the previous section. This slow down in both timescales is not what we might have expected for typical micro-phase separation and suggests that simplified model of MIPS may not hold for our ants at high densities. We can speculate as to the origin of this difference by referring to the speed distribution for the ants at low density in Figure 2.4a, where we saw that increasing the global density did not slow down the ants. It seems that there is always some proportion of ants that are inclined to move, regardless of how crowded their environment is. Adding a rule of this type to traditional MIPS may make for a more successful model with which to predict the behavior of crowds of social creatures.

2.4.2 Activity Cycles

Increasing the density of the ants eventually leads to a spontaneous and brief collective behavior that completely dissolves the stationary clusters and further animates the already active ants. We
call these events “activity cycles,” as first reported by Tennenbaum et al. [13, 97]. Figures 2.14a-f show ants during an experiment with the lowest density for which we have ever recorded an activity cycle, $\phi_{\text{eff}} = 3.6$. Figures 2.14a-c are taken about 4.4 min apart and show the ants during their typical phase separated behavior, for which we saw the two timescales in the previous section. The slow timescale is about $\tau_2 = 9.3$ min at this density, and the clusters are mostly unchanged between the three images.

In contrast, Figures 2.14d-f are taken only 8.3 s apart while the ants are in an activity cycle and all the ants in the cell are actively moving. The clusters are completely broken up by the activity of their moving neighbors.
Figure 2.14: a-c) Three images, each taken $4.4 \text{ min}$ apart $\sim 119 \text{ min}$ into the experiment, outside of an activity cycle. The global packing fraction $\phi = 3.6$ d-f) Three images, each taken $8.3 \text{ s}$ apart $\sim 150 \text{ min}$ into the experiment, during an activity cycle. g) The optical area fraction (gray) and moving ratio (red) as function of time with a packing fraction of $\phi = 3.6$. The strong peak in the moving ratio corresponds to an activity cycle. The black vertical lines correspond to the image sequences shown above. h) The presence (blue) or lack (red) of activity cycles as a function of the packing fraction in the experiment. Some points are offset vertically for clarity.

We can detect the cycle empirically by measuring the ratio of changing pixels over the course of an experiment. We threshold each of the images to separate the ants from the background; $I_{\text{thresh}}(x, y, t) = 1$ if that pixel corresponds to part of an ant and $I_{\text{thresh}}(x, y, t) = 0$ if the pixel is not
part of an ant. This allows calculation of the optical area fraction:

\[ \phi_{\text{opt}}(t) = \frac{\sum_{x,y} I_{\text{thresh}}(x, y, t)}{n_{\text{pix}}}, \]

where \( n_{\text{pix}} \) is the number of pixels that make up the confining cell. We also calculate the proportion of the pixels that change between frames, which we relate to the number of moving ants:

\[ M(t) = \frac{\sum_{x,y} |I_{\text{thresh}}(x, y, t) - I_{\text{thresh}}(x, y, t + \Delta t)|}{2\phi_{\text{opt}}(t)n_{\text{pix}}} \]

For this calculation, we choose to use \( \Delta t = 5 \) s, but the qualitative results do not change for different choices of \( \Delta t \), even up to \( \Delta t = 26.7 \) s. The only change is an approximately constant shift in measured values of \( M(t) \) because active ants have only moved different distances and the stationary ants remain stationary over reasonable choices of \( \Delta t \).

Figure 2.14g shows \( \phi_{\text{opt}}(t) \) and \( M(t) \) over the course of the experiment with \( \phi_{\text{eff}} = 3.6 \). Notice that \( M(t) \) starts off high at the beginning of the experiment but drops over the first 15 min. This is typical behavior because the ants are always completely active just after being handled and loaded into the cell. This same effect was also seen in prior rheological studies [13]. Afterwards, the ratio of moving pixels fluctuates but remains low. The first vertical line at about \( t = 120 \) min represents the moment in the experiment in which Figure 2.14b was captured, while \( M(t) \) is low. The second vertical line at \( t = 150 \) min corresponds to the instant of Figure 2.14e and marks the presence of an activity cycle. Notice the magnitude at the peak in \( M(t) \) from the activity cycle is roughly equal to its value at the beginning of the experiment, which shows that all of the ants in the cell are once again active.

We take the presence of one of these peaks over the course of 3 h experiments to be a quantitative indicator of an activity cycle. Figure 2.14h shows the presence (blue) or lack (red) of activity cycles in our experiments as a function of the effective packing fraction. The activity cycles have only been seen in these cells for packing fractions \( \phi_{\text{eff}} \geq 3.6 \).

Activity cycles represent a marked change in behavior because, as previously seen, increasing
density of the ants typically leads to the slowing of their dynamics.

We noticed earlier that there was no collective-motion in the ants at low density because the ants only experienced an aligning interaction when they were so close that they were likely to stop moving. This preclusion of collective-motion allowed for motility induced attraction, typical of active matter without aligning interactions, which caused the formation of clusters. During an activity cycle, the ants can find themselves both extremely close to their neighbors and in motion. This makes it likely that the ants exhibit collective-motion. This would make ants a particularly interesting active matter system: one that shows MIPS-like behavior at dilute concentrations, but intermittent collective-motion at high densities that turns on and off over time. Unfortunately, the small cells we use in this section make it hard to detect collective-motion, but we will remedy that problem in the next chapter.

2.5 Conclusions

Ants are self-propelled, but they move with variable speeds. Despite complicated social interactions, ants at low densities demonstrate one of the hallmarks seen in ideal active matter, including motility induced attraction and phase separation.

At low densities the behavior of ants in a cell can be successfully modeled, in an analogue to thermal particles, with an effective pair potential that has a minimum at \( r' = 4 \text{ mm} \). For inter-ant distances of less than \( r' \), excluded area effects lead to an effective repulsion, and we have shown that motility effects cause an effective attraction at inter-ant distances of greater than \( r' \). Despite the non-isotropic shape of the ants, there is no collective-motion at low densities because the ants slow down and stop moving when they approach each other.

The motility-induced attraction and lack of alignment instead lead to the formation of clusters of stationary ants, and we demonstrated that these clusters are space-filling. We also showed that the size distribution of these clusters was consistent with a model that assumed the clusters formed and dissolved one ant at time.

The motility effects between the ants result in phase separation into dense clusters of non-
moving ants in a background of actively moving ants, resulting in dynamical heterogeneities. Increasing the density results in a slowing down of the timescales of both phases: the ants in the active “gas” phase slow down because of the increased traffic and dense clusters of ants take longer to break apart and reform.

Finally, as the density is increased past a critical effective packing fraction, which we measure to be $\phi_{\text{eff}} = 3.6$, the phase separated state is punctuated by activity cycles. During an activity cycle all of the clusters are dissolved and all the ants in the cell are actively moving or trying to move, even though the density remains very high, which may put the ants in the collective-motion phase.
CHAPTER 3
ANT WAVES

In this chapter we will introduce our surprise observation of activity waves in dense 2D fluids of ants. We quantify different aspects of the waves and compare them to other moving disturbances in active systems, especially the traveling bands observed near the collective-motion phase transition. In order to explain the ant waves, we propose the types of interactions and behaviors necessary for the ants and test whether these behaviors are enough to generate waves in simulations and an analytical model. While we find that certain observed behaviors do lead to collective waves in our models, they fail to capture the nonlinearity we measure of the waves we measure experimentally. We finish the chapter by speculating on the source of this nonlinearity and show how our models can be changed to reproduce it.

3.1 Experiments

We now turn our attention to the collective behavior of the fire ants in 2D. Whereas the experiments from the previous chapter used on the orders of $10^2$ number of ants, the experiments in this section are all performed with $10^3 - 10^5$ ants.

3.1.1 Set Up

We confine the fire ants to the rectangular acrylic cell shown in Figure 3.1a, after slicking the walls with talcum powder to greatly reduce the ants’ ability to cling to the cell. The cell is made of strips of acrylic sandwiched between two large sheets of acrylic so that a gap with dimensions $(90.0 \pm 0.2)\, cm \times (10.0 \pm 0.2)\, cm \times (1.6 \pm 0.2)\, mm$ is left in the center. Chilled ants are loaded into the cell before the top sheet is placed and then the cell is held together with binder clips along its outer edge. Once the ants are trapped in the cell, we orient it so that its long axis is vertical and shake it vigorously so that the ants are forced to the bottom of the column, where they form the
dense mass at the base of the column shown in Figure 3.1b. In the dense mass, the ants are packed much closer together than \( r' = 3.6 \text{ mm} \), so they are effectively repulsive, which causes the initially dense mass to expand against gravity. The ants continue to expand until they reach a steady state, like the one shown in Figure 3.1c, with high density at the base of the column, gradually lower densities with increasing height, and a nearly level free surface at the top.

Figure 3.1: (a) The large cell used to confine the ants. (b) A dense mass of ants confined to the cell shortly after the ants are knocked down to the base of the cell. (c) The same mass of ants fully expanded. (d) A schematic of the experimental set up. A large fluorescent lamp behind the cell serves as a back-light. (e) The calibration curve used to calculate the density of ants in the cell from the transmission. (f) The space-time surface calculated for the same experiment as images b, c, and d of this figure. Colors closer to red indicate higher density.

To measure the density of the ants in the cell as a function of height and time, we back-light the cell with a fluorescent lamp through a large piece of paper that serves as a light diffuser and take images at 1 \( fps \) from the front, as shown in Figure 3.1d. We then average the measured intensity horizontally in rectangular regions of interest, which are highlighted in yellow in Figure 3.1d. The
first region is the 10 cm wide column of ants, which yields an average observed intensity as a function of height and time, \( I(y, t) \). The second region is 100 pixels (∼3 cm) wide immediately to the left of the outside of the cell, which we use as a background intensity as a function of height and time, \( I_0(y, t) \). A measurement of the background intensity is necessary because the intensity of light from our fluorescent lamp is not completely constant in space nor time. It lightly flickers over the course of a single experiment. From these two intensities, we calculate the transmission as \( T(y, t) = I/I_0 \), which we blur with a moving average over 2 cm in \( y \) and 2 s in \( t \) to reduce the effects of noise.

Finally, we calibrate the transmission to the number density by counting the ants contained in regions 2 cm high and 10 cm wide in a few images and comparing the counts to the corresponding measured \( T(y, t) \). Our observations are shown in Figure 3.1e. The maximum transition we measure in our cell corresponds to \( \rho(y, t) = 0 \), but still \( T(y, t) \neq 1 \), because \( I_0(y, t) \) was measured where the light was only transmitted through paper, while \( I(y, t) \) was measured where the light had to pass through the paper and two layers of acrylic coated in baby powder. On the other end of our calibration curve, for low \( T \), the ants become so densely packed that we cannot distinguish between ants to count them.

Our calibration measurements are well fit by a straight line that we then use to calculate an estimated \( \rho(y, t) \) elsewhere in the cell over the course of the whole experiment from the measured \( T(y, t) \), like the data shown in Figure 3.1f. We extrapolate our straight line to account for higher \( \rho \) at lower \( T \), even though we suspect \( T \) might no longer be a linear function of \( \rho \) at very high densities. High reported densities in this chapter should then be assumed to be relatively correct but not exact. Instead, very high reported densities are likely even higher. For example, a region with reported density \( \rho(y, t) = 23 \text{ ants/cm}^2 \) is definitely more dense than a region with a reported density \( \rho(y, t) = 20 \text{ ants/cm}^2 \), but both reported densities are likely underestimations.

Notice that, at the beginning of the example experiment shown in Figure 3.1f, the height of the free surface increases linearly while the column is expanding. This is the typical behavior of the ants at the beginning of experiments in the large cell. We believe that this constant-rate expansion
is due to the speed at which ants are able to unjam themselves from the dense mass at the bottom into which they were unnaturally forced.

At the free surface in the final steady state, the ants maintain a density of about $\rho^* = 10 \text{ ants/cm}^2$, so each ant has a specific area of $a^* = 0.1 \text{ cm}^2$. Assuming that $a^*$ is roughly a disk, we find its radius to be

$$r^* = \sqrt{\frac{a^*}{\pi}} \approx 1.8 \text{ mm},$$

so that the ants have a center-to-center distance of $2r^* \approx 3.6 \text{ mm} \approx r'$, consistent with the minimum of the ants’ effective potential (see Section 2.2.4). Below the free surface the density increases and the center-to-center distances shrink. Thus, all the ants in the cell primarily interact repulsively to support the weight of the column against gravity.

### 3.1.2 Activity Waves

Remarkably, in a fraction of the experiments, once the final height is reached, the ants spontaneously propagate density waves towards the free surface, such as the wave at about 28,000 s in the space-time plot of Figure 3.1f. In this example space-time plot, darker red colors correspond to higher density and the wave is apparent as a dark red line through the yellow and green background. The waves begin somewhere near the base of the column, persist for several minutes, and propagate until they reach the free surface at the top of the column, which may be tens of centimeters from their origin. The density waves propagate at a constant speed with a constant amplitude. This is immediately evident in Figure 3.2a; this figure shows the same space-time region highlighted in Figure 3.1f rotated and flipped to highlight the speed and amplitude.

Figure 3.2b shows real space images of the wave from Figure 3.2a with intervals of 20 s between them. Notice that the density of the ants immediately above the wave, where it has yet to pass, is relatively inhomogeneous: there are small dense clusters of stationary ants surrounded by less dense, slowly moving ants. On the other hand, the density of ants directly behind the wave is much more homogeneous. This contrast between the density in front and behind the wave was one of the signatures of the activity cycles in Section 2.4.2 and hints at a change in the ants’ behavior.
This helps explain why the waves originate lower in the column. Activity cycles are only initiated when the ants are confined at very high densities, which in the columns is only true near the base. Once started, the activity cycle is able to propagate as a wave through the lower density regions of ants above the cycle’s origin. As a result, it seems that, in addition to being a wave in density, the ant waves are also a front in activity.

Figure 3.2: (a) A three dimensional view of the space time surface for the wave highlighted in Figure 3.1f. The surface has been flipped and rotated to highlight the constant speed and amplitude of the wave. (b) Three real space images of the same wave taken 20 seconds apart. The scale bar 5 cm and the arrows have been added to draw attention to the wave.

3.1.3 Collective-Motion

Unlike particles in linear waves, such as sound waves, ants experience a net displacement as a wave passes through. Figure 3.3a shows the vertical component of three typical individual ants’ velocities as they join and then leave a wave. These are measured by taking a close up video of a wave at a frame rate of 0.5 fps and then locating a few individuals during each frame. A line segment is drawn from the tip of an ant’s abdomen to the tip of its mandibles, as shown in Figure 3.3b, and the position of each ant is calculated as the mid-point along this segment. By following the position of the ants over several minutes as the waves passes through, we calculate the ants’ instantaneous velocities by comparing their positions between frames. We take the vertical component of an ant’s velocity because that is the direction of wave propagation and average it over 15 s because the ants stop and start suddenly in their crowded environment. The resulting average vertical speed of the ants is reported as $v_{0,y}$ in Figure 3.3a.
The ants are swept into the leading edge of the wave, direct their motion in the same direction as the wave for a short time, and then drop out from the trailing edge. This is reminiscent of the simulated waves of intrusion demonstrated by Grégoire and Chaté [24] and discussed in Section 1.3.2, which prompts us to ask whether the ant waves are a case of phase coexistence between an ordered collective-motion phase and a background disordered phase, even though we have not seen evidence of long range order in the ants up to this point.

To check this we measure the alignment of the ants within a region of interest through the duration of a wave. We begin with a region sufficiently far from both the walls and the free surface, through which a wave passes. Our region of interest measures 3.2 \( cm \times 5.5 \text{cm} \) horizontally and vertically, and contains about 500 ants. For each of eight frames, spaced about 20 s apart, the alignment and position of each ant in the region of interest are determined by associating an arrow to each ant that points from the posterior tip of an ant’s abdomen to the center of their mandibles, as seen in Figure 3.4a. We use the mid-point of the line of the arrow as the position of the ant and treat the direction of the arrow as the ant’s orientation. Using this data, we can bin the ants by vertical position and examine the distributions of orientations in each bin, as seen in Figure 3.4b. The wave is harder to see by eye in Figure 3.4b because we are zoomed in on a small part of it, but the density wave is currently located in the center bin. Notice that for the polar histogram of the top bin, where the wave has not yet passed, the ants show a preference for orientations along
the vertical axis, perhaps for gravitational reasons, but low polar order. In contrast the bottom two bins show strong polar order in the direction of the wave.

Figure 3.4: (a) A close up image of ants near the horizontal center of a large cell. The white arrows point from the abdomen to the mandibles of the ant and are used to determine the ants’ positions and orientations. The scale bar is $5 \text{ mm}$. (b) A small region of interest in which the orientations of ants have been measured over the course of a wave. The region of interest has been separated into three bins, and the polar histograms to the sides show the orientational probability distribution functions for each bin. The bottom two bins show the orientational order brought on by the wave. The scale bar is $1 \text{ cm}$. (c) An illustration of images shifted into the wave frame. The red line indicates the center of the density wave, $y' = 0$, and the blue box represents the area of one of the vertical $2 \text{ cm}$ bins used to average the ants density and orientation. (d) The density (blue) and average alignment (red) in each of the bins as a function of the bin’s position relative to the center of the wave, $y' = 0$. $y' < 0$ corresponds to regions through which the wave has already passed, and $y' > 0$ corresponds to regions with ants that have not yet felt the effects of the wave.

To obtain better statistics, we take advantage of our observation that the ant waves travel with a constant speed and shift the positions of the ants into a frame of reference that moves with the wave, $y'$. We consider the average speed of the wave, calculate where the center of the wave should
lie for each frame, and translate the vertical coordinate of that frame so that the center of the wave lies at $y' = 0$. Once the frames are shifted, we treat all of the images as a single data set and bin the ants by $y'$ into ten 2 cm bins, each containing 400 ants on average. An example of this translation to the wave frame and the borders of a bin in two images is shown in Figure 3.4c.

For each of these bins, we calculate the alignment of the ants as $\langle \sin(\theta) \rangle$, where $\theta$ is orientation of an ant’s associated arrow with respect to the horizontal, and the local density as $\rho = N/A_{tot}$, where $A_{tot}$ is the total area available to the ants in that bin. We choose $\langle \sin(\theta) \rangle$ to represent the alignment of the ants, because we know the ants align to face upwards as the waves pass.

The result of binning and calculating the density and order is shown in Figure 3.4d. The region where $y' > 0$ represents regions above the wave, where the wave has not yet passed, so that the wave in Figure 3.4d can be understood to be propagating from the left towards the right of the plot. Here the density is plotted in blue and the orientational order of the ants is shown in red. Notice the strong peak in the ants’ alignment at $y' < 0$, which means that the alignment follows the density. Evidently, in addition to being a wave in density and a wave in activity, an ant wave is also a wave in alignment.

This is evidence of local collective-motion in ants. Recall that, at the low densities of Section 2.2, the ants didn’t exhibit collective-motion because they never fulfilled the requirements of being both dense enough and active enough to align. In the small cells at high density in Section 2.4, the ants sometimes fulfilled these two requirements, but they didn’t have room to exhibit clear collective-motion. Here, we see collective-motion, a hallmark of active matter, in the ant waves in large cells.

### 3.2 Simulation

Inspired by the appearance of local collective-motion in the waves, we set out to check our understanding by building a Vicsek-like simulation with a few added rules and looking for waves. Our simulation, like Vicsek et al.’s, takes place in discrete time and continuous space [17]. We initialize ants regularly spaced on a square grid with unit velocities in random directions. After initializing,
the simulation takes place over a series of time-steps in which each ant interacts with its neighbors as shown in Algorithm 2.

### Algorithm 2 Wave Simulation

1: procedure STEP( )
2: Sort(positions)
3: for each ant do
4: neighbors = FIND_NEIGHBORS
5: \( \vec{v}(t + \Delta t) = \text{ALIGN}(\text{neighbors}) \)
6: \( \vec{v}(t + \Delta t) = \text{GRAVITY}(\text{neighbors}) \)
7: \( \vec{v}(t + \Delta t) = \text{COLLIDE}(\text{neighbors}, \vec{v}(t + \Delta t)) \)
8: \( \vec{v}(t + \Delta t) = \text{CONFINE}(\text{position}, \vec{v}(t + \Delta t)) \)
9: for each ant do
10: \( \vec{r}(t + \Delta t) = \vec{r}(t + \Delta t) + \vec{v}(t + \Delta t) \Delta t \)

3.2.1 Rules

Each step, the particles are sorted into square boxes with edge-lengths set by \( R_I \), the interaction radius, so that we have a list of which ants are present in each box. Physically, we expect the interaction radius to represent the distance an ant can reach with its legs to detect motion. This sorting is a common bookkeeping step used in modern molecular dynamics simulations that makes it easier for particles to locate neighbors that they can potentially interact with [98]. It is not necessary for the physics, but it decreases the run time complexity of the simulation from order \( O(N^2) \) to order \( O(N) \) because each ant only needs to search its own box and the surrounding eight boxes to find all neighbors it can possibly interact with instead of searching the entire simulation area.

After sorting the ants, a loop allows each ant to update its velocity based on the positions and velocities of its neighbors. An ant begins by making a list of neighbors that are within a distance \( R_I \) away, which is a subset of the list of particles in its own and surrounding eight boxes. Then the ant undergoes a series of interactions in the order shown in Algorithm 2. Each type of interaction depends only on the positions and velocities of the ants from the previous time-step, so the order in which the ants update their velocities does not matter. Because each interaction can change the
ant’s velocity, \( \vec{v}(t + \Delta t) \), for this time step, interactions that occur later in the list beginning on Line 5 take precedence over interactions that are earlier in the list.

The first update to the ant’s velocity is the aligning interaction used by Vicsek et al. First, the ant checks the average orientation of all its neighbors:

\[
\langle \theta(t) \rangle_{R_I} = \arctan 2(\langle v_y(t) \rangle_{R_I}, \langle v_x(t) \rangle_{R_I}),
\]

where \( \arctan 2 \) is the two-argument arctangent that unambiguously returns the angle that would result from transforming Cartesian inputs into their equivalent in polar coordinates, and \( \langle v_x \rangle_{R_I} \) and \( \langle v_y \rangle_{R_I} \) are the average velocity of all interacting neighbors within \( R_I \) in the horizontal and vertical directions, respectively.

Then, the ant noisily aligns its own orientation for the next time step, \( \theta(t + \Delta t) \), to \( \langle \theta(t) \rangle_{R_I} \):

\[
\theta(t + \Delta t) = \langle \theta(t) \rangle_{R_I} + \Delta \theta,
\]

where \( \Delta \theta \) is a uniformly distributed angle between \(-\eta\) and \( \eta \). The updated orientation is used to update the ant’s velocity, which becomes:

\[
v_x(t + \Delta t) = v_0 \cos(\theta(t + \Delta t))
\]

\[
v_y(t + \Delta t) = v_0 \sin(\theta(t + \Delta t))
\]

The length \( R_I \) and speed \( v_0 \) can be used to define length and timescales for our simulation. For most of our simulation runs, we choose to use \( \Delta t = 0.1R_I/v_0 \), as in Vicsek et al.’s keystone paper.

If no other interaction rules were present and we implemented periodic boundary conditions, this interaction rule would result in exactly the same simulation as Vicsek et al.’s results in Reference [17].

The next interaction rule the ants follow is related to the effects of gravity in the cell. Since the ants in our column experiments need to be in contact with each other, we introduce gravitational
effects as a limit on their aligning behavior. A simulated ant can self-propel in whichever direction it would prefer, as long as it has a neighbor below it to help support it. If no neighbor is found, the ant is forced to take velocity pointing downwards, in effect imposing both a minimum density and a free surface on our simulated column of ants. An ant searches its list of neighbors within the distance $R_I$ to try to find one below it so that the vector drawn from the ant of interest to this neighbor makes an angle with the horizontal $-\frac{3\pi}{4} \leq \theta_G \leq -\frac{\pi}{4}$. If such a neighbor is found, the GRAVITY interaction makes no change to the ants’ velocity. If no such neighbor is found the ant takes a velocity:

$$v_x(t + \Delta t) = v_0 \cos(\Delta \theta_G)$$

$$v_y(t + \Delta t) = v_0 \sin(\Delta \theta_G)$$

where $\Delta \theta_G$ is a uniformly distributed random angle within $-3\pi/4, -\pi/4$.

Next, the ant looks for excluded area effects in the COLLIDE interaction. This time the ant searches its list of neighbors to find its nearest neighbor and checks if that neighbor is within a distance, $2R_B$, where $R_B < R_I$ is taken to be the ant’s body size. We model the ants as disks, so the factor of two relates the excluded area to the physical size of the ants. If the nearest neighbor is within $2R_B$, the ant changes its direction the minimum amount necessary to ensure that the distance between the two ants is greater than $2R_B$ after one time-step. It does this by first performing a coordinate transformation to rotate its updated velocity, $\vec{v}(t + \Delta t)$, into polar coordinates centered on its nearest neighbor. In these coordinates, $v_r$ is the ant’s velocity away from its nearest neighbor and $v_t$ is its velocity in the counter-clockwise direction orthogonal to $v_r$. After this coordinate transformation, the ant checks the radial component of its velocity to see if it is moving towards its nearest neighbor, $v_r < 0$. If so, the ant of interest receives a small positive radial velocity

$$v_r = \frac{5v_0}{R_I} (2R_B - d),$$

where $d$ is the distance between the two colliding ants. Given a time-step of $\Delta t = 0.1R_I/v_0$, this rule ensures that the two ants will not overlap in the next frame. The orthogonal component,
\(v_t\), remains unchanged. After the radial velocity is updated, \(v_r\) and \(v_t\) are rotated back to original coordinates to become \(v_x(t + \Delta t)\) and \(v_y(t + \Delta t)\).

This has the effect of making sure the ants cannot pass through one another, while avoiding the large amount of computing time needed to retrospectively enforce a true hard body interaction. By not changing the tangential component of the velocity, the effect of this collision is that the ants slide past each other, which only reinforces their alignment. The COLLIDE interaction is necessary in our simulation to prevent the ants from collapsing to form a single flock comprised of all the particles in a single point along the border of our cell, which would occur in Vicsek et al.’s simulations without periodic boundary conditions.

Finally the ants interact with the walls of the cell with the CONFINE interaction on Line 8. Similar to the ant’s COLLIDE interaction, the ant now checks to see if it is both within a distance \(R_B\) of one of the edges of the simulation cell and moving towards that wall. If both of these conditions are met, the ant loses the component of its velocity that would bring it closer to the wall, either \(v_x(t + \Delta t) = 0\) or \(v_y(t + \Delta t) = 0\). This ensures the ants stay in the simulation cell and also causes the particles to slide along the wall, which is typical for active particles, instead of bouncing off.

After all of the ants have updated their velocities by executing the above interactions in order, they update their positions based on their new velocities in Line 10. For time-steps of \(\Delta t = 0.1R_I/v_0\) and a physical choice for the ant’s body size \(R_B = 0.25R_I\), our simulation results exhibit waves that strongly resemble the experimental ant waves. An example of one of these waves is shown in Figure 3.5a. This shows that the simplified interactions we have chosen are enough to explain the waves and give us insight into their nature.
In the simulation, the waves are instances of collective-motion turned on and off by density. At high density, the ants align and flock towards the free surface. However the ants at the backside of the flock are constrained by the minimum density required to support the ants above them and as a consequence change their behavior, resulting in a density wave. Because high densities occur often near the base of the column, the waves are constantly produced.

In the experiments, the collective-motion of the ants is triggered not only by the density, but by the front in activity, so waves are only started by activity cycles. This is why the waves in experiment are not always observed, in contrast to what we find in the simulations. Before ants join the wave, they are relatively inactive with a relatively low density. They are then activated by the wave at a high enough density to flock with their neighbors, but finally, a minimum density is enforced by gravity limiting the motion of the ants at the back of the wave. This is illustrated in experiment by the velocities of the ants marked with green and red arrows in Figure 3.3a. They remain almost motionless before the wave, activate and align to move upwards during the wave, and are then restrained after the wave, when their velocity fluctuates as they try to continue moving up but find themselves at too low density to make further vertical progress.

3.2.2 Parameter Space

We find that the width of the column has no effect on the presence of the waves. Waves propagate in our simulations with similar frequencies and using the same parameters in all tower widths.
we have tested, from \( w = 7R_I \) to \( w = 35R_I \). This may be because, as in traveling bands, the Goldstone mode of collective motion causes the wave to spread out until it spans the width of the column.

In contrast, the waves are sensitive to our choices of both \( R_B \) and \( \Delta t \). If an ant’s body is too large compared to its legs, the ant cannot sense enough neighbors to trigger collective-motion. If \( \Delta t \) is too small, the ants align too quickly and never compress into waves. Instead, the whole simulated column mostly aligns, and the direction of the global collective motion fluctuates rapidly, as the motion of the ants is changed alternately by the CONFINE and GRAVITY interactions. There are only very small density fluctuations because as soon as two ants come within a distance \( R_I \) of each other, they interact rapidly and align, so the distance between them no longer decreases.

It is also worth noting that the waves are actually stronger if we allow the ants to drift down due to GRAVITY at a speed with different speed than they would usually self propel:

\[
\begin{align*}
v_x(t + \Delta t) &= v_g \cos(\Delta \theta_G) \\
v_y(t + \Delta t) &= v_g \sin(\Delta \theta_G)
\end{align*}
\]

where \( v_g < v_0 \), which is perhaps a better physical representation of our experiments, though this complicates the simulation slightly.

### 3.3 Analytical Model

We can use the idea of the waves as constrained instances of collective-motion to construct a one-dimensional, continuum, over-damped model of our system. To do this, we leave the discrete description in terms of ants and describe the active fluid in terms of material particles. We choose to use Lagrangian coordinates to follow the vertical position of the material particles \( y(x, t) \). The equation of motion for the material particles then reduces to:

\[
y_t(x, t) = \frac{F(x, t)}{\mu}
\]
where $\mu$ is the over-damping drag coefficient, $y_t$ is the partial derivative of the vertical position with respect to time, and $F(x, t)$ is the net force acting on a material particle. The variable $x$ is chosen to be the position of a material particle when the system is in an arbitrary state in which the fluid has a uniform mass density $\rho_c$, a critical density above which the ants that make up the fluid experience inter-particle repulsions and resist further compression. At a uniform density, the material particles are uniformly distributed; hence $y_x = \text{const}$. For simplicity, we choose this reference state to have $y_x = 1$, as sketched in Figure 3.6a. Our system is continuous, but for the purpose of illustration, we sketch it discretely. We emphasize that the variable $x$ does not change in time, and thus following the time dependence of $y(x_1, t)$ for a single value of $x_1$ is equivalent to following the trajectory of a material particle.

Figure 3.6: An illustration of the terms used in the solution to our analytical model. a) In the reference state, $y = x$ and the free surface lies at $M/\rho_c$. b) At $t = 0$, all material particles are in the uncompressed state (blue rectangle), where $y_x = \rho_c/\rho_d$. Therefore, $y(x_g) = x_g = 0$. c) At some arbitrary time later, the free surface has drifted downwards a distance $-g't$ and some material particles near the base of the column are compressed (red rectangle). Notice that $x_g(t_1)$ now refers to a different material particle than $x_g(0)$. d) Another arbitrary time, $t_2 > t_1$. Even more of the column is compressed.

Our over-damped model uses three ant behaviors that are switched on and off by the local one-dimensional density, $\rho = \rho_c/y_x$, where $y_x$ represents the partial derivative of $y$ with respect to $x$ from below. We use the one-sided derivative because, as we will see, there are discontinuities in our model and because the ants’ behavior in experiments seems limited by the distance to a
neighbor below it. The first two behaviors are for ants outside of an activity cycle.

When the local density in our model is too low, $y_x^- > 1$, the “inactive” particles that make up the fluid do not resist compression, and they slowly drift downwards affected by nothing but an effective gravity, $F = -mg$, where $m$ is the mass of a material particle. Experimentally, this relates to the tendency of the ants to form dense clusters when they are inactive so the density can increase without resistance. In the model, we can imagine the material particles to be composed of a point mass with two springs on either side that would resist compression. At densities below $\rho_c$, the springs are not in contact, as illustrated in the leftmost schematic in Figure 3.7.

![Figure 3.7](image)

Figure 3.7: The three key ingredients of the continuum model correspond to three density regimes in the over-damped PDE: Uncompressed ants (blue springs) drift down at constant speed due to gravity. Compressed ants (green springs) are subjected to competing gravitational and decompressing forces. For active ants (red springs), gravity and compression effects are ignored and ants travel upwards at a set velocity.

However, as we increase the density past $\rho_c$, $y_x^- < 1$, and the particles begin to resist further compression. In the experiments this is what leads to the density gradient between the top and bottom of the column. In this model, we now consider the springs of the material particles to be in contact and exert a linear spring force on adjacent material particles, as shown in the center schematic of Figure 3.7. The force from the springs below the particle pushes it upwards, $F_1 = ky_x^-$, and the springs above the particle push it down, $F_2 = -ky_x^+$. The net force on the particle is then the limit of the sum of these two competing forces: $F_{\text{spring}} = ky_{xx}$.

Finally, if the density of the ants increases above another density, $\rho_a$, as shown in the rightmost schematic in Figure 3.7, the material particles initiate an activity cycle. Remember that in
experiments, activity cycles seem to start at random as long as the density is above a threshold that corresponds to $\phi_{eff} \approx 3.6$, but for simplicity, in this model we assume they always activate at the threshold density, $\rho_a$. If the material particles compress even further to some critical density, $y_x^- = \rho_c/\rho_a$, they activate, switching from inactive to active, and travel upwards with a constant speed, $y_t = v_0$, disregarding the passive forces around them until they are restricted by the density below them. This happens when $y_x^- = \rho_c/\rho_d$, where $\rho_d$ is the minimum achievable density in our model and represents the state in which the active material particles are as extended as possible. At that moment, the material particle at $y(x, t)$ deactivates, completing its activity cycle.

Altogether, these rules can be written

$$y_t = \begin{cases} 
\Psi(1 - y_x^-) \frac{k y_{xx}}{\mu} - \frac{m g}{\mu} & \text{inactive} \\
v_0 & \text{active}
\end{cases}$$

where $\Psi$ is the Heaviside step-function, and

inactive $\xrightarrow{y_x^- = \rho_c/\rho_a}$ active $\xrightarrow{y_x^- = \rho_c/\rho_d}$ inactive

In summary, the material particles compress solely under the influence of gravity, until $\rho = \rho_c$, when the Heaviside function begins to include spring forces resisting further compression. Then, when $\rho = \rho_a$, the material particles activate and move upwards with a speed $v_0$. They continue moving until $\rho = \rho_d$, at which point they deactivate and repeat the process.

3.3.1 Inactive Solution

To solve this equation analytically we assume time-scale separation:

$$|v_0| \gg \left| \Psi(1 - y_x^-) \frac{k y_{xx}}{\mu} - \frac{m g}{\mu} \right|$$
always. This means that the dynamics of the active material particles happen on much shorter timescales than the compression of the column and that, when we solve for the behavior of the active particles, we consider the inactive material particles to be stationary. In effect, this allows us to break our problem into two parts. The first is the time evolution of the inactive material particles from a uniform fully extended state to a state that triggers activation, and the second is the time evolution of the front caused by the active material particles that returns the system to the fully extended state.

We start with the inactive solution and rescale the gravitational and spring constants for convenience:

\[ g' = \frac{mg}{\mu}, \quad k' = \frac{k}{\mu} \]

Let the system consist of a finite mass, \( M \), of our model fluid, so that the initial state of the system has our material particles uniformly distributed between \( y(0, 0) = 0 \) and a free surface at \( y(M/\rho_c, 0) = M/\rho_d \), as shown in Figure 3.6b. We will limit the position of the bottom of the column, so that \( y(0, t) = 0 \) always. This uniformly distributed state has a density \( \rho_d = \rho_c/y_x \). From this state, all of the material particles begin to drift downwards, due the over-damped force of gravity, \( y_t = -g' \). Immediately after, the bottom reaches a state where \( y_x < 1 \) because the bottom of the column is stationary. This divides our system in two regions. Near the top of the column, in the regions in the dashed blue boxes in Figures 3.6b-d, material particles have a velocity \( y_t = -g' \), but at the bottom, in the regions in the dashed red boxes in Figures 3.6c-d, they have a velocity \( y_t = k' y_{xx} - g' \).

Between these two regions, an interface can be defined that now moves upwards over time as more and more material particles find themselves compressed at the bottom of the column. We find it convenient to solve for the local density in the column and the position of the free surface in terms of the material particle that sits at the interface at any given time, which we call \( x_g(t) \), referring to that material particle’s position in our reference state. Exactly which material particle is sitting at the interface changes as a function of time, as shown by the red, green, and blue material particles.
in the sketch in Figure 3.6a-d. Notice that \( x_g(t) \) refers to different material particles at different times. At \( t = 0 \), \( x_g(t) = 0 \) because the interface is just the bottom particle, as shown in Figure 3.6b. With this notation, the position of the interface is \( y(x_g(t)) \).

In keeping with the over-damped nature of our model fluid, we consider that inertial effects are non-existent and the spring forces that resist compression adjust much faster than material particles move, so that the particles in the compressed region are always in mechanical equilibrium and \( y_t \approx 0 \). Then the only time dependence of the position of the particles comes from the total mass of fluid in the compressed region, which is controlled by \( x_g(t) \). By integration under these assumptions:

\[
\begin{align*}
    y_{xx} &= \frac{g'}{k'}; & y_x &= \frac{g'}{k'} x + C_1; & y &= \frac{g'}{2k'} x^2 + C_1 x,
\end{align*}
\]

where \( y(x,t) \) is the position of the particle that would be at position \( x \) in the reference state at the time \( t \).

Here the time dependent behavior of \( y(x,t) \) is controlled by the behavior of the integration constant, which we can solve for in terms of \( x_g(t) \) by recalling that \( y_x(x_g(t)) = 1 \) by definition, so

\[
C_1 = 1 - \frac{g'}{k'} x_g(t)
\]

We next turn to the uncompressed region, \( y \geq y(x_g(t)) \). Here the density is still set by the initial conditions and thus

\[
\begin{align*}
    y_x &= \frac{\rho_c}{\rho_d}; & y &= \frac{\rho_c}{\rho_d} x + C_2
\end{align*}
\]

Enforcing continuity of \( y(x,t) \) at \( y(x_g(t)) \) allows us to solve for \( C_2 \). We obtain

\[
C_2 = \left( 1 - \frac{\rho_c}{\rho_d} \right) x_g(t) - \frac{g'}{2k'} x_g(t)^2
\]
We can then write the solution for $y$ both above and below $y(x_g(t))$ in terms of $x_g(t)$

$$y(x,t) = \begin{cases} \frac{g'}{2k'}x^2 + \left(1 - \frac{g'}{k'}x_g(t)\right)x & x \leq x_g(t) \\ \frac{\rho_c}{\rho_d}(x - x_g(t)) + x_g(t)\left(1 - \frac{g'}{2k'}x_g(t)\right) & x \geq x_g(t) \end{cases}$$

(3.1)

The solution written this way for the uncompressed region shows that trajectory of a material particle, $x = \text{const.}$, has the same complicated time dependence as $x_g(t)$, scaled by its position in the reference state. Writing the solution for uncompressed region this way is not physically transparent because $x_g$ has no direct effect on those particles’ behavior. We remember that the only force acting on the uncompressed material particles is the damped gravity, so we know their trajectories to be linear in time.

We can use this knowledge to solve for the time dependence of $x_g(t)$. Recall that while the material particles near the base are compressing, the free surface continues to drift downwards unconstrained due to the over-damped gravity term, as show by the gold material particles in Figures 3.6b-d. We use that $y_t(M/\rho_c, t) = -g'$ and employ Equation 3.1 to write down the trajectory of the free surface $y(M/\rho_c, t)$:

$$y\left(\frac{M}{\rho_c}, t\right) = \frac{M}{\rho_d} - g't = \frac{\rho_c}{\rho_d}\left(\frac{M}{\rho_c} - x_g(t)\right) + x_g(t)\left(1 - \frac{g'}{2k'}x_g(t)\right)$$

Solving the above equation for $x_g(t)$, we can write down which material particle is on the interface for a given time, $x_g(t)$, before the wave begins to propagate through the column:

$$x_g(t) = \frac{k'}{g'}\left(1 - \frac{\rho_c}{\rho_d}\right) \pm \sqrt{\left(\frac{\rho_c}{\rho_d} - 1\right)^2 + \frac{2g'^2}{k'}t}$$

One of these two solutions is unphysical: $\rho_c/\rho_d > 1$, so

$$\frac{k'}{g'}\left(1 - \frac{\rho_c}{\rho_d}\right) = \sqrt{\left(\frac{\rho_c}{\rho_d} - 1\right)^2 + \frac{2g'^2}{k'}t} < 0$$

at all times.
The other solution has the correct behavior:

\[ x_g(t) = \frac{k'}{g'} \left( \left(1 - \frac{\rho_c}{\rho_d}\right) + \sqrt{\left(\frac{\rho_c}{\rho_d} - 1\right)^2 + \frac{2g'^2}{k'} t} \right) \]

This shows the material particle at the interface when \( t = 0 \) has a reference position of \( x_g(t = 0) = 0 \), so the first particle to find itself on the interface is the bottom one. After that, the proportion of material particles below the interface increases with \( t \). Equation 3.1 shows that this causes material particles below the interface to continually sink, which will constantly increase the density at the base of the column. On the other hand, the position of the interface, \( y(x_g(t)) \), increases for short times, but eventually decreases. Using Equation 3.1, and substituting \( x_g(t) \) for \( x \) shows us the position of the interface at a given time:

\[ y(x_g(t)) = x_g(t) - \frac{g'}{2k'} x_g^2(t) \]

Because \( x_g(t) \) monotonically increases with \( t \), \( y(x_g(t)) \) has a maximum that occurs when \( x_g(t) = k'/g' \).

As time goes on, more and more of the material particles pass through the interface until either the interface reaches the free surface, \( x_g(t) = M/\rho_c \), in which case the system reaches a steady state with all of the ants in the compressed region, or until the density at the bottom of the column reaches \( \rho_a \), activating the ants at the bottom of the column.

This happens when \( g_x = \rho_c/\rho_a \), which first occurs at the base of the column. Thus, we consider our solution for \( y_x(x, t) \) in the compressed region after substituting \( C_1 \).

\[ y_x(x, t) = \frac{g'}{k'} (x - x_g(t)) + 1 \]

Substituting the time dependent solution for \( x_g(t) \) and solving for the time that \( y_x(0, t) = \rho_c/\rho_a \) allows for the solution for the time that the material particles at the bottom of the column first
activate.

\[
t' = \frac{k'}{2g^2} \left( \left( \frac{\rho_c}{\rho_a} \right)^2 + \frac{2\rho_c}{\rho_d} - \frac{2\rho_c^2}{\rho_a \rho_d} - 1 \right)
\]

An example of the density profile of the column at this instant is shown by the solid black curve in Figure 3.8, which shows the density in the fluid column as a function of height. The fluid at the bottom of the column is compressed, \(\rho/\rho_c > 1\) but there is a region of constant density near the top of the column. The discontinuity between the two regions at \(y\rho_c/M \approx 0.2\) is \(y(x_g(t'), t')\). Though \(y(x, t)\) is continuous, which means there are no positions without a material particle, \(y'(x, t)\) is not, and the density drops from \(\rho_c\) to \(\rho_d\) at \(y(x_g, t)\).

### 3.3.2 Wave Solution

Now, we begin the solution for the propagation of the wave, assuming it moves on a much faster timescale, so that while the wave is propagating, the gravitational relaxation of the column can be ignored. We define a new timescale that begins with \(t_3 = 0\) at the moment the wave has formed and assume that \(t' + t_3 \approx t'\).

The wave begins at the base, \(y(x = 0, t_3 = 0)\), and we solve for the speed of both the front and back ends of the wave by considering the distance that a material particle needs before it can activate the material particle above it. The speed at which the front end of the wave propagates, \(v_f\), is related to the speed of the activated material particles and the relative distance between them. One way to understand this relationship is by examining the time it takes for a wave to travel between two material particles. This time is set by how long it takes the material particle, moving at speed \(v_0\) to compress the region above it from \(1/y^+_{x}\) to \(\rho_c/\rho_a\):

\[
\frac{y^+_{x}}{v_f} = \frac{y^+_{x} - \rho_c/\rho_a}{v_0}
\]

This implies that the front of the wave travels at the speed

\[
v_f = \frac{dy_f}{dt} = v_0 \frac{y^+_{x}}{y^+_{x} - \rho_c/\rho_a}
\]
Near the base of the column \( y_x^+ > \rho_c/\rho_a \), but only slightly larger. This makes the denominator of Equation 3.2 small and causes the front end of an activity wave to move explosively at first. The front of the wave continues to slow down as the denominator of Equation 3.2 increases with increasing \( y \). This slow down continues until the time at which the front of the wave reaches \( y(x_g(t')) \), which we call \( t_3 = t^* \) and solve for by separating and integrating Equation 3.2:

\[
\int_{t'}^{t+t^*} v_0 dt = \int_0^{y(x_g(t'))} \frac{y_x^+ - \rho_c/\rho_a}{y_x^+} dy_f = \int_0^{x_g(t')} y_x^+ - \frac{\rho_c}{\rho_a} dx_f
\]  

(3.3)

We then use that, at \( t = t' \),

\[ y_x(x, t') = g' \sqrt{x} + \frac{\rho_c}{\rho_a} \]

and

\[ x_g(t') = k' \frac{g'}{g_f} \left( 1 - \frac{\rho_c}{\rho_a} \right) \]

to solve the far right integral in Equation 3.3:

\[
\int_0^{x_g(t')} \frac{y_x^+ - \rho_c/\rho_a}{y_x^+} y_x^+ dx_f = \int_0^{x_g(t')} \frac{g'}{k^2} x_f dx_f = \frac{g'}{2k} x_g(t')^2 = \frac{k'}{2g'} \left( 1 - \frac{\rho_c}{\rho_a} \right)^2
\]

which results in

\[ t^* = \frac{1}{v_0} \frac{k'}{2g'} \left( 1 - \frac{\rho_c}{\rho_a} \right)^2 \]

After \( t_3 = t^* \), the spacing between material particles in front of the wave is constant, \( y_x^+ = \rho_c/\rho_d \), and by Equation 3.2, the wavefront travels with a constant velocity, as in the experiments and simulations:

\[ v_f = \frac{dy_f}{dt} = v_0 \frac{1/\rho_d}{1/\rho_d - 1/\rho_a} \]

Similarly, material particles at the back end of the wave need to move far enough to decrease from \( y_x^- = \rho_c/\rho_a \) to \( y_x^- = \rho_c/\rho_d \) to deactivate. We can make the same comparison of the distance traveled by material particles and the distance traveled by the back end of the wave as we did to
calculate the speed of the front edge. In this case,

\[
\frac{\rho_c/\rho_d}{v_b} = \frac{\rho_c/\rho_d - \rho_c/\rho_a}{v_0}
\]

where \( v_b \) is the speed of the trailing edge of the wave. This leads the trailing edge of the wave to move at a speed

\[
v_b = \frac{dy_b}{dt} = v_0 \frac{1/\rho_d}{1/\rho_d - 1/\rho_a}
\]

which is the same speed the front of the wave travels in the uncompressed region.

Before the wave reaches \( y(x_g(t^*)) \), the front of the wave travels faster than the back of the wave. This leads to the wave increasing in width from a single point, when \( t_2 = 0 \), until it incorporates a mass of ants

\[
m = \rho_a (y_f(t^*) - y_b(t^*)) = \frac{k'}{g'} \left[ \rho_c - \frac{\rho_c^2}{\rho_a} - \frac{\rho_d\rho_a}{2(\rho_a - \rho_d)} \left( 1 - \frac{\rho_c}{\rho_a} \right)^2 \right]
\]

by \( t_3 = t^* \).

Figure 3.8 shows the density, \( \rho_c/y_x \), of our model as a function of \( y \) during the formation and propagation of a wave. After the wave is triggered at the bottom of the column at \( t_3 = 0 \) (black solid line), the front of the wave initially moves much faster than the tail, giving the wave finite width at \( t_3 = t^*/2 \) (red dashed line). The width of the wave continues to grow until the wave reaches \( y(x_g(t')) \) and then the wave travels with a constant speed and amplitude through the constant density region of our model (blue dotted and green dashed lines). The wave continues to move at a constant speed with a constant mass until the back of the wave reaches the free surface, when the density of the column is once again \( \rho_c/y_x = \rho_d \) everywhere. At this point, the column of ants has returned to the fully extended state used as initial conditions and the cycle repeats itself.
Figure 3.8: Analytical solutions to the PDE for four different times, where $t^*$ is the time that the wave reaches the interface between the compressed and uncompressed regions and we have chose $\rho_a = 5/4$, $\rho_d = 4/5$, and $\rho_c = 1$. The curves have been offset vertically relative to the $t = 0$ curve for clarity. The first curve is immediately after the density at the bottom of the column reaches the critical density $\rho_a$. The second shows the wave during the expansion phase in the compressed region. The final two show the wave traveling at a constant speed through the uniform density background.

3.4 Wave Nonlinearity: Speed Dependence on Amplitude

To measure the speed and amplitude of the waves in our simulations, we parallel our original experimental analysis of the waves and create space-time diagrams. We do this by sorting our simulated ants by their vertical positions into bins that are each $0.1R_I$. We smooth our bin counts with a moving average over a distance of $2R_I$ and repeat the process for every tenth frame of our simulation. The result is that we have a count of the number of particles in each bin as a function of both space and time, as shown in Figure 3.9a. From the space-time diagrams, we can see the leading edges of the waves by the sudden jumps in density for a given height. We fit these points with a line for each wave to measure the wave’s speed, $s$, and average the counts detected in these bins as a measure of the wave’s amplitude, $\rho_{max}$. Figure 3.9b shows that, though there is some variation in the speed and amplitude of the waves, the variations do not appear correlated. Hence, $s$ is not a function of $\rho_{max}$. 
In our analytical model, the speed of the waves is determined by the chosen parameters, $v_0$, $\rho_d$, and $\rho_a$. For a given choice of these parameters, the waves have a set density of $\rho_a$ and a speed

$$s = v_0 \frac{1/\rho_d}{1/\rho_d - 1/\rho_a}.$$ 

In our model, a wave is always formed when $\rho(0,t) = \rho_a$, but we can easily imagine that a more realistic picture would be that $\rho_a$ varies from wave to wave because we know that, in experiments, the waves form spontaneously. For values of $\rho_a/\rho_d$ only slightly larger than 1, $s$ strongly decreases with increasing $\rho_a/\rho_d$, but as $\rho_a/\rho_d$ is increased, the speed of the wave plateaus towards $v_0$, as shown in Figure 3.9c. Realistic values for $\rho_a/\rho_d$ from our experimental measurements are $1.25 < \rho_a/\rho_d < 1.5$. However the densities in our analytical model are 1D instead of 2D, so they may not map well to the our experiments.

### 3.4.1 Experimental Measurements

Motivated by the simulation findings, namely that $s$ is not a function of $\rho_{max}$, and by the analytical expectations, $s \approx v_0$ for large $\rho_a/\rho_d$, we return to our experiments.

To measure the amplitude of the experimental waves, we compare two images separated by a 40 s interval, such as those shown in Figure 3.10a and Figure 3.10b. We choose this interval to ensure that the wave has traveled more than its own width in between the chosen frames, but our
results do not qualitatively change if we select a slightly different interval. We blur the transmis-
sions of each of these two images by convolving the images with 2 cm disks and convert them to
2D densities using the calibration curve shown in 3.1e.

The resulting difference highlights regions where the density has recently changed, as shown
in Figure 3.10c. The blue areas in Figure 3.10c are regions in which the density has recently
decreased and the red areas are regions in which the density has recently increased. The wave is
immediately apparent as the darkest red region.

We take the maximum change in density, marked with an × in Figure 3.10c, $\Delta \rho_{max}$ to be the
instantaneous amplitude of the wave during the later frame. We find that the amplitude remains
relatively constant over the last 50 s before the wave hits the free surface but typically increases
slightly over time, like the example shown in Figure 3.10d, so we average it over the last 50s to
take a measure of the amplitude of the wave, $\langle \Delta \rho_{max} \rangle$, and report the error as ±10%.

The amplitudes of our waves vary from $\langle \Delta \rho_{max} \rangle = 4 \text{ ants/cm}^2$ to $\langle \Delta \rho_{max} \rangle = 10 \text{ ants/cm}^2$. 

Figure 3.10: a) An image of a wave propagating through the column of ants. b) Another image of the same wave 40 s later. c) The result of blurring and subtracting the two intensities, then converting to the transmission difference into a density difference. The × indicates the location with the peak change in density between these two frames, $\Delta \rho_{\text{max}}$. d) The peak change in density between frames as a function of time for the last 50 seconds before the wave reaches the free surface. The solid line indicates the average amplitude of the wave, and the dotted lines indicate the error. e) A measurement of the speed of the wave. The red line is a linear fit of the position of the peak of the wave. Its slope is used as the speed of the wave, $s$.

Next, we measure the speed of the waves from their space-time diagrams, like the one in 3.1f. For each measured height, which is each vertical pixel from our images, we find the time that corresponds to the maximum density, $t_m(y)$, which is demonstrated by the black curve in Figure 3.10e. From $t_m(y)$, we pick the lowest height above which the wave travels at a clearly constant velocity, $h_b$, and then fit $t_m(y)$ from $y = h_b$ to the free surface with a line, like the red line in Figure 3.10e.

We find that the speed of the waves also varies from $s = 0.4 mm/s$ to $s = 1.5 mm/s$ which is comparable to an ants’ walking speeds at very low densities, in approximate agreement with our simulation and analytical model.
However, plotting the speed of the waves against their amplitude in Figure 3.11a shows that the activity waves are strongly nonlinear. The speed of the wave exhibits a strong dependence on amplitude.

![Figure 3.11:](image)

We find that we can correct our simulation to achieve the macroscopic nonlinearity of the waves by adding a microscopic nonlinearity to the speed of the particles. In the simulations, we do this by modifying the ALIGN interaction so that \( v_0 = C_N N \), where \( C_N \) is a constant that ensures that the average speed of the particles is still \( v_0 = 1 \), and \( N \) is the total number of ants within range \( R_I \). This makes the ants’ self-propulsion speed increase when it detects more nearby neighbors. After remeasuring the speed and amplitude of the waves, we find that this microscopic nonlinearity produces a macroscopic nonlinearity in the waves’ speeds, as shown in Figure 3.11b.

Similarly, allowing \( v_0 \propto \rho_a \) in the analytical model is enough to overcome the decrease of \( s \) with \( \rho_a \) and cause \( s \) to increase with \( \rho_a \), provided \( \rho_a/\rho_d > 2 \), as shown in Figure 3.11c. In both cases, it seems that a local nonlinearity is enough to lead to the nonlinearity of the waves.

### 3.4.2 Possible Source of the Nonlinearity

To address the physics associated to the proposed local nonlinearity, we return to Figure 3.4d, which shows our waves in order and in density, and compare it to Figure 1.9c, which shows the waves in order and density for Chaté et al.’s traveling bands. In their simulations of traveling...
bands, the density and order parameter mirror each other without delay, which is a hallmark of linear responses. The lack of delay is easy to understand on the leading edge of the traveling bands. The disordered phase is at low density compared to the waves, so as soon as the wave overtakes a particle in front of it, the particle feels the alignment interactions with all of its new neighbors pressuring it to move in the same direction. In just a single time-step, the particle aligns and joins the wave, so that the particle has provided negligible resistance to the passing of the wave.

In contrast, in our ant waves, the wave in order lags the wave in density. When an activity wave overtakes ants in the cell, the disordered phase is already at a relatively high density, so the active collectively moving ants cannot progress through the disordered phase until the disordered phase has been activated and aligned. The wave stalls behind the inactive ants until they activate. Modeling by Tennenbaum et al. [13] showed that the activation rate for ants was dependent on the number of active ants already active in the area. This helps explain the local source of the nonlinearity. The stalled ants pressure the inactive ants in front of the wave to activate, and the more stalled active ants there are, the faster they can pressure the inactive ants to activate.

### 3.5 Conclusions

We have shown that activity cycles can trigger activity waves in vertical columns of fire ants. The waves are caused by the ants switching their behaviors to activate and briefly exhibit collective-motion. Before the wave passes, the ants are dense but mostly inactive. When the wave overtakes the ants, the ants begin to activate, and the combination of high density and high activity causes the ants to align and undergo local collective-motion. Then, at the back end of the wave, the ants are constrained by a minimum density required to support their vertical motion. They remain active for a time but are unable to continue moving vertically upward and the collective motion is arrested. The ants that have been left outside of the wave then gradually become caged in by their deactivating neighbors, until they too deactivate.

Measurement shows that these waves behave nonlinearly, and we ascribe this to a nonlinear
activation rate. We propose that denser waves are able to pressure the ants in front faster to activate them and allow the wave to continue moving upwards.
In the last chapter, we saw the striking effects of the ants’ activity manifest as activity waves in 2D. Now, we turn our attention to the ants’ behavior in 3D, using an experimental set up inspired by work on the behavior of corn in a silo. Granular materials, like piles of corn, are passive systems, but they are out of thermodynamic-equilibrium because their grains are large enough to be athermal. This makes the comparison between passive granular systems and ants, another out-of-thermodynamic equilibrium system, compelling. Frictional forces often frustrate the motion of granular systems, preventing them from rearranging and therefore flowing. In this chapter, we’ll begin by considering the difference in behavior between a liquid and a granular system in a simple container. We’ll find that friction causes the walls of the container to support the grains against compression. However, for sufficiently narrow columns, we find an intermediate regime where we obtain the opposite effect. We will then repeat these experiments with active grains, in this case ants, instead of passive grains. We know that friction allows the ants to build connections to behave like a solid on short timescales like a granular system, but also that they flow on long timescales [78, 81] like a fluid. An interesting question is whether activity fluidizes a system of ants so that the walls of a container can no longer support their weight. Can fluctuating connections still support an active granular system?

4.1 Review of Janssen’s Experiment

In 1895, H.A. Janssen studied the forces that corn kernels exert on the walls of a silo and found a striking deviation from fluid behavior [99, 100]. His experimental system was a silo of corn with walls supported independently of the base, which was separately connected to a scale, as shown in Figure 4.1a. This allowed him to measure the force the corn applied directly to the base and then extrapolate to predict the rest of the forces the corn exerted on a silo.
4.1.1 Fluid Columns

Consider the behavior of an isotropic fluid in hydrostatic equilibrium confined to a cylindrical container with diameter $D$.

The Cauchy stress tensor, $\sigma$, for the fluid is a scalar matrix, defined by the pressure, $p$:

$$\sigma = -p I$$

Due to the symmetry of our problem, we will choose cylindrical coordinates. Hence:

$$\begin{pmatrix}
\sigma_{rr} & \sigma_{r\theta} & \sigma_{rz} \\
\sigma_{\theta r} & \sigma_{\theta\theta} & \sigma_{\theta z} \\
\sigma_{zr} & \sigma_{z\theta} & \sigma_{zz}
\end{pmatrix} =
\begin{pmatrix}
-p & 0 & 0 \\
0 & -p & 0 \\
0 & 0 & -p
\end{pmatrix}$$ (4.1)

We can calculate the pressure at a given depth by enforcing mechanical equilibrium on a thin slice of the fluid with thickness $dz$ at a distance $z$, from the free surface, like the one shown in Figure 4.1b. The only forces acting on the slice are the body force due to gravity and the contact forces on each of its surfaces. The force from stress on the top of the slice is

$$\vec{F}_{\text{top}} = \int_A \sigma \cdot d\vec{S} = \sigma(z) \cdot (-A\hat{z}) = -\sigma_{zz}(z) \pi D^2 \hat{z} / 4,$$

where $-\hat{z}$ is the unit vector normal to the top surface. Likewise, the force from stress on the bottom of the slice

$$\vec{F}_{\text{bottom}} = \sigma(z+dz) \cdot A\hat{z} = \sigma_{zz}(z+dz) \pi D^2 \hat{z} / 4.$$

Finally, the gravitational force exerted on our slice is

$$\vec{F}_g = mg\hat{z} = \rho (\pi D^2/4) gz dz.$$

Mechanical equilibrium requires

$$\sum \vec{F} = \vec{F}_{\text{top}} + \vec{F}_{\text{bottom}} + \vec{F}_g = \rho \frac{D^2}{4} g\hat{z} dz - \sigma_{zz}(z) \frac{D^2}{4} \hat{z} + \sigma_{zz}(z+dz) \frac{D^2}{4} \hat{z} = 0,$$

which, after considering that $\sigma_{zz}(z+dz) - \sigma_{zz}(z) = d\sigma_{zz}$ and rearranging, yields:

$$\frac{d\sigma_{zz}}{dz} = -\rho g.$$

We can integrate this equation to find $\sigma_{zz}(h)$, the stress at the base of a fluid with a height $h$,
assuming that the stress at the free surface is \( \sigma_{zz}(0) = 0 \):

\[
\int_0^{\sigma_{zz}(h)} d\sigma_{zz}' = -\int_0^h \rho g dz
\]

\[
\sigma_{zz}(h) = -\rho gh.
\]

Since \( \sigma_{zz} = -p \), we find \( p = \rho gh \). Due to the isotropy of \( \sigma \), we also have \( \sigma_{rr} = \sigma_{\theta\theta} = -\rho gh \).

The total force detected by a scale placed at the bottom of the column is \( -\sigma_{zz}(h)A \), which we relate to what we call the apparent mass, \( m_{\text{app}} \), using the acceleration of gravity, that in this case coincides with the added mass of fluid in the column, \( m_{\text{add}} \):

\[
m_{\text{app}} g \equiv F_{\text{net},z} = \rho gh \frac{\pi D^2}{4} = m_{\text{add}} g.
\]

Figure 4.1: a) Janssen’s original setup. The base of his column is connected to a balance that he loaded with varying mass (marked G), and the column walls are connected to the floor with springs (marked S). b) An illustration of a horizontal slice in a cylinder that is in mechanical equilibrium. Possible forces on the slice in include gravity, compressive stresses, and frictional forces with the walls. c) Navy line: The theoretical prediction for the apparent mass detected by a scale for a fluid in Janssen’s setup. Navy and red points: Two trials from Janssen’s original experiments using a square column with side length \( l = 60 \text{ cm} \). Black line: The theoretical prediction assuming a homogeneous isotropic elastic material at the Coulomb limit with its walls.

4.1.2 Granular Systems

When Janssen measured \( m_{\text{app}} \) under a column of corn, he found that the apparent mass detected by his scale saturated as more and more corn was added, as shown by the points in Figure 4.1c,
which are Janssen’s experimental results for corn in a square prism with side length $L = 60 \text{ cm}$ [99, 100]. Note the deviation from fluid behavior, which we show with the blue line in the same figure, reflecting that, for the fluid case, $m_{\text{app}} = m_{\text{add}}$. The only possible explanation was that the walls of the column were supporting the mass of the grains against gravity.

To understand this, Janssen modeled the granular system as an elastic continuum solid interacting with the walls via friction and made a few simplifying assumptions. To illustrate Janssen’s reasoning, let’s first consider an elastic solid, confined to our cylindrical container without friction between the solid and our container walls. For a homogeneous and isotropic elastic solid, the Cauchy stress tensor, $\sigma$, is set only by the strain tensor, $\epsilon$, and two independent material constants, the Lamé coefficients: $\lambda$, an elastic modulus, and $\mu$, the shear modulus. Explicitly:

$$\sigma = 2\mu\epsilon + \lambda \text{Tr}[\epsilon] I$$  \hspace{1cm} (4.2)

where $I$ is the identity tensor. Equation 4.2 is an expression of Hooke’s Law in three dimensions and it holds for any coordinate system.

In our case, we expect the weight of the solid, due to the gravitational force, to compress it, resulting in a vertical strain, $\epsilon_{zz} > 0$. Additionally, the walls of the cylinder restrict the solid so that there is no radial strain. We can now recalculate the forces exerted on a slice of the column, as we did in the previous section. All of the forces exerted on the slice are the same and we retrieve the same result for the $h$ dependence of $\sigma_{zz}$:

$$\sigma_{zz} = -\rho gh$$

However, if we use Equation 4.2, instead of Equation 4.1, we find that

$$\sigma_{rr} = \sigma_{\theta\theta} = \frac{\lambda}{\lambda + 2\mu} \sigma_{zz} \equiv k \sigma_{zz}.$$  \hspace{1cm}

For most materials, $\lambda$ and $\mu$ are positive constants, so $k$ is also a constant and $k < 1$. 

110
This change, however, does not effect the readout of a scale at the base of the column without friction:

\[ m_{\text{app}} = -\frac{\pi D^2}{4g} \sigma_{zz}(h) = m_{\text{add}} \]

Now, let us allow for friction between the elastic solid and the walls of the cylinder. Unlike fluids in hydrostatic equilibrium, our elastic solid can support the shear stress that results. The radial component of the stress tensor \( \sigma_{rr} \) provides the normal force required for friction, which in turn provides a shear stress that resists further compression of the elastic solid. Following Janssen, we now make two simplifying assumptions to analytically solve the problem. The first is that even in the presence of friction, \( \sigma_{zz} \) is a function only of \( z \):

\[ \sigma_{zz} = f(z) \quad (4.3) \]

This amounts to \( \sigma_{zz} \) being constant in our cross-sectional slice. The second assumption is that friction is maximally utilized: \( F_f = \mu_s F_N \), where \( F_f \) is the force of friction, \( \mu_s \) is the static coefficient of friction, and \( F_N \) is the normal force. In the literature, this assumption is often called the Coulomb limit. Under these two assumptions,

\[ \sigma_{zz} = \mu_s k \sigma_{zz}. \]

We now apply the condition of mechanical equilibrium to our slice from 4.1b. The forces on the top and bottom of the slice remain the same as before, as does \( \vec{F}_g \). However, now there is a new vertical component to the friction between the solid and the walls:

\[ \vec{F}_f = -\sigma_{zz} \hat{z} \pi Ddz = \mu_s k \sigma_{zz} \hat{z} \pi Ddz \]

Mechanical equilibrium applied to our slice then reads:

\[ \sum \vec{F} = \rho \frac{\pi D^2}{4} g \hat{z}dz - \sigma_{zz}(z) \frac{\pi D^2}{4} \hat{z} + \sigma_{zz}(z + dz) \frac{\pi D^2}{4} \hat{z} + \mu_s k \sigma_{zz} \hat{z} \pi Ddz = 0, \quad (4.4) \]
Rearrangement yields
\[
\frac{d\sigma_{zz}}{\rho g D/4\mu_s k + \sigma_{zz}} = -\frac{4\mu_s k}{D} dz
\]

We make the substitution \( u = \rho g D/(4\mu_s k) + \sigma_{zz} \) and integrate:

\[
\int_{\rho g D/4\mu_s k}^{\rho g D/4\mu_s k + \sigma_{zz}(h)} \frac{du}{u} = -\frac{4\mu_s k}{D} \int_0^h dz
\]

\[
\ln \left( 1 + \frac{4\mu_s k}{\rho g D} \sigma_{zz}(h) \right) = -\frac{4\mu_s k h}{D}
\]

\[
\sigma_{zz}(h) = \frac{\rho g D}{4\mu_s k} \left( \exp \left( \frac{-4\mu_s k h}{D} \right) - 1 \right).
\]

Using that the magnitude of the force at the base of the cylinder is \( |\sigma_{zz}(h)A| = m_{app} g \), we find:

\[
m_{app} = -\sigma_{zz}(h) \frac{\pi D^2}{4g} = \frac{\rho \pi D^3}{16\mu_s k} \left( 1 - \exp \left( -\frac{4\mu_s k h}{D} \right) \right)
\]

It is now convenient to define the apparent saturation mass,

\[
M = \frac{\rho \pi D^3}{16\mu_s k}
\]

so that

\[
m_{app} = M \left( 1 - \exp \left( -\frac{m_{add}}{M} \right) \right) \tag{4.5}
\]

Notice the limits of this equation. For \( m_{add} \gg M \), \( m_{app} \to M \), while for \( m_{add} \ll M \), Taylor expansion shows \( m_{app} \to m_{add} \), whereby we recover fluid-like behavior.

The black curve in Figure 4.1c shows that this theory fits Janssen’s results very well, but it is important to remember that it rests on assumptions that oversimplify the situation. Nevertheless, the model does capture the role of friction between the grains and the walls, which helps support the weight of the column as the column becomes increasingly tall.

Today we understand compressed granular systems to be made up of networks of “force chains,” in which a subset of the grains experience compressive forces that link them together. These force
chains can link grains in a cylinder and push the grains against the wall. Typically these force chains form arches or “cones of equal pressure,” so that the force of friction at the wall pushes the particles upwards and supports the mass of the column. In this sense, our assumption that $\sigma_{zz}$ was constant across the slice in Figure 4.1b was wrong. More sophisticated models may take into account that $\sigma_{zz}$ is constant across a cone with a given slant. As an example, Figure 4.2a shows force chains in a system of photo-elastic disks, that can highlight regions with internal stress when placed between cross-polarizers. Notice that only some subset of the grains are being compressed. Figure 4.2b shows how force chains can result in stresses orthogonal to $\sigma_{zz}$ that result in friction with the walls, which ultimately supports the grains.

The assumption that friction is maximally utilized can be met to varying degrees depending on details of the experimental set up. In Janssen’s original experiments, he detected the apparent mass by loading a balance and then filling the column until his balance began to move, which meant the grains were just beginning to slip. This meant that he was able to measure near the Coulomb limit. If the grains along the wall don’t reach the Coulomb limit, $m_{app}$ is expected to lie somewhere between Janssen’s theory and $m_{app} = m_{add}$, between the black and navy lines in Figure 4.1c.
4.2 Scaling Down

To perform Janssen’s experiment with ants, we need to scale it down since we are limited by the number of ants in a single colony. Likewise, we will be limited in the loading protocol because we don’t have freedom to move the column of ants to achieve the Coulomb limit. Therefore, we begin our study of Janssen’s effect with passive particles in columns with small ratios of $D/a$, where $a$ is the particle diameter, and use a simple loading protocol in which we add particles in small steps from the top of the column. The intent of these experiments is to form a baseline for our comparison to ants.

4.2.1 Set Up

We first use spherical plastic beads with an average mass and diameter of $m = (113 \pm 1) \text{ mg}$ and $a = (5.94 \pm 0.02) \text{ mm}$, respectively. The distributions for these quantities for a sample of 200 plastic spheres show that the particles are quite monodisperse, see Figure 4.3a-b. The beads are dropped into open glass and plastic cylinders with a diameters, $D$, that we vary from $D/a = 29.0$ down to $D/a = 3.7$. We have measured the coefficient of friction between the beads and the plastic or glass cylinders by placing a tripod of particles in a horizontal cylinder and gently tilting the cylinder until the tripod first slips, as shown in Figure 4.3c. The tripod configuration prevents the beads from rolling down the tube, so that we can measure the sliding coefficient of friction. The angle at which the first slip occurs, $\theta_{slip}$ is related to the coefficient of static friction, $\mu_s = \tan \theta_{slip}$. We find $\mu_s = 0.31 \pm 0.01$ in both the glass and plastic cylinders, so we conclude that the glass and plastic are equivalent for the sake of these experiments and won’t distinguish between experiments that use glass or plastic cylinders, from this point on.
Figure 4.3: a) The probability distribution function (PDF) of the single particle mass, $m_0$, of the plastic beads we use in our passive Janssen experiments, measured from $N = 200$ beads. b) The PDF of the diameter, $a$, of the plastic beads. c) To measure the coefficient of friction between the beads and the cylinders, we tilt a column containing a tripod (shown in the inset) of particles until it first slips and use the angle of the cylinder with respect to the horizontal to calculate the coefficient of friction. Inset: a top down view of the tripod of particles we use to make these measurements.

We use two slightly different experimental set ups to keep our scales out of contact with the cylinder, depending on $D$. For thin enough cylinders, we use a clamp to hold the cylinder stationary with its lower edge slightly above a scale’s surface, as shown in Figures 4.4a-b. However, for larger cylinders, we find the force exerted by the grains on the cylinder can overcome the clamp, so that the cylinder slips slightly downwards. When this happens, the force of friction between the grains and the cylinder pushes the grains downwards, so that the scale begins to support some of the weight of the cylinder. In these cases, $m_{app}$ never saturates and instead increases linearly in proportion to $m_{add}$.

To avoid this effect in larger cylinders, we use a pair of wooden beams that are not in contact with the scale to support the cylinders and mount a platform inside the cylinder that serves to bring the particles into contact with the scale, as shown in Figure 4.4c.
Figure 4.4: a) An experiment using a thin cylinder and a clamp that holds the cylinder away from the scale. b) A closer look at the base of the column showing that the walls of the cylinder are not in contact with the scale. c) An alternate set up used for large diameter tubes. Wooden beams support the weight of the open cylinder, and a platform inside the cylinder supports the beads and puts them in contact with a scale. d) Three trials in which the mass was measured by the scale, \( m_{\text{app}} \), for \( D/a = 29.0 \) using clamps to hold the cylinder off the scale. The cylinder is not held tightly enough and eventually it slips so that some of its mass is supported by the scale. The \( m_{\text{app}} \) never saturates, regardless of \( m_{\text{add}} \). e) Three trials with \( D/a = 29.0 \) using wooden beams to support the cylinder. In all three trials, \( m_{\text{app}} \) eventually saturates as expected. The black line shows the results of Equation 4.5 with \( M \) set as the average maximum \( m_{\text{app}} \) from the three trials.

In either set up, we add the beads incrementally by dropping beads from the top end of the cylinder, using few enough at a time that the average height of the free surface never changes by more than \( 2a \) in a single step and then wait two to five minutes for the system to approach mechanical equilibrium before we measure the apparent mass and add another increment.

Figures 4.4d-e show results from trials using a column with \( D/a = 29.0 \), the largest ratio we used. In Figure 4.4d, we attempted to clamp the column, only to have it slip at some point in the experiments so that \( m_{\text{app}} \) never saturated. Notice that the column slips later in the trial represented by the red points, but eventually the scale supports a set fraction of any additional added mass.
The trials in Figure 4.4e were all performed using wooden beams to support the column and the apparent mass saturates as expected.

4.2.2 Overshoot

The results in Figure 4.4e can be directly compared to the narrowest square columns used by Janssen, $L/a \approx 29$. For this moderate cylinder diameter, our filling protocol only leads to Janssen-like behavior, instead of Janssen behavior, because it doesn’t guarantee that the grains at the boundary find themselves at the Coulomb limit. Still the apparent mass follows the fluid line for low added masses and eventually approaches a saturation mass asymptotically. The black line in Figure 4.4e is Janssen’s theory, Equation 4.5, using the average largest $m_{app}$. Using our measured values of $\mu_s$ and $D$, the $M$ we measure corresponds to $k = 0.60 \pm .02$, in general agreement with Janssen’s original results for corn, which found $k = 0.67 \pm .05$. For intermediate $m_{add}$, our measurements of $m_{app}$ are higher than what others have measured with filling protocols that guaranteed the maximum friction between the grains and the cylinder [99, 103, 104], as expected.

However, as we begin to examine experiments with smaller $D/a$, we discover a surprising new effect. Consider for example, Figure 4.5a, which shows the apparent mass for a column with $D/a = 8.4$. At low $m_{add}$, $m_{app}$ still follows fluid behavior, but then there exists a range of $m_{add}$ for which $m_{app} > m_{add}$, implying that the friction force between the grains and the walls is actually pushing the grains downwards, compressing them, before $m_{app}$ eventually saturates as expected.
Figure 4.5: a) An overshoot in which $m_{\text{app}} > m_{\text{add}}$ is visible in the results for cylinders with $D/a = 8.2$. b) The overshoot for $D/a = 3.7$ varies from experiment to experiment. c) A set up in which the base of the column is supported by a wooden beam and the cylinder is in contact with the scale. d) $\Delta m_c$ for the column (cyan) and the corresponding calculated apparent mass of the grains, $m_{\text{app}}$ (red) for $D/a = 3.7$. Experiments directly measuring $m_{\text{app}}$ are shown as empty points. e) Height of the overshoot as a function of $D/a$. The line shows $D^{-1}$.

This effect becomes more apparent when we decrease $D/a$ further, as seen for $D/a = 3.7$ in Figure 4.5b. In this case, there are few enough grains in the cylinder that the results vary between independent trials, depending on the exact structure of the force chains formed, but the “overshoot” is always apparent.

We can verify that the frictional forces in the system are compressing the grains during the overshoot by measuring the apparent mass of the open cylinder, $m_{\text{app,c}}$, as a function of the added mass. We do this by modifying the set up so that the walls of the cylinder come into contact with the scale via two outer blocks, while the platform inside the cylinder that supports the beads is kept off the scale by a beam that runs to stages set on the counter-top, as shown in Figure 4.5c. Then we calculate $\Delta m_{\text{app,c}} = m_{\text{app,c}} - m_c(0)$, where $m_c(0)$ is the mass of the empty cylinder and
supporting blocks, which shows how much of $m_{add}$ the cylinder is supporting.

Our measurements, represented by the cyan points in Figure 4.5d, show that $\Delta m_{app,c}$ is negative for intermediate $m_{add}$ so that the cylinder appears lighter during the overshoot before it begins to support all of the additional added mass, as expected in the saturation regime of the Janssen model. The decrease in $m_{app,c}$ for intermediate $m_{add}$ must be due to the same frictional forces that compress the grains for that regime. Indeed, if we consider that $m_{add} = m_{app} + \Delta m_{app,c}$ because all of the added mass must be supported either by the scale or by the cylinder, we can calculate the corresponding values of $m_{app}$ for our granular column, which are shown by the red points in Figure 4.5d. Comparing this calculated value of $m_{app}$ to the directly measured values (black empty points) shows qualitative agreement, demonstrating conclusively that the overshoot is not an artifact.

We believe this overshoot has been missed in previous studies because it is only visible for small $D/a$. In fact, across our experiments we find that the magnitude of the overshoot, which we define as $\left(\max\left(\frac{m_{app}}{m_{add}}\right) - 1\right)$ scales as $1/D$; see Figure 4.5e.

To unravel the physics of the overshoot, we have collaborated with Shivam Mahajan and Massimo Pica Ciamarra, who performed computer simulations [105]. Their simulations verified the existence of compressive force chains that are prevalent only for a narrow range of depths from the free surface. Figure 4.6a shows the strongest 5% of frictional forces in their simulated three dimensional column with $D/a = 8$ and a low filling height, projected into two dimensions. It is clear here that the strongest force chains angle from the walls towards the base of the column. These forces push the cylinder upwards, and the reaction pushes the grains down, compressing them. Figure 4.6b shows the frictional forces near the base in the same simulation after many more particles were added, raising the free surface out of the displayed region. Now many of the force chains angle upwards from the wall, supporting the grains against further compression by their neighbors above them. We can see that the effect of friction changes with filling height. Near the free surface, friction compresses the grains, but at greater depths, friction supports the grains.

The origin of the compressive frictional forces can be understood by considering secondary rearrangements near the free surface. Whenever a new grain is added, it slightly rearranges its
neighbors, which may result in the upward motion of grains near the walls. Due to friction however, the wall pushes back, compressing the grains downwards. To illustrate the presence of these rearrangements, consider the addition of the gray particle into a 2D system of frictionless disks; see Figure 4.6c. When the gray particle is added, it displaces the blue striped particle to the left which in turn forces the surface particle on the left upwards. In the presence of friction, this rearrangement is prevented only if the wall exerts a downwards force on the striped particle, increasing $m_{app}$ relative to $m_{add}$. As more particles are added on top, they can increase the strength of these force chains until they have a significant effect on the apparent mass.

Figure 4.6: a) A projection of the force chains in Mahajan et al.’s simulations into 2D showing the strongest 5% of the force chains present in the simulation. The force chains generally point downwards from the wall. Notice the filling height is low compared to the radius of the cylinder. b) Force chains later in the simulation, after enough particles have been added to fill the cylinder to more than double the height presented. Now the force chains are mostly supportive. c) An example in 2D of a secondary rearrangement that results in compressive frictional forces. d) A fit of Mahajan et al.’s theory against our experimental data for $D/a = 8.3$. This figure is modified from Reference [105].

Importantly though, after even more grains are added, the compressive forces due to gravity begin to dominate and the frictional forces change direction to help support the grains, causing $m_{app}$ to saturate in a Janssen-like manner.

Through collaboration with the group of Pica Ciamarra, we were able to use these insights to
modify Janssen’s model and replicate the overshoot by allowing the frictional forces to depend explicitly on the depth, \( z \), instead of only through its dependence on \( \sigma_{zz} \). In parallel to equation 4.4, our proposed equation for mechanical equilibrium on a thin horizontal slice is

\[
0 = \rho g \frac{\pi D^2}{4} dz - d\sigma_{zz} \frac{\pi D^2}{4} - \pi D \tau(z) dz
\] (4.6)

so that the stress due to friction is now \( \tau(z) \) instead of \( \mu_s k \sigma_{zz} \).

Based on simulations and experimental evidence, \( \tau(z) = 0 \) is chosen for depths less than \( h_{\text{start}} \), which marks the beginning of the overshoot, which is reasonable by inspection of Figure 4.5 a-b and consistent with Janssen’s model, in which friction effects are weak in shallow columns. To simplify the model, we then choose \( \tau(z) = C \), with \( C \) a positive constant so that friction compresses the grains, up until \( h_{\text{max}} \), the height that corresponds to \( \max \left( \frac{m_{\text{app}}}{m_{\text{add}}} \right) \), and finally we let \( \tau(z) \) approach \( \mu_s k \sigma_{zz} \) asymptotically in a Janssen-like manner for greater depths:

\[
\tau(z) = \begin{cases} 
0 & z < h_{\text{start}} \\
-C & h_{\text{start}} \leq z < h_{\text{max}} \\
\mu_s k \sigma_{zz} \left( e^{-\lambda(z-h_{\text{max}})} - 1 \right) & h_{\text{max}} \leq z,
\end{cases}
\] (4.7)

where \( \lambda \) is a fitting parameter that controls how quickly the model returns to the Janssen-limit with increasing \( z \). These choices for \( \tau(z) \) allow the model to reproduce the overshoot seen in experiments, as shown in Figure 4.6d. In this figure, the purple points are experimental measurements in a cylinder with \( D/a = 8.3 \), the blue line is Janssen’s model, and the black line is a fit to equation 4.6 with our assumptions about \( \tau(z) \).

Because of the simple choices of \( \tau(z) \) before and during the overshoot, we can solve Equation 4.6 for the maximum magnitude of the overshoot. Rearrangement of 4.6 yields

\[
d\sigma_{zz} = \left( \frac{4}{D} \tau(z) \right) - \rho g
\]
Integration from 0 to \( h_{\text{max}} \) and substitution of our choices in 4.7 then shows

\[
\int_0^{\sigma_{zz}(h_{\text{max}})} d\sigma'_{zz} = \int_{h_{\text{start}}}^{h_{\text{max}}} \frac{4C}{D} dz - \int_0^{h_{\text{max}}} \rho gdz
\]

\[
\sigma_{zz}(h_{\text{max}}) = \frac{4C}{D} (h_{\text{max}} - h_{\text{start}}) - \rho gh_{\text{max}}
\]

Using, as before, that \( m_{\text{app}} = -\sigma_{zz}(h)\pi D^2/4 \), \( m_{\text{add}} = \rho gh\pi D^2/4 \), and our definition of \( h_{\text{max}} \), we rearrange to find

\[
\max \left( \frac{m_{\text{app}}}{m_{\text{add}}} \right) - 1 = \frac{4C}{\rho g D} \left( \frac{h_{\text{start}}}{h_{\text{max}}} - 1 \right). \tag{4.8}
\]

Based on our earlier observation that \( \max \left( \frac{m_{\text{app}}}{m_{\text{add}}} \right) - 1 \propto 1/D \) (see Figure 4.5e), this model predicts that \( h_{\text{max}}/h_{\text{start}} \) is independent of \( D \). This seems reasonable because most rearrangements we expect to lead to the overshoot should be localized near the walls. Moreover, we can support this claim experimentally. It is hard to measure \( h_{\text{start}} \) experimentally when the overshoot is small, but we can easily detect \( h_{\text{max}} \). We find that \( h_{\text{max}} \) is independent of \( D/a \) in experiments (see Figure 4.7), lending further credence to our model.

![Figure 4.7: The height of the column during the maximum overshoot of \( m_{\text{app}} \) in our experiments with spherical plastic beads for varying \( D/a \).](image)

### 4.2.3 Time Dependence

In Section 4.2.1, we mentioned that we wait \( 2 - 5 \) min between adding more mass and measuring \( m_{\text{app}} \) so that the grains approach mechanical equilibrium. In actuality, our grains never quite seem
to reach mechanical equilibrium and we always detect a time dependence in measurements of $m_{\text{app}}$. In the short term, the time dependence leads to a decrease in the apparent mass, regardless of the ratio $D/a$. Figure 4.8a shows $m_{\text{app}}$ (red) and $m_{\text{add}}$ (blue) as a function of time during a trial in a cylinder with $D/a = 29.0$, as $m_{\text{app}}$ begins to saturate. The apparent mass jumps whenever more mass is added to the cylinder but then decays smoothly.

Figure 4.8: a) Time dependence of $m_{\text{app}}$ in an experiment with $D/a = 29.0$. The red curve is $m_{\text{app}}$, and the blue curve is $m_{\text{add}}$. b) The same time dependence in an experiment with $D/a = 3.7$ during the beginning of the overshoot. c) Time dependence of the mass of the cylinder relative to its initial mass, $\Delta m_c$.

The same process seems to be at play in the smaller cylinders even during the overshoot. Figure 4.8b shows the apparent mass as a function of time while it first departs from fluid behavior into the overshoot.

Apparently micro-rearrangements slowly increase friction between the grains and the walls of the cylinder so that the cylinder supports more of the mass over time. We also pick up the opposite of this effect when measuring the apparent mass of the column, as shown in our time dependent measurement of $\Delta m_{\text{app},c}$ in Figure 4.8c. Note that, in this case, $\Delta m_{\text{app},c}$ increases with the addition of grains. We believe this demonstrates that the time dependence is not an artifact of our measurements.

When we wait hours or even days between additions of mass, we find that the time dependence is not monotonically decreasing; instead it changes over the course of hours. We believe this is due to micro-rearrangements that occur because many of the grains remain out of mechanical equilibrium for long periods of time. This effect is present even in trials for which we used glass.
beads instead of plastic beads to remove concerns about the effect of humidity. This may be an interesting effect that deserves more work in the future.

4.2.4 Stability of the Overshoot

The time dependence of $m_{\text{app}}$ brought about by our loading protocol might suggest that the overshoot could be unstable against perturbation, but we find that this is not the case. Tapping or even physically shaking the cylinder does not allow the column to relax enough to remove the overshoot. Figure 4.9 shows a trial in which glass beads in a cylinder with $D/a = 14.8$ were vigorously shaken back and forth immediately after each addition of grains. Interestingly, $m_{\text{app}}$ dips to values near $m_{\text{add}}$ while the column is being shaken, which suggests that we are temporarily negating the effects of friction. However, in the overshoot regime, $m_{\text{app}}$ always returns to a level consistent with an overshoot instead of relaxing to approach Janssen’s theory, showing that the overshoot is stable.

![Figure 4.9](image-url)

Figure 4.9: The $m_{\text{add}}$ (navy) and $m_{\text{app}}$ (red) of glass beads in a column with $D/a = 14.8$. Every time new beads were added, the column was shaken. After each shaking event, $m_{\text{app}}$ returns to a level consistent with an overshoot.

4.2.5 Other Passive Particles

So far, all of the experiments we have discussed have used monodisperse spherical beads. We now check to see if the overshoot persists and if so, whether it is still well described by our model if the grains have more complicated shapes and are polydisperse. To do this we repeat the experiments with dried garbanzo and mung beans. The garbanzo beans have average widths and standard
deviations of \((8.2 \pm 0.5) \text{ mm}\) by \((5.8 \pm 0.4) \text{ mm}\) by \((5.6 \pm 0.4) \text{ mm}\), along their three principle axes, as shown in Figure 4.10a, and a mass of \((160 \pm 30) \text{ mg}\), as shown in Figure 4.10b. The mung beans have roughly cylindrical symmetry around their long axis. They measure \((5.6 \pm 0.6) \text{ mm}\) along their long axis and \((4.3 \pm 0.3) \text{ mm}\) along their short axis, with an average mass and standard deviation of \((81 \pm 15) \text{ mm}\) (Figures 4.10c-d).

![Figure 4.10](image)

Figure 4.10: a) The measured size distributions of our garbanzo beans along their three principle axes. b) The mass distribution of the our garbanzo beans. c) The size distribution of our mung beans along two principle axes. d) The mass distribution of our mung beans.

We find that these types of beans produce overshoots when loaded into our cylinders with small \(D/a\). For example, Figure 4.11a shows the results of three experiments with garbanzo beans (tan points) and three experiments with mung beans (green points). It seems that both types of beans exhibit similar secondary rearrangements near the walls, suggesting that the overshoot may be a general behavior for granular systems.
Figure 4.11: a) An example of an overshoot in three trials using mung beans in a column with $D/a = 5.1$ (green points) and an overshoot in three trials using garbanzo beans with $D/a = 5.1$ (tan points). b) The magnitude of the overshoot as a function of $D/a$ for three types of particles: spherical plastic beads (navy), mung beans (green), and garbanzo beans (tan). The points represent the average, and the error bars represent the maximum and minimum values measured in different trials. c) The height at which we recorded the peak magnitude of the overshoot for all three types of particles.

However, when we compare the results using these beans between experiments with varying $D/a$, we find that $\max\left(\frac{m_{\text{app}}}{m_{\text{add}}}\right) - 1$ is no longer proportional to $D^{-1}$, as it was for the spherical beads. Figure 4.11b shows that $\max\left(\frac{m_{\text{app}}}{m_{\text{add}}}\right) - 1$ seems to decrease faster than $D^{-1}$ (black line) for both types of beans. This implies that we are missing some of the essential physics for these more complicated grains. One hint on where to look for modifications to our model can be found by measuring $h_{\text{max}}$ for experiments with the beans, as shown in Figure 4.11F. Unlike the experiments with the spherical plastic beads, we find that $h_{\text{max}}$ decreases precipitously when we increase $D$. This may mean that compressive force chains form in a narrower range of depths when $D/a$ is large. The result is that Equation 4.8 no longer leads to $\max\left(\frac{m_{\text{app}}}{m_{\text{add}}}\right) - 1 \propto D^{-1}$ because $h_{\text{max}}/h_{\text{start}}$ is no longer constant.

We conjecture that the narrower depths that support compressive force chains may have to do with the specifics of packing non-spherical particles. Ellipsoidal grains have been shown to pack more tightly than spherical grains, and interestingly, it seems that random packing may be denser than ordered packing for ellipsoidal grains [106]. It has also been proven that the packing fraction of ellipsoidal particles increase with container size because particles near the walls tend to be more ordered [107]. These extra effects of container size may be behind the steeper $D/a$ dependence of the magnitude of the overshoot in Figure 4.11a; the larger packing fractions expected for non-
spherical grains compared to spherical grains might be the cause of the less prominent overshoot for increasing $D/\sigma$ in these systems. We hope that future simulations will be able to shed light on this.

4.3 Fire Ants in Cylindrical Containers

We now turn to the behavior of ants in the same cylinders. To perform these experiments, we collect the largest obtainable volume of sterile female ants from a single colony and then divide these ants into 25 roughly equally massed “chunks.” We coat 25 cups in talcum powder, weigh the cups, then add a chunk of ants to each cup and weigh the cups again to determine the mass of ants we have added to each cup.

In our early experiments, we had trouble with ants squeezing themselves into gaps between the cylinder and the scale or between the cylinder and a platform inside it, which brought the cylinder into contact with the scale. To avoid this, we slightly modified the cylinder set up to include a thin polyethylene sheet over the bottom opening of a $D = (22.2 \pm 0.05) \text{ mm}$ cylinder, which corresponds to $D/L \approx 6$, where $L$ is the average length of an ant. The sheet is sealed over the opening with tape, but we leave enough slack that the sheet can rest on the scale for an area equal to the cross-section of the cylinder without tension, as shown in Figures 4.12a-b.
We slick the walls of the cylinder with talcum powder to greatly reduce the ants’ ability to climb on it and clamp it into place with its bottom edge about $0.5\ cm$ away from the scale so that the polyethylene sheet is resting on the scale. We then add the ants chunk by chunk, from the top of the cylinder every two minutes, so that the experiment lasts for $50\ min$. Each chunk is chilled for two minutes in a refrigerator immediately prior to its addition to make the ants easier to handle. Figure 4.12c shows the ants in the experimental set up immediately after the end of an experiment, demonstrating that we have enough ants to achieve large $h/a$. While we add the ants, we continuously sample the apparent mass at $3.75\ fps$, so that we record $m_{app}$ as a function of time for each trial. A sample of this data can be seen in Figure 4.12d, which shows the cumulative mass of all the added chunks, $m_{add}$, in blue and $m_{app}$ in red.
Figures 4.12e-f show $m_{app}$ measured 110 s after each chunk is added as function of $m_{add}$. In Figure 4.12e, the color of the points for each trial represents the amount of time the ants used have been in captivity before the trial, showing that we don’t see any kind of pattern. Figure 4.12f shows the same data, but this time the trials that are colored the same use ants from the same colony. Experiments using the same colony are, in fact, quite reproducible, though the results can slightly vary between colonies.

Aside from the reproducibility, the most striking feature of these plots is that $m_{app}$ saturates, indicating the presence of supportive force chains in the ants and a strong deviation from fluid behavior. Apparently, even though these force chains must be continually broken and reformed as the ants rearrange and crawl past each other, the ants never relax enough to behave like a fluid.

Another noteworthy feature of this plot is the lack of the large overshoot we would expect for passive grains with a similar $D/a$ ratio. Evidently, the ants resist the kind of secondary rearrangements that we saw in the simulations of passive particles, perhaps due to their activity or the strength of the connections they form with their legs and mandibles.

### 4.3.1 Time Dependence

Like the passive plastic beads, we find that the ants show a strong time dependence in their $m_{app}$. However, instead of relaxing into a state in which friction supports more of the ants’ mass, as was the case for the passive beads, the ants relax to support more of their mass against the base. This contrast can be seen clearly by considering Figures 4.13a-c. Figure 4.13a and Figure 4.13b show the net change in $m_{app}$ over the 60 s immediately after the addition of new mass for passive spherical beads in experiments with $D/a = 29.0$ and $D/a = 3.7$, respectively. In both cases, $m_{app}$ decreases after each addition, and the effects are strongest when at intermediate values of $m_{add}$. This indicates that more supportive force chains are forming or that existing chains are becoming stronger.
Figure 4.13: a) The change in $m_{\text{app}}$ during the 60 s immediately following the addition of more particles for an experiment with passive spherical grains and $D/a = 29.0$. The apparent mass decays over time. b) The same measurement for an experiment with passive spherical grains and $D/a = 3.7$. c) The change in $m_{\text{app}}$ for the ants over 60 s after adding more mass, after $m_{\text{app}}$ was averaged over 1.33 s. d) The apparent mass of the ants when 17.5 g of ants are added all at one time at $t = 0$. The apparent mass continues to increase until it approaches the apparent mass for the same mass added in an experiment with incremental chunks.

To calculate the equivalent measurement for the ants, $\Delta m_{\text{app}}$, we first smooth our measurements of $m_{\text{app}}$ with a moving average over 1.33 s (5 frames) to reduce the impact of short-time fluctuations. Then we calculate $\Delta m_{\text{app}}$ as the change in smoothed measurements over the 60 s after a chunk of ants is added. An typical example of the result for a single experiment is shown in 4.13c. The $m_{\text{app}}$ of the ants also shows the strongest time dependence at intermediate values of $m_{\text{add}}$, but the ants almost always increase their $m_{\text{app}}$ over time; this indicates that, in the ants, some of the force chains disappear over time, a fact that could be related to the ants’ activity.

This effect can be seen clearly if we add a large quantity of ants at once and observe how the apparent mass changes with time, as shown in Figure 4.13d. In this case, 17.5 g ants (blue line) are added to the cylinder at $t = 0$. The red line shows the expected value of $m_{\text{app}}$, which corresponds to a measurement of $m_{\text{app}}$ taken 110 s after the chunk was added to make $m_{\text{add}} = 17.5$ g in an
experiment with incremental chunks that used the same ants about an hour before the experiment in Figure 4.13d. The green points show the $m_{\text{app}}$ detected by the scale at the bottom for 2000 s after the 17.5 grams were added. Over the first $\sim 900$ s, $m_{\text{app}}$ increases by more than 25% and then plateaus at a level consistent with the expected $m_{\text{app}}$.

We believe this effect is due to the ants’ hesitancy to support shear stresses. While the ants can temporarily support large shear stresses because of the links they make between each other with their legs and mandibles, it is theorized that they rearrange faster in areas under stresses so that they continually relax [81]. This would result in initially strong force chains gradually weakening, so that individual ants support less stress.

Either the ants are always able to support some small shear or the ants coordinate to reduce $m_{\text{app}}$ for high $m_{\text{add}}$ by forming purposely forming more force chains in order to prevent the ants at the base from being crushed.

4.3.2 Fluctuations

On top of the time dependent increase in $m_{\text{app}}$, the constant motion of the ants leads to short term fluctuations in $m_{\text{app}}$ (red line in Figure 4.12d). To isolate these fluctuations, we calculate the running average of $m_{\text{app}}$ using a window of 13.3 s, which is significantly longer than the timescale of the fluctuations. We then take the difference between $m_{\text{app}}$ and its running average to be the measure of the fluctuations

$$\delta m_{\text{app}} = m_{\text{app}} - \langle m_{\text{app}} \rangle$$

.
Figure 4.14: a) The fluctuations in $m_{\text{app}}$ as a function of time. Notice that the fluctuations become larger as more ants are added to the column. b) The root-mean-square fluctuations of the apparent mass during each interval between adding chunks of ants across three trials using the same colony. c) Averaging across all three experiments, we measure the average fluctuations for various $m_{\text{app}}$. The black line represents $m_{\text{app}}^{0.5}$.

Figure 4.14b shows $\delta m_{\text{app}}$ as a function of time for the first 1000 s of an experiment. At early times, additions of mass are clearly visible as successive positive and negative peaks because of our averaging window, and the fluctuations appear to grow stronger over time as more mass is added. We measure the root mean square (RMS) fluctuations, $\sqrt{\langle \delta m_{\text{app}}^2 \rangle}$, for each added mass in our trials to see how the fluctuations scale. We are careful to choose our averaging window to span from 10 seconds after a chunk of ants is added until 10 s before the next chunk to avoid including the artificial peaks in $\delta m_{\text{app}}$ that result from the averaging window for $\langle m_{\text{app}} \rangle$ and the sudden additions of mass.

Figures 4.14b show the RMS fluctuations for each step during three trials that each used the same colony and that are represented by maroon points in Figure 4.12b. As expected, the size of the fluctuations increases with $m_{\text{app}}$. To improve our statistics we next bin the measurements from all three experiments into to $2 \, g$ bins by $m_{\text{app}}$ and average them. The log-log plot in Figure 4.14c show the averages for each bin as navy points which should be compared to the black line line that
represents $\sqrt{m_{\text{app}}}$. We can see that the size of the fluctuations scale as $\sqrt{\langle \delta m_{\text{app}}^2 \rangle} \propto \sqrt{m_{\text{app}}}$, in keeping with the Law of Large Numbers (LLN).

This reminds us that the LLN does not only apply to equilibrium systems. Recall from Section 1.2 that the LLN governs sums of independent random variables with finite variance and doesn’t apply to sums in which all of the variables are correlated. Because $\sqrt{\langle \delta m_{\text{app}}^2 \rangle}$ scales as predicted by the LLN, we conclude that the varying friction forces that together support the ants against compression act independently and are not a system-wide organized effect. The force chains that form do not seem correlated.

We next measure the autocorrelation function of $\delta m_{\text{app}}$ for each of the 2 min intervals between chunks:

$$R(\Delta t) = \frac{\langle \delta m(t + \Delta t)m(t) \rangle_t}{\langle \delta m(t)^2 \rangle_t}.$$ 

We bin our measurements of $R(\Delta t)$ for all of the intervals for which $m_{\text{app}}$ fell into different windows (specified in the legend) to create the plot shown in Figure 4.15a. Unlike the RMS magnitude of the fluctuations, the autocorrelation does not change with increasing $m_{\text{app}}$, suggesting again that the ants are not changing their behavior as more mass is added to intentionally avoid crushing the ants on the bottom.

Observing this, we average $R(\Delta t)$ over all of the 2 min intervals to produce the black points shown in Figure 4.15a. Fitting the first few points, as shown by the black dotted line, yields a timescale of about 2.5 s, which we speculate measures the average amount of time that two neighboring ants remain linked forming a segment of a force chain.

Our measurements also demonstrate an anti correlation in $\delta m_{\text{app}}$ for $\Delta t$ on the order of seconds, which indicates that the ants may overcompensate when they form or break force chains. We speculate that this occurs because the ants are likely to break connections that are especially stressed.
Figure 4.15: Average autocorrelations $\delta m$ for various measurements of $m_{app}$ (colored points). The autocorrelation times do not seem to depend on $m_{app}$. The black points are autocorrelation measurements averaged over all added chunks.

4.4 Conclusions

In this chapter, we compared ants to passive granular particles in Janssen’s experiment, a classic experiment in which grains in a cylinder are supported against further compression by frictional forces with vertical walls that support force chains in the grains. In this experiment, the mass measured by a scale at the bottom of a cylinder eventually saturates, regardless of how much mass is further added to the cylinder. Because we were limited by the size of our ant colonies, we had to scale down Janssen’s experiment in order to achieve depths for which Janssen’s effect should become visible. We began our scaled down experiments with spherical plastic beads and found a new effect in which frictional forces compressed the grains instead of supporting them, resulting in an overshoot; the mass measured by the scale at the bottom was higher than the mass we added to the cylinder. We called this situation an “overshoot” and found that it was stable against tapping or shaking the cylinder. We found that the relative strength of the overshoot scaled inversely to the diameter so that it disappears in larger cylinders. Simulations showed that these compressive frictional forces were due to secondary rearrangements that forced particles near the walls upwards and that these secondary arrangement only occurred at a narrow range of depths near the free surface.

We then tried other non-spherical grains in our narrow columns and found that the overshoot disappeared more quickly with increasing cylinder diameter than it did for spherical particles.
We conjectured that this was due to denser complex packing.

We also measured the change in apparent mass over time and found that the passive particles actually support more of their mass against the walls as time goes on, hinting that there must be continual micro-rearrangements that build new or strengthen existing force chains.

After we completed our study of passive grains in narrow tubes, we turned our attention to experiments with ants in narrow tubes. We found that the apparent mass on the bottom scale did saturate as we added more ants to a cylinder, much like it would for an experiment with passive grains. This is somewhat surprising because we know that ants at these densities can flow in response to an applied shear, so we might expect them to relax and behave more like a fluid instead of supporting themselves via shear stresses at the walls with friction. We found that the apparent mass of the ants on the scale increased over time. This is opposite to the behavior of the passive particles, but it is consistent with the ants’ motion. However, this time dependence plateaus, meaning that the ants cease to flow and continue to support themselves via force chains.

On top of this long timescale relaxation, we also detected strong fluctuations in the apparent mass of the ants on the scale, as individual ants move and make and break connections with each other. We measured these fluctuations and found that they scale the same as equilibrium fluctuations, hinting that the chains that support the mass of ants against compression act independently.

Taken together, our results highlight the rich and often unexplored physics in systems with friction.
CHAPTER 5
SHAKEN POLAR GRAINS

5.1 Introduction

Simulations of equilibrium systems routinely run up against “finite-size” effects, which cause their simulation of thousands, or millions, of particles to behave very differently from the physical systems they model, which usually consist of on the order of $10^{23}$ particles. These effects are especially important near phase transitions. For instance, simulations of the Ising spin model often do not explicitly show a discontinuous phase transition from complete disorder to sudden symmetry breaking. Instead, the measured magnetic moment switches back and forth from $M$ to $-M$, even at temperatures somewhat above the expected phase transition. The root cause is that the scale of the fluctuations in the simulations approaches the dimensions of the simulation [108].

For equilibrium systems, and indeed for phase transitions in active matter, the goal is typically to reduce these effects or use techniques to compensate for them.

However, in active systems, finite size effects may be more than just a limitation. Many active systems, including the systems of ants that we have discussed so far in this thesis, consist of only thousands or tens of thousands of particles. With this in mind, we will try to examine the collective-motion phase transition in a system of just hundreds or thousands of shaken polar grains. We will check whether the phase transition appears continuous in these small systems and whether they exhibit the phase separation discussed in Section 1.3.2.

Then we will use the shaken polar grains to test some of the predictions from simulations of passive polymers embedded in active baths.
5.2 The System

In the previous chapters we have studied the physics that results from the variable activity of ants. The social pair interactions of the ants and their activity cycles made ants an interesting system with which to test the universal claims of active matter theory. However, these same complications can make exact tests of active matter theory difficult. In this chapter, we use one of the most well-known active matter systems, shaken polar grains, to test some of the theoretical predictions for ideal active systems and also compare the behavior of the disks to the ants.

The grains and shaker set up we use are the same used to produce the results on giant number fluctuations in Section 1.2.2 [19, 46, 109, 110]. Each grain is a circular copper-beryllium disk with diameter $d_0 = (4.00 \pm 0.05) \text{ mm}$, supported by two asymmetric legs, as shown in Figure 5.1a. The back leg is a wide rubber skate, while the front leg is a made of copper-beryllium and comes near to a point, as shown in the bottom view of a grain in Figure 5.1b. Both legs are the same length, so the top of each grain is a level surface $(2.0 \pm 0.1) \text{ mm}$ above the bottom plate. When shaken vertically with a high enough amplitude, the asymmetry between these two legs biases the landing of the disks and they execute directed motion in the direction of their metal leg, towards $\vec{n}$, which is represented by a red arrow in Figures 5.1a-c.
Figure 5.1: a) Side view of one of the polar grains. When shaken the grain self propels in the direction \( \vec{n} \). b) Bottom view of the polar grain showing that the back foot is much wider than the front foot. c) An image of the grains at low density after image processing and tracking. The grains have been highlighted with white circles, red arrows indicate the grains’ orientations, and colored lines indicate the trajectories of the particles. Parts of this figure are adapted from References [19, 46].

To shake the disks, we place them between two glass plates with a gap height of \( \sim 3 \text{ mm} \) and drive them with a servo-controlled shaker attached to the bottom plate, whose specifications can be found in Reference [46]. We control the gap height between the glass plates with three screws around the edge of the top plate to make sure that the gap is even across the entire setup. If the gap is too large, the polar disks tend to flip over or stack by slightly overlapping, which frustrates their motion. In contrast, if the gap is too small, the self propulsion of the grains is greatly reduced or stopped.

We create different confining cells for the grains by cutting the geometries out of 2 mm thick crafting foam board and attaching them to the top plate with tape. Once we have confined the particles between the plates and inside our chosen geometry, we turn on the shaker and then adjust the stiffness of four damping springs around the edges of the bottom plate to ensure that the vibrations are uniform across the experimental area. We continue to adjust the springs until the
particles appear evenly distributed to the eye and remain that way for several minutes. The grains tend to vacate areas with stronger vibrations and accumulate in regions with softer vibrations due to motility effects like those discussed in Section 1.5.1.

Subject to suitable vibrations from the bottom plate, the grains exhibit low rotational diffusion and long persistence lengths, as shown in Figure 5.1c. Here the orientation of the particles, $\hat{n}$, are shown with red arrows and the colored lines show the tracked trajectories of the grains, demonstrating their long persistence lengths.

The grains’ orientations experience an angular noise that we quantify by measuring the orientational diffusion constant, $D_\theta$. For small deviations in the orientations, $\Delta \theta$, of the grains

$$\langle (\Delta \theta)^2 \rangle = 2D_\theta t$$

In practice, we fit the first $1/3$ s of $\langle (\Delta \theta)^2 \rangle$ with a line, and use the slope to calculate $D_\theta$, as shown for two examples in Figure 5.2a. Note that the long-time limit, $\langle (\Delta \theta)^2 \rangle$ does not appear diffusive. This is because, in this experiment, the particles are undergoing collective motion around a ring.

![Figure 5.2: a) The mean squared rotation of the particles as a function of time in experiments with $a_0/g = 2.4$ (solid blue curve) and $a_0/g = 2.6$ (solid red curve). The dashed lines are the best fits of a line to the first ten points in each experiment and are used to calculate $D_\theta$. b) Measurements of $D_\theta$ for varying $a_0/g$.](image)

We find that we can vary the orientational diffusion constant by varying the magnitude of the acceleration of the vibrated bottom plate. We keep the frequency of the vibrations constant at $f = 120 \ Hz$ and vary the amplitude, which we measure as the maximum acceleration of the bottom plate, $a_0$. At $f = 120 \ Hz$, a maximum acceleration of $a_0 = 2g$, where $g$ is the acceleration
due to gravity on Earth, corresponds to the plate shifting distances of $\pm 34 \mu m$ each cycle.

Increasing $a_0/g$ typically results in a linear increase in the rotational diffusion constants of the grains, as shown Figure 5.2b. Unfortunately, the relation in Figure 5.2b does not hold quantitatively between separate noise sweeps because the behavior of the grains is extremely sensitive to the alignment of the shaker and the stiffness of the damping springs, which can vary based on humidity, temperature, and the arbitrary adjustments made when setting up the experiment. In fact, many attempts at the noise and density sweeps discussed in this chapter had to be discarded because the alignment of the set-up shifted over the course of the sweep. We therefore measure $D_\theta$ independently for each experiment.

Collisions between the grains tend to bring about alignment, even though the grains have an isotropic shape. The cause of the alignment is a torque that acts on the particles when their velocity $\vec{v}$ is not aligned with orientation of the particle, $\hat{n}$. This torque is in a direction that brings $\hat{n}$ in line with the $\vec{v}$. A collision with another particle sometimes forces the grain to move in a direction that’s different from $\hat{n}$. In this case, we model the effect of torque on the particle to be:

$$\tau_n \frac{d\hat{n}}{dt} = \left( \hat{n} \times \frac{\vec{v}}{|\vec{v}|} \right) \times \hat{n}$$

$$\tau_v \frac{d\vec{v}}{dt} = \hat{n} - \frac{\vec{v}}{|\vec{v}|} + \text{collisions}$$

where $\tau_n$ and $\tau_v$ are timescales that control the rates at which the orientation of the particle changes to match the velocity and how strong of an effect the orientation has on the velocity, respectively.

The result of this torque in a simulation with $\tau_n/\tau_v = 0.5$ is shown in Figure 5.3. This figure shows the trajectories of a particle that is originally oriented towards the right for three different perturbations in its velocity. Here the orientation of the particle is shown in black arrows and the velocity is shown in colored arrows that match the color of the trajectories. These proposed forces have successfully captured the behavior of the vibrated grains in previous simulations [111].

140
5.3 Collective-Motion Phase Transition

In Section 1.3, we discussed the debate between Vicsek et al. and Chaté et al. about the nature of the phase transition. At first, Vicsek et al. maintained that the phase transition was continuous [17, 50], but eventually Chaté et al. were able to show that the transition was discontinuous [24, 49]. The latter argued that simulations using Vicsek’s interaction rules only showed the discontinuous nature of the phase transition for large system sizes, but using a vectorial noise term instead of an angular noise term resulted in a clear discontinuous phase transition even in smaller systems.

Our vibrated polar grains are known to exhibit collective-motion [19], but previous studies have used confining geometries that didn’t allow for long term unidirectional collective-motion, which made it hard to examine the phase transition in detail. We would like to directly check whether the phase transition for our polar disks appears discontinuous by confining our active grains to a geometry conducive to collective-motion and then systematically varying the noise or density. Recall that collective-motion is known to occur at high densities and low noises so that the phase transition can be measured by varying either.
To facilitate steady collective-motion, we confine the polar grains to a ring with an inner diameter of $D_{in} = (180 \pm 2) \text{ mm}$ and an outer diameter of $D_{out} = (260 \pm 2 \text{ mm})$ so that the grains are free to explore a channel $40 \pm 2 \text{ mm}$ wide, as shown in Figures 5.4a-b. After we turn on the shaker, we image the setup from above at a rate of 30 $fps$ and analyze the video to track the grains. In each frame, we locate the grains by using a Hough transform (see Reference [112]), which is an algorithm used to find regular shapes, to search for perfect disks of the appropriate size and then determine the orientations of the disks by searching within each disk to find the dark spot marked on its surface. This position and orientation information allows us to track the grains from frame to frame using Algorithm 1, which we described in Section 2.2. In this case, there are very few errors that need to be considered because the Hough transform is efficient at finding the grains and the frame rate is high enough that the grains take about 10 frames to move a distance $d_0$.

From the tracking, we can calculate the $x$ and $y$ components of the velocity of the grains by taking the difference in their position between two time steps

$$v_{i,x}(t) = \frac{x_i(t + \Delta t/2) - x_i(t - \Delta t/2)}{\Delta t}$$

$$v_{i,y}(t) = \frac{y_i(t + \Delta t/2) - y_i(t - \Delta t/2)}{\Delta t}$$
Ideally, we might like to measure the instantaneous velocity by choosing $\Delta t$ to be only one frame. However, choosing $\Delta t$ too small can cause us to miss the directed motion of the grains and instead pick up short term fluctuations in their position caused by the shaker. To best detect their directed motion, we instead use $\Delta t = 1 \text{ s}$. After measuring $\vec{v}_i(t)$, we project it onto the direction tangential to the ring because we know any collective-motion that we detect must be either clockwise or counter-clockwise around the ring. The projected tangential velocity is

$$v_{i,\theta}(t) = -v_{i,x}(t) \sin(\theta) + v_{i,y} \cos(\theta),$$

where $\theta$ is the angular position of the particle on the ring with respect to the $x$-axis.

Next, we take the average of this tangential velocity over all the grains over the duration of a trial to be the global drift velocity for that trial,

$$u_0 = \langle v_{i,\theta}(t) \rangle_{i,t},$$

which is the average speed of the collective-motion phase if different from zero.

Finally, we normalize the drift velocity by the average speed of the particles,

$$\langle v \rangle = \sqrt{\langle |\vec{v}_i|^2(t) \rangle_{i,t}},$$

to determine an order parameter:

$$\Psi = \left| \frac{u_0}{\langle v \rangle} \right|,$$

which spans from $\Psi = 0$, when the grains’ motion is completely disordered, to $\Psi = 1$, when the grains are completely aligned so that their motion is exactly clockwise or counter-clockwise around the ring.

We begin to look for the phase transition in a density sweep with 500 s long trials, varying the
average area fraction,

$$\phi_0 = \frac{N \pi d_0^2}{\pi (D_{\text{out}}^2 - D_{\text{in}}^2)}$$

while keeping $a_0/g = 2.2$. Measurements of $D_\theta$ decrease with density because collisions can slow down rotational diffusion, but for the purpose of comparison, we report that when $\phi_0 = 0.39$, $D_\theta = 1.2 \text{ s}^{-1}$. The trials must be relatively short because the longer the sweep takes, the more likely it is for the shaker set up to lose alignment so that no further trials can be made in that sweep. The number of grains we used in these trial varies from $N = 1488$, on the high end, down to $N = 278$. We find that $\Psi$ increases smoothly as we increase $\phi_0$, as shown in Figure 5.5a. This could be consistent with a continuous phase transition. However, we do not probe low enough $\phi_0$ to detect the phase transition from $\Psi = 0$ to $\Psi \neq 0$.

We also measure the PDFs of instantaneous measurements of $\Psi$, while it fluctuates in time. Recall that in a discontinuous phase transition, it is typical to detect two peaks in the order parameter very near to the phase transition as the system switches back and forth between the two phases. In our case, we only detect a single peak in each trial. Increasing the density causes the single peak to smoothly move towards greater order, in agreement with Vicsek’s argument for a continuous phase transition; see Figure 5.5b. We emphasize, however, that in these experiments, we do not measure sufficiently close to the phase transition. Hence, any conclusions about its character should be made with care.
Figure 5.5: a) The global order parameter $\Psi$ over a range of densities when the amplitude of the bottom plate’s vibration is $a_0/g = 2.2$. The order parameter decreases smoothly for decreasing $\phi_0$ but never reaches $\Psi = 0$. b) The global order parameter over a range of $D_\theta$ with $\phi_0 = 0.4$. c) The probability distributions of measurements of $\Psi$ during experiments with varying $\phi_0$. d) The PDFs of $\Psi$ during experiments with varying $D_\theta$.

We then conduct a noise sweep. In this case, the average order parameter decreases smoothly as we increase $D_\theta$ while holding the area fraction constant at $\phi_0 = 0.4$, as shown in Figure 5.5c. Importantly, one of the sweeps achieved an average order parameter very near $\Psi = 0$, indicating that we are sampling the region near the phase transition. We also see that the PDFs of $\Psi$ again only contain a single peak and that the peak moves gradually towards lower order as we increase the noise, as shown in Figure 5.5d.

The results of our noise sweep suggest the phase transition from the disordered phase to the collective-motion phase is continuous. In our physical system consisting of a relatively small number of particles, however, the discontinuous phase transition could become widely stretched out due to finite-size effects, and even stretch over all of our measured packing fractions and noise levels. Part of this stretching effect could be a result of our confining walls. It is relatively unlikely for
the grains to completely turn around once they start moving either clockwise or counter-clockwise without running into another grain and experiencing a collision along the wall.

Note that, from a practical point of view, the results of finite size effects are important. In fact, many of the physical systems that we attempt to describe as active matter, such as swarms or herds, only consist of hundreds or thousands of individuals. In these systems also, some collective-motion should be expected due to finite size effects, even when their density is very low and the noise level quite high, despite the discontinuous character of the collective-motion phase transition for very large systems. This stands in contrast to equilibrium systems, in which we often disregard finite size effects because systems contain on the order of $10^{23}$ particles.

5.4 Propagating Fronts

Simulations described in Section 1.3.2 predict traveling polar bands near the phase transition, in which dense bands of particles with high order propagate through a low density disordered background. These bands are examples of phase coexistence between the dense collective-motion phase and the disordered phase. Recall that in equilibrium systems, phase coexistence occurs for discontinuous phase transitions. To check for coexistence in our density and noise sweeps, we create windows in our ring geometry that are $21^\circ$ wide and overlap by $20^\circ$. The exact window size does not qualitatively affect our results, but we want to include enough area to capture meaningful area fractions for each window. For each frame and window, we measure the local area fraction, $\phi(\theta,t)$, using $N(\theta,t)$, the number of particles that are located in the window centered at $\theta$:

$$\phi(\theta,t) = \frac{N(\theta,t)}{A} \frac{\pi d_0^2}{4A},$$

where $A = 16.2 \text{ cm}^2$ is the area of a window. We also calculate local average drift speed:

$$u(\theta,t) = \frac{1}{N(\theta,t)} \sum_j v_{j,\theta},$$

where the sum runs over all of the grains in the bin.
Figure 5.6: a) A space-time plot of the area fraction, $\phi(\theta, t)$, showing dense bands traveling in the $-\hat{\theta}$ direction in a trial with $\phi = 0.39$ and $a_0/g = 2.2$. b) The space-time plot of $u(\theta, t)$ for the same trial. To the eye, the regions with the lowest $u(\theta, t)$ correspond to more rarefied regions in (a).

Plotting these two quantities as functions of space, parameterized by $\theta$, and time reveals that the grains naturally organize into traveling bands, as can be seen in the space-time plots for a trial with $\phi_0 = 0.39$ and $a_0/g = 2.2$ in Figures 5.6a-b. In the space-time plot of $\phi(\theta, t)$ in Figure 5.6a, the traveling waves appear as denser, yellow bands that propagate at a nearly constant speed through the significantly lower density background. The bands are less clear in 5.6b, which shows $u(\theta, t)$, but it is clear that the regions in space-time that have the lowest $u(\theta, t)$ in Figure 5.6b also have some of the lowest $\phi(\theta, t)$ in Figure 5.6a. This separation into dense regions with collective-motion and high density, and regions of low order and low density is reminiscent of the traveling bands seen near the discontinuous phase transition. However, unlike traditional traveling bands in large active systems, our bands continually form and dissipate as they are subject to the large fluctuations that dominate finite systems.

In theoretical work and in some large experimental systems such as those using colloidal
rollers, the traveling bands move at speeds very close to the speed of the directed motion of a particle. This is because the bands are dense enough for the alignment of the particles to overcome their rotational noise and because the disordered background is almost devoid of particles. However, the phase transition for our grains spans across a large range of densities and the waves sometimes reach peak densities that approach \( \phi = 0.82 \), which is the random close packing limit for monodisperse disks [113]. It seems improbable that the waves could travel just as quickly, regardless of their density, so we speculate that the speed of the waves in the grains might vary over the transition.

To measure the typical speed of the waves in each trial, we calculate the dispersion relations for \( \phi(\theta,t) \) using discrete Fourier transforms in space and time:

\[
\Phi(q,\omega) = \sum_{\theta} \sum_{t} \phi(\theta,t) e^{-2\pi(q\theta+\omega t)}
\]

We then normalize \( q \) by the average length of a lap around the ring:

\[
q' = 2\pi \left( \frac{D_{in}+D_{out}}{2} \right)^{-1} q
\]

The resulting dispersion relation for the \( \phi_0 = 0.39 \) trial from Figure 5.6 is shown in Figure 5.7a. Measuring the slope of the peaks, \( c_0 \), in the dispersion relation around \((q,\omega) = (0,0)\), shown by the red line in Figure 5.6a gives us the speed of the dominant long-wavelength fluctuations in our system, which are the traveling bands.

We find that \( c_0 \) does, in fact, change with both \( \phi_0 \) and noise level, as shown in Figure 5.6b and Figure 5.6c, respectively, and is always above \( u_0 \). This suggests that the changes in \( \Psi \) shown in Figure 5.5 are driven not only by the ratio of particles inside the bands compared to outside the bands, but also by the changes in order inside the waves themselves. We believe this may be the result of finite-size effects, as it would not be the case for large systems experiencing a discontinuous phase transition. For large enough systems, the change in \( \Psi \) would be driven only by the fraction of particles in each each phase.
Figure 5.7: a) Dispersion relation for $\phi_0 = 0.39$. The slope of the peaks through $q = 0, \omega = 0$, demonstrated by the red line, corresponds to the typical propagation speed of the traveling bands. b) The average global drift speed, $u_0$, of each trial in the density sweep (blue circles) and the speed of the propagating fronts, $c_0$, in those same trials (red triangles). c) The drift speed and the wave speed during one of the noise sweeps.

Notice, however, that the waves vanish when $u_0$ (and therefore $\Psi$) approaches 0 in the noise sweep in Figure 5.6c, much like they are expected to in the large system limit. Also notice that for low $\phi_0$ in the density sweep, both $u_0$ and $c_0$ increase with increasing $\phi_0$, but that they decrease with $\phi_0$ at high densities. This change in behavior hints that there are two competing mechanisms governing the speed of the waves, which we propose are aligning effects and crowding effects.

We can quantify these two effects by measuring “local equations of state” for our grains. We measure the average alignment, $\langle p(\theta,t) \rangle$, and mean squared velocity, $\langle v^2(\theta,t) \rangle$, of the grains in each of the averaging windows we discussed earlier:

$$p(\theta,t) = \frac{1}{N(\theta,t)} \sum_j \left| -n_{j,x}(t) \sin(\theta_j) + n_{j,y}(t) \cos(\theta_j) \right|$$

and

$$v^2(\theta,t) = \frac{1}{N(\theta,t)} \sum_j v_j^2,$$

where $\vec{n}_j$ is the unit vector defining the particles orientation and the summations run over all of the particles in the given window.

In the interest of obtaining better statistics and a broader range of $\phi(\theta,t)$, we now group together all measurements of $\langle p(\theta,t) \rangle$ and $\langle v^2(\theta,t) \rangle$ from our entire density sweep. Then we bin...
measurements of $\langle p(\theta,t) \rangle$ and $\langle v^2(\theta,t) \rangle$ by $\phi(\theta,t)$ and take their average (blue points) and standard deviation in each bin. This allows us to directly examine the effects of local density on the alignment and the average speed of the grains separately. The result shows that $\langle p \rangle$ smoothly increases with $\phi$ up until $\phi = 0.55$, at which point the alignment of the particles saturates, as shown in Figure 5.8a.

On the other hand, Figure 5.8b shows that $\langle v^2 \rangle$ monotonically decreases as a function of $\phi$, showing that crowding effects are important at all of our measured densities. This explains the local maximum that we saw in $c_0$ in Figure 5.7. For waves with a local density $\phi(\theta,t) < 0.55$, denser waves tend to move more quickly because increased alignment more than compensates for slowing effects of crowding. However, for $\phi(\theta,t) > 0.55$ any increase in density does not further align the particles and crowding effects begin to slow the bands down.
Figure 5.8: a) The average polarity of the grains, \( \langle p \rangle \), as a function of the local area fraction, \( \phi \). The error bars represent one standard deviation from the mean and show the typical spread of the distribution. b) The average squared speed of grains \( \langle v^2 \rangle \) as a function of \( \phi \). The error bars are one standard deviation from the mean. c) A heat map showing our measured values of \( u(\theta,t) \) and \( \phi(\theta,t) \) over the course of the density sweep. d) A zoomed in view of a region in the space-time plot in Figure 5.6a, showing a jam and the resulting band.

To explain the origin and break up of our bands, we now sort all of our measurements of the local drift speed, \( u(\theta,t) \), and \( \phi(\theta,t) \) during the density sweep into 2D bins and then count the number of measurements in each bin. The result is shown in Figure 5.8c, in which the color represents the number of counts in the bin, with red representing the most counts and blue representing the fewest.

Notice that there is a clear peak in the most likely \( u(\theta,t) \) for each \( \phi(\theta,t) \) with a distribution of likely \( u(\theta,t) \) that becomes narrower as \( \phi(\theta,t) \) increases. Notice also that our running average of \( u(\phi) \) never approaches \( \langle u(\phi) \rangle = 0 \) for small \( \phi \), highlighting a stark difference between our small system and larger systems that show traveling bands.

Because of our ring’s finite size, fluctuations dominate the dynamics, causing temporary traffic...
jams, like the one shown in the space-time diagram in Figure 5.8d. The regions marked $A$ through $D$ in this figure correspond to instances of specific densities and speeds that are also marked $A$ through $D$ in Figure 5.8d. In the region $A$ in Figure 5.8d, the fluctuations have driven the local speed relatively low, which has two effects on the regions around it. First, the slowdown leads to a decreased density downstream, and second, it causes a traffic jam behind it as particles start to pile up and slow down because they can’t pass by the disordered region. The pile up finds itself with high density and a relatively low speed, as shown by the points labeled $B$. The longer the grains in the pile up remain very dense, the more collisions they experience and the more strongly they align. When the traffic jam eventually clears, the grains are suddenly released with a higher velocity than the average $u(\phi)$, as shown by the points labeled $C$. These particles form a traveling band that quickly travels through the rarefied region with less resistance than usual so that they can maintain their high speed, as shown by the points labeled $D$. Then, these bands continue to disperse as they become less aligned over time and disappear unless another traffic jam compresses and aligns them again.

In a larger system, these jams and sudden relaxations would quickly lead to long lived traveling bands because the density in the rarefied regions would be low enough to drop the local to $\Psi(\theta, t) = 0$ and the density in the compressed regions would be high enough that the particles in them would be maximally aligned without dispersing. However, in our system, the bands disperse, so they require fluctuations to strengthen them and the waves move at various speeds depending on the typical size and duration of the jam.

At low densities, we see the wave speed increases rapidly with the global area fraction in Figure 5.7b because the increasing the global density increases the values of $\langle u(\phi) \rangle$ that can be reached by fluctuations. Consider Figure 5.9, which shows $\langle u(\phi) \rangle$ plotted with the left axis and probability distribution functions of the local packing fractions for several global packing fractions on the right axis. Here, the blue curve shows the PDF of local packing fractions when the global packing fraction is 0.17. Notice that, at this packing fraction, fluctuations that result in packing fractions higher than the average result in regions that travel faster than the average drift speed. Slightly
increasing the global packing fraction shifts the PDF to the right, which allows access to new local packing fractions with even higher average drift speeds. However, as the global packing fraction increases further, the fluctuations are able to achieve the maximum allowed velocities in the system, so that the speed of the fronts as a function of density plateaus in Figure 5.7 because increasing the global packing fraction no longer gives access to higher local drift velocities. Consider the green and yellow curves in Figure 5.9, which represent the trials with $\phi_0 = 0.39$ and $\phi_0 = 0.53$, respectively. The fronts in both trials have access to the densities that result in the fastest traveling waves, and so the speed of the fronts in the two trials is almost equal.

![Figure 5.9: The average speed, $\langle u(\phi) \rangle$, as a function of local area fraction (left axis) compared to the probability distribution functions for the local area fraction for 4 different trials with global packing fractions (right axis): $\phi_0 = 0.17$ (blue), $\phi_0 = 0.39$ (green), $\phi_0 = 0.53$ (yellow), $\phi_0 = 0.66$ (red). The spread in density shows how trials with global packing fraction below $\phi_0 = 0.5$ can still exhibit traveling waves with moving at the maximum allowed speed.](image)

At very high densities, fluctuations due to jams can no longer increase the order of the grains behind them. Instead, they can only decrease the speed of the grains behind them, as shown by the black arrow in Figure 5.10. The result is that a jam travels backwards and increases in density until it approaches random close packing, skewing the probability distribution for $\phi_0 = 0.66$ to the right in Figure 5.9. Eventually the jam’s position becomes stationary, like the perpetual jam shown in Figure 5.10b, and releases periodic bursts of particles from the downstream ends that result in traveling bands.
Figure 5.10: (a) The same heat map as was shown in Figure 5.8 with an arrow indicating that fluctuations that lead to local packing fraction of $\phi(\theta, t) > 0.6$ tend to self-reinforcing, pushing the density even higher. (b) The stationary jam from the $\phi = .66$ trial in the density sweep. The direction of the collective motion is counter-clockwise.

5.4.1 Conclusions

We have examined the collective-motion phase transition for a system of hundreds of active particles in a ring and found that it appears continuous and proposed that this was due to finite size effects. This mirrors the early mistaken conclusions about the character of the phase transition in simulations that did not include enough particles.

Near the discontinuous phase transition in the large system limit for simulations of the Vicsek model, there are traveling bands that represent coexistence between the collective-motion and disordered states. The speed of the bands is constant, and the background is always disordered. As the density is increased or the noise is decreased, the proportion of the active particles in the bands increases until no disordered regions remain. In our finite system, we observe similar bands with notable differences. The speed of our bands are variable, depending on their density, and the background through which the waves travel is not disordered, just less ordered. Both of these behaviors are consistent with the continuous appearance of our phase transition. The nonlinearity of these waves is reminiscent of the nonlinear waves we saw in columns of ants, but we believe the source of the nonlinearity is different. In the case of the disks, the nonlinearity is directly controlled by a “local equation of state” that relates flow speed to density, while for the ants, we believe the nonlinearity was related to the activation time required to add new ants to the wave.
We find that the traveling bands themselves are created by density fluctuations that temporarily frustrate the motion of the flowing grains. These frustrations lead to packets of high density particles that then propagate at a high speed. This frustration too reminds of the ants. Recall that, in our simulations of the ants, if we reduced the time-steps too far, the waves disappeared because most of the ants were always moving in the same direction.

However, all in all, we believe the traveling bands in this chapter should be compared most closely with the traveling bands discovered by Chaté et. al. because they are also controlled by a “local equation of state” that relates the coarse-grained velocity to the density, so that they are waves in both density and order with constant activity. On the other hand, the waves in columns of ants were waves in three local parameters: density, order and activity.

We have mentioned several times in this chapter that our system is controlled by finite size effects, and so one may conclude that our results are less general than results in the infinite-size limit. However, we believe the continuous nature of the phase transition and the nonlinear waves in these small systems are important because many of the active systems of interest have only hundreds or thousands of individuals, whether they are flocks of birds, bits of tissue, or crowds of people. While we are used to considering the infinite-size limit for equilibrium systems, finite size effects may be especially important in active systems. More work on how these finite size effects scale and eventually disappear with increasing system size could prove interesting.

5.5 Polymers in an Active Bath

Polymer physics has remained an exciting field since the correct diagnosis of natural rubber’s component molecules as “macromolecules” by Hermann Staudinger in 1920 [114, 115]. Since then, a wide variety of synthetic polymers have been created, such as those used in plastics, and more have been identified in biological systems. Polymers are still often studied by physicists because of the importance of entropic effects. Because of the large size of polymer molecules, there are often enough possible configurations for entropic effects to compete with or dominate the energetic interactions between component monomers.
A linear polymer in a good solvent at equilibrium has conformations that correspond to a self-avoiding random walk with a step size equal to the Kuhn length, $b$, which is the typical length between monomers needed for their orientations to be considered independent. The Kuhn length is related to the bending stiffness of the polymer, $\kappa$, and the temperature:

$$b = \frac{2\kappa}{k_B T}.$$ 

In terms of $b$, the average length-scales of the polymer, such as the average end-to-end distance and the average radius gyration, $R_g$, scale as $\langle R \rangle \propto N_b^\nu b$, where $N_b$ is the length of the polymer divided by the Kuhn length, $N_b = L/b$, and $\nu$ is a scaling exponent, called the Flory exponent. For a random walk in 2D, $\nu = 3/4$ [116].

However, not all polymers are suspended in equilibrium baths. For example, in biological systems, polymers can be exposed to baths that are held far from equilibrium, where the usual theory for polymer configurations does not necessarily apply [11]. Consider, for example, the important process of protein folding in cells. Proteins are synthesized by piecing together individual nucleic acids (monomers) into long polypeptide chains (polymers), but the chains are not biologically useful until they have folded into a complicated, but stable, tertiary structure in space that is held together by hydrogen bonds between specific sequences of amino acids in the chain. If two sequences of amino acids are never brought into close enough contact to bond, then certain folding configurations of the polypeptide chain are prohibited, making a knowledge of the possible chain conformations incredibly important when interested in the final tertiary structure of the protein.

Since protein folding occurs in the cytoplasm of the cell, which may be modeled as an active fluid [117], there is interest in studying the behavior of passive polymer chains in an active bath. The predictions are that polymers with $\kappa = 0$ slightly swell when the activity of the active bath is increased [118], while polymers with $\kappa \neq 0$ shrink [11].

However, to date, all of these studies have used simulations. Our goal in this section is to test some of the simulation outcomes, using our polar grains as an active bath.
5.5.1 Set Up

We use stainless steel ball chains as our passive polymers and embed them in baths of either passive or active grains confined to a flower shaped cell, as shown in Figure 5.11a. We have artificially highlighted the outside borders of the cell in white for visibility. The flower shaped cell prevents active particles from accumulating on the walls by returning wall-following particles back to the bulk [46]. We consider our chain equivalent to a flexible polymer in a good solvent because the only forces between monomers from different parts of the chain are hard body interactions that enforce an excluded area. The passive grains we use here have the same size as their active analogues but have only a singular circular foot, as shown in Figure 5.11c-d. Previous studies have shown that instead of directed motion, these passive grains undergo random walks and diffuse without a bias in direction when shaken between glass plates [19], as shown in Figure 5.11e.
Figure 5.11: a) The individual balls in our chain marked with blue to yellow from one end of the chain to the other. This example shows the first image in trial with a polymer with $N = 90$ monomers in an active bath. The flower shaped cell prevents active particles from accumulating on the walls. b) The last image from the same trial, showing that none of our monomers have switched labels. c) A side view of one of our passive grains, showing only a single isotropic foot. d) A bottom view of a passive grain. The inside of the circular foot is concave so only its edges come into contact with the glass plate. e) A sample of the trajectories of some passive particles showing that they undergo Brownian motion. Parts of this figure are reproduced from Reference [19].

We run each trial, with either type of grain, for 10 min, while imaging from above at a rate of 30 fps, and track the position of each of the individual balls in our chain, which act as our monomers. These monomers are located as very bright spots in our images because their spherical shape ensures that they always reflect light back at the camera. We use a slightly modified version of Algorithm 1 to track the balls, but only apply a cost to distance because the monomers are uniform in size and apolar. The algorithm now also checks to make sure the measured distance between adjacent monomers stays within acceptable bounds. When the distance becomes too large, the algorithm asks the user to identify the mistake in a series of images and recalculates the distances for those frames. These mistakes tend to come about when non-adjacent balls briefly come into contact as the chain folds.
With these modifications to Algorithm 1, we can reliably track the monomers over the course of an experiment, as shown in Figures 5.11a-b. Figure 5.11a shows a chain of \( N = 90 \) monomers with the monomers highlighted with a color scale that runs from blue for the first monomer in the chain to yellow for the last monomer in the chain. Figure 5.11b shows the last frame in that trial, 10 min later, showing that the order of the monomers has been maintained and that the tracking algorithm never switched the tracks of two monomers.

5.5.2 Polymer Shrinking

We are interested in checking whether a polymer swells or contracts in an active fluid, relative to a passive fluid. From the tracked positions of the monomers, we can calculate the radius of gyration of our polymers for each time step by first computing the gyration tensor:

\[
S_{xy} = \sum_{i=1}^{N} \Delta x_i \Delta y_i
\]

where \( \Delta x_i \) and \( \Delta y_i \) are the \( x \) and \( y \) positions of a monomer relative to the center of mass of the polymer, respectively, and the index runs over all monomers.

Because the gyration tensor is symmetric, i.e.

\[
S = \begin{pmatrix}
S_{xx} & S_{xy} \\
S_{xy} & S_{yy}
\end{pmatrix}
\]

there is an axis rotation, that is, a basis change, that diagonalizes it [119]. In this case,

\[
S' = RSR^{-1} = \begin{pmatrix}
\lambda_1 & 0 \\
0 & \lambda_2
\end{pmatrix}
\]

where, by convention, \( \lambda_2 > \lambda_1 \) so that \( \lambda_2 \) is the axis that corresponds to the maximum 1-dimensional radius of gyration. Here, \( R \) is the rotation matrix associated to changing basis to the basis of the eigenvectors; it is orthogonal, and thus \( R^{-1} = R^T \).
We define the radius of gyration as

\[ R_g^2 = Tr(S) = \lambda_1 + \lambda_2 \]

This is equivalent to the more familiar definition of the radius of gyration

\[ R_g^2 = \sum_{i=1}^{N} (\Delta x_i^2 + \Delta y_i^2), \]

where \( N \) is the number of monomers in our chain, since the trace of a matrix is invariant under rotations. Recall that the radius of gyration is often used in relation to the moment of inertia:

\[ I = \int_V \rho(r)r^2dV = mR_g^2. \]

In other words, a rotating object has the angular momentum as if all of its mass was concentrated in a particle a distance \( R_g \) away from the axis of rotation.

In Figure 5.12a-b, we show the PDFs of our measured radii of gyration for passive and active baths, respectively, where we scale the radius of gyration by \( N\sigma \) where \( \sigma \) is the average separation between adjacent monomers. With this scaling, the maximum end to end length of the polymer is \( R_{ee}/(N\sigma) = 1 \), which corresponds to a maximum radius of gyration equivalent to that of an infinitesimally thin rod with length \( N\sigma \):

\[ \frac{R_g}{N\sigma} = \left( \int_{-1/2}^{1/2} x^2dx \right)^{1/2} = \sqrt{1/12} \approx 0.29. \]

In a bath of passive particles, the shortest chain with \( N = 32 \) monomers acts like a very stiff polymer, with \( R_g \) near the maximum limit; see the red points in Figure 5.12a. In this case, the chain is almost always fully extended. As we increase the number of monomers, the relative radius of gyration \( R_g/(N\sigma) \) begins to shrink. This is because our polymer with a normalized length \( L/(N\sigma) = 1 \) appears softer because it contains more Kuhn lengths; we are increasing the polymer
length relative to its persistence length.

Figure 5.12: a) The probability distribution functions for $R_g/(N\sigma)$ for three different lengths of chains in a passive bath. The color represents the number of monomers in the chain. In these units, higher $N$ would correspond to lower $\kappa$ in equilibrium systems, which results in lower radii of gyration. b) The probability distribution functions for $R_g/(N\sigma)$ for five different lengths of chains c) $\langle R_g \rangle/\sigma$ for passive (blue) and active (red) baths. The black line represents $R_g^{0.75}$, corresponding to a self avoiding random walk.

When we instead immerse the chain in the active bath, the polymers have a much lower $R_g/(N\sigma)$ for a given length, indicating that to leading order, the chain behaves like it has a decreased Kuhn length in the active bath, making it more flexible. This agrees with the simulation predictions in References [11, 118]. Because the physical stiffness of our chain has not decreased, the increased flexibility in the chain can be taken as a sign of an increased effective thermal energy, $k_B T_{eff}$, of the active bath. As we discussed in Section 1.4.1, high effective temperatures are often one the main effects of activity in low density systems.

Figure 5.12c shows that the average $R_g$ for the chain immersed in the active bath is always shorter in the active bath (red points) compared to the passive bath (navy points). Note that within the contour lengths we are able to study, the chain approximately follows the expected Flory Law. Here, the black line is Flory behavior in 2D, $R_g \propto N^{3/4}$. This is similar to the results found in simulations by Kaiser and Löwen [118]. They predicted that long polymers would continue to scale with $R_g \propto N^{3/4}$ in an active bath. In their model, active particles were able to intrude between monomers, which stiffened the chain, and caused the polymer size to scale faster than $R_g \propto N^{3/4}$ for small polymers. These intrusions are absent in our experiments, so we wouldn’t expect to find this fast scaling in our experiments.
Though it looks like the primary effect of the active bath is to increase an effective temperature, we will see that this simple explanation breaks down when we examine the typical shapes the chains adopt in the two baths.

5.5.3 Hairpins and Spirals

Simulation work by Harder, Valeriani and Cacciuto [120] has predicted that, in addition to shrinking, polymers in an active bath are much more likely to find themselves in a hairpin configuration, in which the polymer contains a single prominent bend and is otherwise extended, such as the conformation shown in Figure 5.13c. One way to measure the prevalence of this type of configuration is to measure the acylindricity of the polymer and compare it to the radius of gyration. We define the acylindricity as

\[ A^2 = \frac{\lambda_2 - \lambda_1}{R_g^2} \]

where \( \lambda_2 \) and \( \lambda_1 \) are the one-dimensional radii of gyration along the two principal axes. The acylindricity measures the relative difference of the radii of gyration, so if they are equal, as in the case of a uniform circle or square, then \( A = 0 \). The maximum limit of \( A \) corresponds to a line: when \( \lambda_1 = 0 \), the acylindricity is 1.
Simultaneous measurements of $A^2$ and $R_g$ allow us to detect hairpin-like configurations. For example, consider the three configurations in Figures 5.13a-c. We have added lines that represent the lengths and orientations of their principle radii of gyration: each cyan line has a length of $2\sqrt{\lambda_2}$, and each red line has length $2\sqrt{\lambda_1}$. The configuration in Figure 5.13 has $A^2 = 0.95$ and $R_g/(N\sigma) = 0.26$, which are near the maximum possible values. The acylindricity is large because $\lambda_2 \gg \lambda_1$, and $R_g$ is large because the polymer is fully extended. In comparison the configuration in Figure 5.13b is spread out more isotropically, $\lambda_1 \approx \lambda_2$, which corresponds to a smaller acylindricity, $A^2 = 0.31$. At the same time, the chain is more compact, which reduces the radius of gyration, $R_g/(N\sigma) = 0.11$.

Figure 5.13c shows an example of the hairpin configuration that we want to detect with $A^2 = 0.98$ and $R_g = 0.14$. The acylindricities of these configuration are very high because the hairpins are highly anisotropic but $R_g$ is much lower than it would be for a fully extended chain. This is the signal we look for to detect the prevalence of hairpins.

When we check the shapes of our polymers using these measures, we find many more conformations corresponding to hairpins in our trials with an active bath than we do in trials with a
passive bath. Figures 5.14a-c show the probability of conformations of our polymer in a passive bath as a function of $R_g$ and $A$, for polymer lengths of 32, 45, and 78, respectively. The color scale represents the probability of the polymer having the given $R_g$ and $A^2$, and probabilities for each trial are scaled by the trial’s maximum probability so that all the plots could be shown with the same color scale. Notice that the radius of gyration and acylindricity appear closely related. This is because our polymers are usually close to being fully extended, so any reduction in $R_g$ is also accompanied by a reduction in $A^2$.

In contrast, Figures 5.14d-f show the probability of conformations for the same polymers in an active bath. Compared to the passive trials, there are many more conformations with low $R_g$ for a given $A^2$, indicating that those conformations are more like hairpins, in agreement with the Harder et al.’s simulations [120].

We also note here that many of our trials with the passive particles reached a steady state quickly and were therefore not useful for calculating conformation probabilities. In some trials,
the polymer depleted to the wall, as shown for a chain with $N = 100$ monomers in Figure 5.15a. This doesn’t occur in trials with active particles because there are always some active particles on the wall that can eventually push the chain back into the bulk.

In other trials, the polymer collapsed into a coil, as shown in Figure 5.15b. This coiled configuration is the same as the one found by Liu, Jiang, and Hou [121] in their simulations for a polymer embedded in a chiral active bath. In their simulation, however, the bath was formed by self-propelled particles that had an average angular velocity in addition to their average directed motion. Further trials would be necessary to determine if there is some hidden chirality in our set up with passive particles or if the spiral configuration is just stable in 2D when the density is low enough that no particles from the bath are trapped within a loop configuration of the polymer.

![Figure 5.15: a) An example of a chain in a passive bath depleting to the boundaries. Once the chain has depleted, it never returns to the bulk. b) An example of a chain coiling in a passive bath. This is also a steady state as the grains can’t exert any force that would uncoil it.](image)

5.6 Conclusions

In this chapter, we observed finite-size effects on the discontinuous phase transition from disordered to collective-motion in a system of vibrated polar grains. We found that finite-size effects caused the discontinuous phase transition to appear continuous but that the system still contained the traveling bands reminiscent of those in the infinite-size limit. In the infinite-size limit, traveling bands represent phase coexistence between ordered and disordered regions, and they travel
at a constant speed set by the collective-motion phase. However, we found that finite-size effects allowed these bands to travel with density-dependent speeds because a “local equation of state” relating coarse-grained velocity to local density was continuous. We also found that the low density regions were still ordered. We think it is likely that, in a larger system with smaller relative fluctuations, our bands would continue to grow and the low density phase would become truly disordered. Our results show that, for many practical active systems, which are composed of hundreds or thousands of individuals, the phase transition may appear continuous so that incremental changes in density can cause incremental changes in average alignment.

Finally, we quantified the behavior of a model polymer in a bath of polar vibrated disks. At a first glance, it seemed that the primary effect of activity on the polymer was to increase the effective temperature of the active bath. However, closer examination showed that the most likely conformations also changed towards more “hairpin” configurations, which may have implications for the most likely protein folding configurations. Both of these results were in strong agreement with predictions from polymer simulation groups.
CHAPTER 6
POSSIBLE FOLLOW-UP WORK

There is always more work we could have done, and many of our results have prompted new questions. Here we discuss a few experiments that would be strong starting points to continue the research presented in this thesis.

In Chapter 3, we discussed the motility-induced phase separation that the ants appear to undergo at moderately high densities. The ants developed dynamic heterogeneities that we quantified in the small cells, using up to \( N \approx 1000 \) ants. We can also observe examples of these heterogeneities in the vertical cells that we used to detect the waves in Chapter 4. One interesting extension of the work might be to look at the phase separation in much larger cells placed perpendicular to the gravitational direction to see the typical size and duration of the heterogeneities. We also noticed in this chapter that \( U_{e_{f_{v}}} \) and \( U_{e_{f_{g}}} \) had the same shape, but that we needed to adjust \( U_{e_{f_{v}}} \) by a constant to make them equivalent. Can we use the knowledge that the ants spend about half their time moving regardless of the density to analytically predict this constant? If so, it would make measurements of \( U_{e_{f_{v}}} \) more useful in practical applications. Perhaps we could measure the behavior of a small crowd in an exhibition hall and then predict the behavior of a larger crowd, using this scaling.

We have also observed in the small cells that there are often “corridors” between stationary clusters so that moving ants can pass between them. This prompts us to wonder whether the geometry of the cell could control the location of the clusters or corridors. For instance, if we create a narrow opening connecting two halves of the cell, do the ants avoid stopping near the door? Does this effect change with density? Is there a similar effect near corners? Finally, can we model this kind of effect with an external effective potential? This might have applications in designing buildings. Perhaps we could identify geometries where people are more likely to pause and mingle. If humans, like ants, prefer to move a certain percent of the time and socialize for
a certain percent, it may help crowd flow to include geometries that promote mingling alongside geometries that promote walking in the desired corridors.

In Chapter 4, we saw that the activity cycles could lead to activity waves that propagated towards the free surface. These waves occur because the ants are confined at high density and are all actively moving. We stated that this was evidence of local-collective motion for the ants. However, the collective-motion was always in the same direction: towards the free surface. We also saw that, even before the wave passed, the ants were predisposed to align their long axis with gravity, so that they tended to face roughly up or down. This may have made it easier for the ants to align and manifest waves. A simple and interesting experiment to verify true collective-motion due only to aligning effects would be to confine the ants at $\phi_{eff} > 3.6$ to a small ring, in parallel to the one used for the disks in section 5.3. During an activity cycle, do the ants naturally begin to flow clockwise or counterclockwise around the ring each cycle? If we increase $D_{out} - D_{in}$ do the ants begin to form channels that flow different directions and do these channels have a typical width? On the other hand, does increasing $D_{out} - D_{in}$ begin to destroy the collective-motion because the ants have more freedom to move in arbitrary directions? One could also try to more closely analyze the results in the small cells to try detect the scaling of number fluctuations inside and outside of activity cycles. Recall that giant number fluctuations, $\langle (\Delta N)^2 \rangle \propto \langle n \rangle^{1.6}$, often accompany the collective-motion phase even in geometries that restrict uniform collective motion. If there is collective-motion during the activity cycles, one might be able to detect giant number fluctuations that turn off and on over the course of the cycle, clearly showing that the ants turn on and off their collective-motion by varying activity.

We tried to trigger activity waves in the vertical column of ants by moving the base wall of the cell to compress the ants at the base of the cell. However, we were only able to generate very small amplitude waves this way, and we still needed to wait for the ants to trigger an activity cycle. Another interesting experiment to look at the nonlinear activation rate of the ants would be to create another cell with a width that narrows as the height increases. Then any waves that start and travel upwards would naturally compress. If our theory about the nonlinear activation rate of
the ants is correct, would the waves now speed up? Other geometries could also be used, such as a column that suddenly widens at some height. Does the wave then disperse outwards away from the opening instead of just traveling upwards?

In Chapter 5, we discussed our measurements for ants in Janssen’s set-up. We found that the ants generally behaved more like passive grains than like a fluid, but also that their $m_{\text{app}}$ slowly increased over time, as shown in Figure 4.13. We wonder if this is due to the activity of the ants. Recall that in the small cells that we used to look for the activity cycles in Chapter 3, the ants always started off active when we placed them in the cell and then slowly deactivated over the first $\approx 10 \text{ min}$. It would be interesting to see the effect of activity cycles in the Janssen set-up, though this may involve letting a large mass of ants sit in the cell for more than hour, which means the ants could wipe off all of the talcum powder and begin to firmly grip the walls. A priori, one might expect activity cycles to increase $m_{\text{app}}$ because the connections between the ants are changing quickly, decreasing the elasticity of the ant-collective. However, the ants may be less likely to break connections with each other because of high strain when they are active so that they can try to crawl in their preferred direction.

Of course, another interesting experiment to perform with the ants in the Janssen set-up would be to use a larger cylinder diameter and see if the saturation mass increases, as it does for passive grains. However, this would require the collection of even larger colonies of ants than we were able to procure.

In our measurements of the typical sizes of the polymer chains in Chapter 6, we stated that the most immediate effect of activity seemed to be an increase in the effective temperature. However, to make this case strongly, it would be desirable to directly measure the changes to the effective temperature in the chain correlation length in the active and passive baths. Unfortunately, we did not have enough data to quantitatively measure the persistence length in either bath. The difference in orientation, $\Delta \theta$, between two segments of a polymer scales as $\langle \cos(\Delta \theta) \rangle = \exp - (L/P)$, where $P$ is the persistence length and $L$ is the distance between the two segments when the chain is completely extended. In an equilibrium system, $P$ is inversely related to the thermal
energy: \( P = \kappa/(k_B T) \), where \( \kappa \) is the bending stiffness. Therefore, a measure of \( P \) would give a direct comparison of the effective temperature in the two baths. However, in the passive bath, the chain changed configurations relatively slowly and many trials needed to be discarded because the chain depleted to the boundary, so we did not have enough statistics to accurately calculate the correlation length. In the active bath, we had more configurations, but also found anticorrelations at lengths that suggest they might be due to the chain being dragged by individual particles. In either case, longer trials that are promptly reset when the chain depletes to the wall should allow for the direct measurement of correlation lengths that might directly show a large change in the effective temperature.
The goal of this thesis was to examine the universal predictions of active matter theory, especially in regards to our main system, groups of fire ants. We began by quantifying the interactions between small numbers of fire ants with an effective potential. The effective potential had two parts: an attractive part that is best modeled as a motility-induced effect and a repulsive part that stems from the ants’ tendency to avoid crowding each others legs. We observed that the attractive part of the effective potential led to motility-induced phase separation in the ants at high densities. At these densities, the motility-induced phase separation is not disturbed by alignment interactions because the ants are rarely both active and aligned. However, we believe existing theory needs should expanded to fully account for mechanism that halts motility-induced phase separation in systems with social interactions. The simple rule we propose as a starting point is that each individual spends a certain percentage of its time engaged in social interactions and the rest of the time moving, regardless of the number of neighbors.

We also found that the ants undergo activity cycles if they are confined at extremely high densities. During these cycles, which the ants begin spontaneously, the ants are both active and close together, so aligning interactions lead to collective-motion. We saw the effects of the activity cycles and alignment in the large waves that the ants propagate against gravity in Chapter 4. This was the first observation of a new type of wave in active systems: an activity wave in which a spontaneous activity cycle propagated upwards towards the free surface. The wave in activity led to a wave in density and in alignment because dense active ants align. Eventually, we hope models of this kind of contagious activity may help explain crowd behavior in panics and stampedes.

Next we looked for the effects of activity by comparing ants to passive grains in Janssen’s experiment in Chapter 5. We found that the apparent mass of the ants in the cylinder saturates with enough added mass, so the ants behave more like a passive granular system than a passive fluid.
This was somewhat surprising because we didn’t expect ants to jam against each other to form the permanent force chains that occur in passive granular systems. Instead, their activity typically allows them to flow past each other and rearrange. However, it tentatively seems that, in the case of the ants, activity actually increases the strength of the force chains. We could see this in the slow increase of $m_{app}$ each time we add more ants. We speculate that activity may allow the ants to tolerate higher stresses on their interconnections as they try to actively move.

Finally we turned our attention to a well-known active system, vibrated polar grains, and examined the phase transition to collective-motion in Chapter 6. We found traveling bands in these experiments near the phase transition that behaved nonlinearly and also that the background through which the bands propagated was ordered. This is in stark contrast to the traveling bands that are usually described near the collective motion phase transition, which are phase separation between ordered and completely disordered phases. We measured a local equation of state in our system and proposed that our relatively short-lived bands were a result of finite-size effects, namely that the bands were created and dissipated by strong fluctuations. The role of finite-size effects in active systems has not been systematically studied yet, but it may prove an important new direction because many active systems of interest consist of on the order of only $10^2$ to $10^5$ individuals, instead of on the order of $10^{23}$ individuals typical in equilibrium systems, and fluctuations are larger in active systems than in equilibrium systems.

In all of the work presented in this thesis, we have built on current active matter theory to explain our results. For example, motility-induced phase separation, collective-motion, and traveling bands have been described theoretically and observed in different experimental systems. In these cases, our results verify the universal nature of active matter predictions, even in systems with complicated particle interactions like the interactions between social insects.

On the other hand, we have also introduced new active matter effects, such as the social interaction twist on motility-induced phase separation and the finite-size effects near the collective motion phase transition. We hope that these new observations and unanswered questions will prompt other groups to explain and explore these effects theoretically and in their own experimental systems.
We started this work with the goal of advancing the field of active matter and hope that this thesis will be considered a step forward.
REFERENCES


179


Caleb J. Anderson was born in Princeton, NJ, to Karl and Dawn Anderson and he grew up in
Houston, TX. He graduated from Texas A&M University with a Bachelor of Science degree in
Physics in 2013. He joined the Soft Condensed Matter Lab at Georgia Tech in 2014, when he
began his work with ants. In 2017, he was awarded the Chateaubriand Fellowship by the French
government to study vibrated polar grains in France. He was also awarded the FLAMEL fellowship
during his time at Georgia Tech to advance his study of materials and computing.