TESIS DOCTORAL

Flocking and pattern formation in active particles and epithelial tissues

AUTORA
CAROLINA TRENADO YUSTE

Tesis depositada en cumplimiento parcial de los requisitos para el grado de Doctor en Ingeniería matemática.

UNIVERSIDAD CARLOS III DE MADRID

DIRECTORES
Luis L. Bonilla y Ana Carpio

Leganés, Noviembre, 2020

A Rufo

Agradecimientos

A través de estas líneas quiero expresar mi agradecimiento a todas las personas que me han apoyado y colaborado en la realización de este trabajo de tesis.

En primer lugar, me gustaría agradecer a mis directores de tesis Luis L. Bonilla y Ana Carpio, por brindarme la oportunidad de realizar el doctorado, por vuestra dedicación e ilusión por la investigación, paciencia y orientación durante el trascurso de la tesis. Agradezco al Prof. I. Hambleton, director del Fields Institute, su invitación y apoyo durante el Workshop organizado por la Prof. A. Carpio: *Workshop on Modeling Biological Phenomena from Nano to Macro Scales* que tuvo lugar en 2018 en el Fields Institute en Toronto, Canadá. Agradezco al Dr. Rastko Sknepnek su ayuda con el software SAMoS y al Prof. Antonio Marquina por su colaboración con las ecuaciones hiperbólicas del modelo de Vicsek. Mi agradecimiento al Prof. Russel Caflisch, director del Courant Institute of Mathematical Sciences de la New York University, por su invitación a realizar una estancia de investigación y por su hospitalidad. Todo esto no habría sido posible sin la ayuda del programa propio de la Universidad Carlos III para mi estancia. La investigación de esta tesis ha sido financiada por los proyectos de investigación del Ministerio de Economía y Competitividad (ahora FEDER/Ministerio de Ciencia, Innovación y Universidades – Agencia Estatal de Investigación) No. MTM2014-56948-C2-2-P y No. MTM2017-84446-C2-2-R.

También me gustaría dar las gracias a mis compañeros de departamento, os agradezco que me hayáis aconsejado en cada momento de la tesis. A todos mis compañeros de despacho Juan, Miguel, Sergei, Rafa, Valle, gracias por haber compartido espacio y hacer que la convivencia entre nosotros fuera muy fácil y agradable. No me puedo olvidar de mis compañeros del despacho "de al lado", gracias Emanuel por compartir grupos de docencia, conversaciones y consejos, y en particular me gustaría agradecer a mi amiga y compañera Rocío por todos los momentos que hemos vivido en el doctorado. Gracias por haberme entendido en cada momento de nerviosismo y por haber podido realizar parte de este camino a su lado. Sin ti todos los congresos y escuelas de verano no hubieran sido igual. Gracias a los profesores del departamento por el buen ambiente de trabajo, Manuel, Jesús, Aurora, Eduardo, Antonio y en especial a Filippo. También agradezco los buenísimos momentos que he compartido con los compañeros del departamento de fluidos que me han dado muchos momentos de risa y felicidad.

A todas mis amigas, gracias por estar a mi lado y haberme apoyado y alegrado en los momentos que más necesitaba.

Por último quiero dar las gracias a toda mi familia y en especial a mis abuelos, porque sé que están muy orgullosos de su nieta. Por supuesto, no puedo terminar estas palabras sin agradecer a mis padres el apoyo constante durante toda la tesis (y vida), por haberme acompañado, aconsejado de la mejor manera y haberme dado las herramientas necesarias

para conseguir todo lo que me propongo. Muchos de mis logros os los debo a vosotros. Para mí sois los mejores.

Alex, gracias por enseñarme que la ilusión por la investigación tiene sus frutos y por apostar siempre por mi. Gracias por darme la fuerza que necesito cada día, por ser pura ilusión y cariño constante.

List of Publications

References

[1] Bonilla, L. L., Trenado, C. (2018). Crossover between parabolic and hyperbolic scaling, oscillatory modes, and resonances near flocking. Physical Review E, 98(6), 062603.
This publication is totally included in Chapter 2 of this PhD thesis.

[2] Bonilla, L. L., Trenado, C. (2019). Contrarian compulsions produce exotic time-dependent flocking of active particles. Physical Review E, 99(1), 012612.
This publication is totally included in Chapter 3 of this PhD thesis.

[3] Bonilla, L. L., Carpio, A., Trenado, C. (2020). Tracking collective cell motion by topological data analysis. arXiv preprint arXiv:2009.14161. Accepted in PLOS Computational Biology.
This publication is totally included in Chapter 4 of this PhD thesis.

Abstract

Collective behavior and, more specifically, flocking are phenomena observed in living systems, from bacterial colonies and spermatozoa, to larger systems such as insects and birds. These organizations exhibit changes from disordered to coherent behavior, which are examples of spontaneous symmetry-breaking out of equilibrium. Collective migrations in these systems can be predicted by simple models such as the Vicsek model (VM) or its variants, in which particles tend to align their velocities to an average of their neighbours'. The change from a disordered state to an ordered state can occur continuously or discontinuously and a variety of resulting patterns are possible. The study of mathematical models of these systems may reveal these changes to be bifurcations in their governing equations.

We consider a system of particles moving within a two dimensional box with periodic boundary conditions. In Chapter 2, and following Ihle's approach, we derive a kinetic equation for a one-particle distribution function in the limit of infinitely many particles by assuming molecular chaos. The kinetic equation is discrete in time and space and it always has a simple uniform solution that corresponds to the disordered state of the system. We have carried out a linear stability analysis of this state and studied the possible bifurcations issuing from it. In the usual case, particles align their velocities to their average velocity with any other particles within a circle of influence plus some angular noise, which has a uniform probability density. The spectrum of the linearized equation has always one multiplier on the unit circumference and there is another one that moves from inside to outside the unit circle as a control parameter crosses a critical value. We use bifurcation methods to derive amplitude equations that describe solutions issuing from the disordered state. The amplitude equations comprise a conservation law for a density disturbance coupled to a two dimensional vector equation for a current density. Analysis and numerical simulations of these equations show that their solutions exhibit an interplay between parabolic and hyperbolic behavior in two different time scales when the distance to the critical value of the bifurcation control parameter goes to zero. In this limit, there appear oscillation frequencies that give rise to resonance phenomena if the alignment rule contains a periodic function of time. Direct simulation of the VM confirms the existence of these resonances.

In Chapter 3, we use the same methodology to study the effect of modifying the probability density of the noise in the alignment rule by which VM particles change their velocities. The mechanism of velocity synchronization consists of: active particles may be conformist and align their velocities to the average velocity of their neighbors, or be contrarian and move opposite to the average angle. Depending on the weights of conformist and contrarian or almost contrarian rules, we study the ordered state solutions of the amplitude equations corresponding to period-doubling, Hopf, or pitchfork bifurcations of the disordered state.

In Chapter 4, we consider the collective migration of epithelial cell monolayers moving on a surface. This phenomenon is crucial for many relevant processes including wound healing,

morphogenesis, and cancer-cell invasion during metastasis. There are many experiments on confluent cellular motion and different mathematical and computational models in the literature. A convenient model based on the physics of foams considers the cells as non-overlapping two dimensional convex polygons. In the active vertex model we study, the cell centers are in a Delaunay triangulation and are subject to forces that constrain them to have target areas and perimeter length, other forces that try to align their velocities to neighboring cells (as in the VM), friction with the substrate, inertia, and stochastic forces. We have simulated numerically this model in two different cases related to wound healing and to invasion of one cell collective by another one: (i) a cellular monolayer spreading on empty space, and (ii) the collision of two different cell populations in an antagonistic migration assay. For (i), we discuss how inertia is necessary to explain the larger size of cells in the boundary with respect to those in the interior of the layer. For (ii), we discuss which parameters of the model produce results that agree with experiments by P. Silberzan's group. In both cases, the interfaces that separate cells from empty space or cells belonging to different populations are quite rough and may shed and absorb islands as time elapses. To analyze both images from experiments and results of numerical simulations, we use topological data analyses of the interfaces.

Contents

List of Abbreviations

VM	Vicsek Model
SVM	Standard Vicsek Model
SPP	Self-Propelled Particle
AMA	Antagonistic Migration Assay
BU	Backward Update
FU	Forward Update
AVM	Active Vertex Model
MDCK	Madin-Darby Canine Kidney
NRK	Normal Rat Kidney
HEK	Human Embryonic Kidney
TDA	Topological Data Analysis

Chapter 1
Introduction

The term *active matter* describes diverse systems across scales, from bacteria colonies [1, 2, 3] and spermatozoa [4] to larger scales as insects [5, 6] or birds [7, 8]. Active matter comprises a novel class of non-equilibrium living and non-living systems that are characterized by (i) having a large number of self driven units (active particles) and (ii) each agent conforming the system is capable of transform free energy into systematic motion. A recent review is [9].

Anyone to have observed the dynamics of the coordinated motion of a group of starlings, or the synchronous motion of a school of fish swimming together, can understand the motivation for the study of active matter; individual units that move in group or *flock* and generate large-scale complex dynamical motions. The rules governing the principles of physics that define the behavior of each individual, the mathematical models describing those behaviors and the exhibition of novel phases (i.e., stable configurations of the flocks) and phase transitions different from those in thermal equilibrium are topics which deserve to be studied.

1.1 Flocking phenomena

The interaction between active particles produces collective motion, i.e., individual particles move coherently forming patterns and shapes, a phenomenon called *flocking*. As it has been detailed above, there is a myriad of examples in living systems that seem similar at different scales. Flocking has been a lively subject of investigation since the mid 1980s. One of the first studies in the field is the pioneering work of [10], which studies the motion of birds by analyzing a recorded movie. He concluded that any unit can initiate a collective motion, which is propagated along the system as a wave. Typically entities in a flock have a nonzero polarization (the arithmetic mean of the orientation of their velocities). In a more recent experiment [11], their authors have concluded that starling flocks have scale-free correlations (they have a correlation length that scales with the flock size) that decay as a power law with a surprisingly small exponent. As the interaction range between birds is of the order of a few bird sizes, the interaction range may be much smaller than the correlation length, which scales as the flock size [11].

In nature, motion usually arises collectively because its effectiveness. For example, colonies or groups are stronger against a predator or are more efficient to obtain food. The collective behavior is a consequence of the interaction between the units conforming the system, which can be simple or more complex as combination of simple interactions. As

a result, a wide range of generic behaviors arise, including non-equilibrium order-disorder transitions, characteristic instabilities or pattern formation.

The main challenge when developing a study describing the observations of a collective motion, is to obtain the best definition of the individual trajectories of the entities. This work can be hard due to the large number of particles and the dimension of the space where they move. Accordingly, there exists a wide range of mechanisms to study collective motions. For example, in a well-defined area, the optical technique called Particle Image Velocimetry (PIV) obtains the velocity field of a fluid or cells using averages. Another technique is that of [12] who study experimentally a school of fish inside a shallow box, thereby approximating their motion to a two-dimensional configuration.

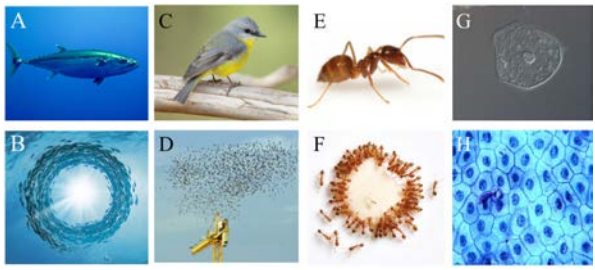

Figure 1.1: Collective organizations arising in nature. (A,B) Individual and collective swimming of a fish and a school of fish respectively. (C,D) Individual fly and flocking migration. (E,F) Individual and cooperative behavior of ants. (G,H) Single epithelial cell on a substrate and a monolayer of epithelial cells conforming a frog skin. (A) Adapted from (https://thefishsite.com/articles/how-climate-change-is-effecting-top-marine-predators). (B) Adapted from (http://simontucket.com/_Portfolio/PortPages_Hi/Il_FishSchool.html). (C) Adapted from (https://www.jjharrison.com.au/). (D) Adapted from (https://pixabay.com/photos/flock-of-birds-migratory-birds-sky-2150470/). (E) Image from [13]. (F) Adapted image from (https://commons.wikimedia.org/wiki/File:Ant_and_honey4.jpg). (G) Image from [14]. (H) Image from https://www.uwlax.edu/biology/zoo-lab/lab-2--microscopy-and-the-study-of-tissues/.

To understand flocking, it is convenient to pose and study mathematical models with minimal ingredients, yet capturing the essence of the phenomena. One of the earliest is the Vicsek model (VM) [15], which takes ideas from statistical mechanics into a non-equilibrium context. The VM comprises a set of self-propelling particles moving with unit speed so that each particle adapts its velocity to the average of the direction of neighboring particle velocities except for some alignment noise. The particle motion does not depend on the environment, which seems to be the case for birds within visual contact but not for fish or bacteria swimming in water. The VM is an example of "dry active matter" [9].

The analysis of active matter borrows many ideas and methods from equilibrium systems. For example, the formation of a flock from a disordered state is similar to the ferromagnetic transition in equilibrium statistical mechanics. We can define an order parameter, the polarization, that plays the same role as the magnetization. In the disordered state of a system of infinitely many particles, there is no flock and the mean orientation of the particle velocities is zero. Depending on the interaction range and parameters characterizing the particles behavior, the polarization may become positive, which characterizes the formation

of a flock. It is also possible to define correlations and correlation lengths that probe deeper into flocking [11]. In Section 1.3, we discuss several models of flocking phenomena [9, 16, 17, 18].

1.2 Tissue-cell migration

The discipline of active matter studies systems across scales, ranging from sub-cellular processes to the dynamics of tissues and organs. Concerning tissue cells, interesting dynamics occurs when an epithelial cell aggregate advances through an empty space, as in the case of wound healing, or it collides and encroaches a different tissue, as in cancer invasion. Advancing cellular fronts may display wave phenomena [20, 21], finger formation [22, 23, 24], or the interpenetration between oppositely moving fronts [25, 26]. Additionally, collective cell migration possesses challenging fundamental questions within the fields of soft and active matter, namely the characterization of fluid, solid or glass-like behavior associated with flocking and jamming-unjamming transitions [9, 27, 28, 29, 30, 31, 32, 33, 34]. The behavior of these systems is dictated by processes occurring at the individual cell level and also by collective phenomena arising in a group of cells.

On a smaller scale, cell motility has been widely investigated. The individual motion of one cell is a well-known phenomenon studied *in vivo*, as for example the migration of an individual dissociated cancer cell presented in [35]. However, there are some differences between single-cell motion and that of a group of cells migrating. During the migration of a single cell, its membrane is retractile. In contrast, for an epithelium, cells maintain their junctions and the retraction of the individual cell is no longer possible. Thus, the mechanisms are completely different during collective-cell migration, where the cell-cell contact must be conserved maintaining the integrity of the layer to get their purpose. Indeed, during a collective migration, there also exist relative displacements within the tissue and a reorganization between the cells. For illustration, Fig. 1.2.a shows the individual motion of a single cell. In Fig. 1.2.b the motion of cell 1 depends on its own activity and on the behavior of cells 2 and 3.

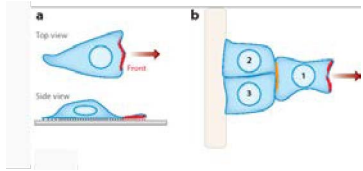

Figure 1.2: Individual and collective migration. a) Single cell on a 2D substrate. b) Collective cell migration. Figure from [42].

In particular, the dynamics, the morphogenic response and the collective strategies of eukaryotic cells are crucial in processes like morphogenesis [35, 36], tissue repair [37, 38], development [39] and embryogenesis [40]. Such living systems move in response of external signals to accomplish their functions. These stimuli are sent in the form of chemical gradients or notch signaling among others. While directional migration is dictated by these external influences, cell motility is also restricted by physical constraints. Cells, tissues or embryos need to study the environment to make their way squeezing through gaps, obstacles or to propel in a rheologically complex external matrix, with the main goal of finding the most efficient route. Processes as metastasis, branching morphogenesis [36] or embryogenesis

[40] are examples where living systems can deform their shape or anticipate their motion to facilitate their migration. In fact, it has been recently shown using experiments *in vivo*, that cancer cells have the capacity to distinguish between different trajectories [41] and use channels that guide the cells towards the best decision point.

Cell polarity is also another process which has raised attraction among researchers for many years. Several experiments explain the existence of leader cells who guide the rest of the cells, which are located at the back and called *followers*. Its only purpose is to follow the preferred direction guided by the leader cells. Indeed, followers are more compact than leader cells and instabilities processes may occur, giving rise to the formation of fingers in the front. Fig. 1.3 shows different types of cell migrations. Depending on the context, epithelial

Figure 1.3: Collective migration. a) Cells on a 2D substrate. b),c),d),e) and f), collective cell migration in 3D in different configurations and biological processes. Figure from [42].

cells can migrate in a single-layered epithelium 2D as in Fig. 1.3.a or by multi-cellular groups in a 3D configurations as in Fig. 1.3.b-f. Multicellular filaments in 3D can distinct between an external monolayer of cells and an internal cavity called lumen which is going to be an internal tube structure in the future. This organization is also possible during angiogenesis Fig. 1.3.c. Cells can also migrate in a finger shape with a poorly organized structure Fig. 1.3.d, move independently in clusters as in Fig. 1.2.e or penetrate a tissue as in Fig. 1.3.f.

Different aspects of these phenomena have been studied by models ranging from macroscopic continuum mechanics to detailed sub-cellular agent-based models [9, 22, 43, 44, 45]. In particular, the relevant processes of wound healing and confluent migration assay are explained in detail in Subsection 1.2.1 and Subsection 1.2.2 to introduce the protocols carried out in the experiments.

1.2.1 Wound Healing

Once one epithelium is formed, tissue cells remain in a continuum monolayer which is not motile unless the monolayer suffers a defect on it, namely a scratch or wound. Due to that, motion starts and cells try to fix the lesion keeping the junctions between them together. One of the experiments to study the phenomenon *in vivo*, is the so-called *scratch experience*. It consists of growing a culture epithelium on a petri dish and create a scratch using a pipette or scalpel.

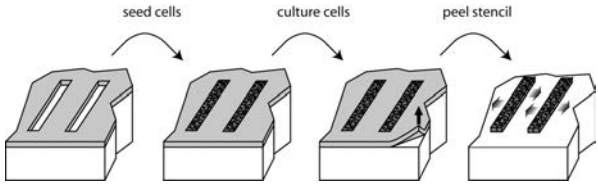

Figure 1.4: Experimental setup based on a stencil removal. Figure from [38].

In particular, in Chapter 4, we study the migration of a monolayer based on an slightly different experiment explained briefly below. The experimental setup consist of a stencil made by micro-fabrication which limits the grow of the cell culture equally distributed. Once the cell culture is ready, the stencil is removed. The advantages of this experimental protocol unlike the other experiment are, (i) cells are not damaged and (ii) the edge of the monolayer has no defects. Fig. 1.4 illustrates the experimental configuration and the steps involved in the experimental protocol.

The physical features of collective migrating cells under this experimental study are, (i) the instabilities appearing at the border of the monolayer, (ii) the velocity field inside the bulk and (iii) the different behaviors of the monolayer during the migration. Fig. 1.5 shows a sequence of images using the experimental protocol explained above. Following the processes developed

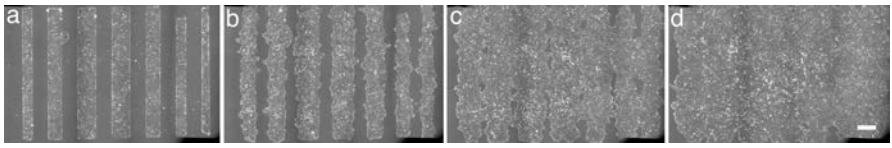

Figure 1.5: Snapshots showing the progression of seven monolayers of cells with different widths at different times, a) t = 90 min, b) t = 13 h, c) t = 25h and d) t = 37h from [38].

in Fig. 1.5 from [38] they observe, after one hour of the stencil removal, the apparition of leader cells. These special cells are born at the edge of the monolayer, and they have a different shape and size and loose their epithelial morphology. In particular, they never loose their contact within the follower cells and are characterized by having a very high directionality and velocity. All these features produces structures called fingers. The remaining cells in the monolayer do not spread out independently, and their growth and proliferation help to close

the injury quickly. This experimental setup is going to be reproduced and analyzed by means of numerical simulations of the active vertex model in Chapter 4.

1.2.2 Confluent migration assay

Cells are always competing for their space to ensure the resources and tissue cohesion. In general, tissue configuration is stable. However there are some specific situations where tissue becomes unstable as in cancer invasion. For example, recent studies have observed that precancerous cells can become highly competitive and try to kill healthy cells [46]. There is another recent experiment that has connected metastasis in colorectal cancer to wound healing and tumor invasion of tissue using molecular markers [47]. Recently, Moitrier *et al.* have reported confrontation assays between antagonistically migrating cell sheets [26]. In their experiment, each cell population grows into one of the compartments separated by a cell-free gap (Fig. 1.6.a). Once the culture insert is removed, the two confluent cellular monolayers advance toward the intermediate empty space, collide and Ras cells population displaces the wild type one (see Fig. 1.6.b). In particular, in these experiments, the invasive population behaves fluid-like and the invaded monolayer behaves solid-like.

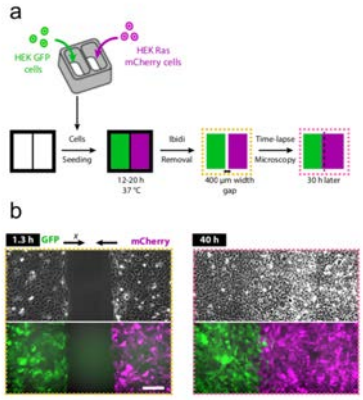

Figure 1.6: a) Experimental setup carried out in [26]. b) Side view of the AMA where purple population (Ras cells) invades green population (wild type).

Thus, our description of spreading of cellular tissue and antagonistic migration assays using a modified active vertex model might be relevant for metastatic cancer and will be studied in Chapter 4. In particular, we study the role of cellular junction tensions in cell invasion, agglomeration and segregation. Understanding precise biochemical mechanisms influencing cell-cell contact and confluent cellular tissue may help develop therapies for metastatic cancers.

1.3 Physical Models of Collective Migration

To understand some of the principles involved in the dynamical organization and motion of living systems, together with the quantification of some specific observables, a large variety of theoretical and numerical models have been derived. Specifically, we describe three different types of models, namely the discrete models usually referred to as *agent-based models*, the *continuum hydrodynamic-like models*, and the so-called *active vertex model* that studies the mechanics of confluent epithelial tissues. Below we give a brief review and description of the three approaches.

1.3.1 Agent-based models

The first studies that gave a minimal framework to explain the flocking behavior of living systems were the agent-based models; see Refs. [9, 16]. In particular, the 2D VM, describes the overdamped dynamics of N particles having the same constant speed, an alignment rule for their velocities, and moving on a box with periodic boundary conditions. As the particles have a given speed, the VM and its variants lie within the so called *self-propelled particle models* (SPP). The velocity of each entity adopts the average direction of its close neighbors plus a perturbation considered as a stochastic factor which affects directly to the motion of the individuals [15, 16].

In the literature, it is possible to find a wide range of variations of the VM along three specific pathways, namely, the variation in the polarity of the particles, the interaction between them (repulsion or attraction forces) and the external medium where they move. The direction of motion originally proposed by the VM depends on the average direction of the particles considered in the neighborhood defined as follows

$$\theta_j(t+\tau) = \mathrm{Arg}\left[\sum_{\langle j,l\rangle} e^{i\theta_l(t)}\right] + \xi_j(t), \tag{1.1}$$

where $\xi_j(t)$ is a random number chosen inside an interval $[-\eta\pi, \eta\pi]$, and the noise strength η has a maximum value of 1. This particular way to update the direction of each particle is called *angular noise*.

A different way to define the orientation dynamics presented in [48] is the so-called *vectorial noise*. The noise is not defined as an external factor as in Eq. (1.1), it arises from the interaction between the particles and their neighbors. Thus, instead of Eq. (1.1), the direction of each particle is updated as

$$\theta_j(t+\tau) = \mathrm{Arg}\left[\sum_{\langle j,l\rangle} e^{i\theta_l(t)} + n_j e^{i\xi_j(t)}\right], \tag{1.2}$$

where n_i is the number of neighbors of the particle j.

Regarding the equation for the particle position, the original updating scheme proposed by Vicsek et al. defines the speed as a backward difference (BU). However, several authors have implemented a slightly different version of the VM, using a forward difference (FU) as the presented in Eq. (1.3)

$$\mathbf{x}_j(t+\tau) = \mathbf{x}_j(t) + \tau\mathbf{v}_j(t+\tau). \tag{1.3}$$

Although there should not be major changes in the behavior of the system with these variations, recent studies have found different phase transitions [48, 49, 50]. In particular, the VM describes a non-equilibrium phase transition from disorder to order: particles move randomly

at low densities and high noise, and move ordered at high density and low noise. This phase transition will be detailed in Chapters 2 and 3.

There exist more intricate SPP models as the studied in Ref. [51]. In this work, instead of applying an averaging rule, the self-propelled particles adjust their direction towards the direction of the net-force acting on them. It is important to emphasize that the VM alignment rule is not able to maintain all the particles together in an open space without extra forces. However, there exists a generalized point particle model proposed by Sepúlveda et al. [24], which describes the collective motion of cells in an epithelium. In this model, cells are actively motile point particles with inertia, Vicsek-like alignment "forces", inter-particle and random "forces". The acceleration is a consequence of the collective motion of cells and the interaction with the environment. The inertia in this model is a consequence of the interaction of the cells with the medium and it is not given by the mass of a single cell. Particles are not self-propelled, so that they can stop their motion and start moving again if there are missing cells in their neighborhood and the active force is zero. The noise that drives the particles is taken to be an Ornstein-Uhlenbeck process. In particular, we will use this cellular dynamics within an active vertex model to describe collective cell migration in Chapter 4, where it is explained in depth.

1.3.2 Hydrodynamic equations

As we have highlighted throughout the introduction, SPP consume energy and dissipate it when they move, producing collective motions at large scales. In particular, another promising framework to describe theoretically the macroscopic properties and dynamics of active matter, is the so-called generalized hydrodynamic approach. It consists of developing a set of continuum equations that describe active matter systems as fluids or liquid crystals. The use of hydrodynamic fields allows us to give a macroscopic description of the non-equilibrium large-scale physics that epitomize the complex behavior of active matter. This framework does not takes into account the discrete features of these active systems, and thus the microscopic details and interaction are not explicitly considered.

One pioneering hydrodynamic theory is that of Toner and Tu [52, 53]. They consider the same fields appearing in the Navier-Stokes equations, namely the density and the velocity fields, $\rho(\vec{r},t)$ and $\vec{v}(\vec{r},t)$, respectively. The density obeys the usual continuity equation because the total number of particles is constant. Then they added to the Navier-Stokes equation for the velocity a source term containing linear and cubic terms in the velocity (so that there is a solution with constant speed $|\vec{v}(\vec{r},t)|$), a constitutive law for the pressure as a power series in the density disturbance $(\rho(\vec{r},t)-\rho_0)$, and terms that are at most quadratic in gradients and cubic in the velocity field. The velocity equation also contains a white noise force term. Higher order terms are excluded because they are irrelevant in the sense of the renormalization group [52]. The terms added to the velocity equation should be invariant with respect to translations and rotations. With the extra terms included in the Navier-Stokes equations, the Toner-Tu equations are sufficiently general to account for many patterns observed in direct simulations

of the VM and in experiments with flocks of animals [54].

Many works have derived different forms of the Toner-Tu equations from kinetic theories inspired in the VM; see e.g., the reviews [9, 16]. Among them, Ihle derived directly from the VM a Enskog type kinetic equation that is discrete in time and space. This equation for the one-particle distribution function comes directly from the N-particle distribution of the VM by assuming molecular chaos [17, 18, 19]. Coarse-graining the kinetic equation produces hydrodynamic equations similar to Toner-Tu's [18]. In our work, we study the linear stability of the uniform solution of the kinetic equation (constant one-particle distribution, corresponding to the disordered state of the flock) and obtain amplitude equations for the phases bifurcating from it. The amplitude equations are forms of the Toner-Tu equations but its coefficients have been obtained from the kinetic equation. Having understood linear stability and bifurcation theory for the kinetic equation of the VM, we have been able to propose modifications that yield exotic rotating wave and period doubling phases. These features are explained in Chapters 2 and 3.

1.3.3 Active vertex model

The first theoretical and numerical studies of epithelial tissues had relied on self propelled particle models [15, 24]. However, this approach is not enough to understand tissue dynamics, because these models does not take into account cell-cell junctions, which have an extreme importance in the behavior of tissue migrations. Indeed, these junctions allow the cells within epithelial tissues to remain together, but also permit the tissue to change its collective behavior from fluid-like to solid-like. For instance, during many biological processes, cells can organize and move together as in wound healing or move more fluid-like as in embryonic development. Hence, another alternative to study such mechanisms and properties of epithelial tissues is the usually referred to as active-vertex model, which has been proven to be a better alternative since it considers cell-cell junctions. In particular, this framework combines the vertex model and the dynamics of punctual particles described in Section 1.3.1.

The origins of the vertex model dates back to the physics of foams in 1970s [55]. Nonetheless, the vertex model considered within the active framework contains some simplifications in comparison to original models of foams. Such vertex model assumes that neighboring cells only share one single and straight edge. Therefore, the tissue is a polygonal mesh with polygons considered as cells, edges as cell junctions, and the vertices of the polygons as those points where two or more cells meet. Additionally, the vertex model is a quasi-static model that tries to maintain the sheet in mechanical equilibrium. Within the vertex model the system is characterized by means of an energy called E_{VM}, to be minimized. In particular, the E_{VM} relates the area, perimeter and junction of each polygon or cell. Hence, the relationship between the center of the cells and the edges of the polygons is established directly by computing the negative gradient of the energy. Indeed, it is necessary to introduce the dynamics of the AVM for the position and velocity of each cell which is computed in the center of each polygon.

Similarly to the agent-based models, recent studies have found that tissues exhibit phase transitions [34]. They have found that the way to quantify the fluid-solid transition depends on the cell shape index p_0: the ratio of the cell perimeter to the square root of the cell area. The value $p_0^* = 3.812$, which corresponds to pentagons, separates solid-like and fluid-like behavior of the tissue. For $p_0 < p_0^*$, cortical tension is prevalent over cell-cell adhesion, cells do not exchange neighbors and the monolayer is solid-like. For $p_0 > p_0^*$, cell-cell adhesion dominates, neighbor exchanges occur, and the cellular tissue behaves like a fluid, see Fig. 1.7 to illustrate. The detailed description of the active vertex model considering one of the particle

Figure 1.7: (a) Phase diagram showing the velocity vs cell shape. (b) Cell trajectories for fluid and solid-like behavior. Figure from [34].

dynamics presented briefly in Section 1.3.1 will be explained in Chapter 4.

1.4 Active Matter concepts

To characterize active systems, we need concepts such as phases, phase transitions and order parameters that arose in the study of equilibrium systems. As usual, a phase is a stable configuration of the system and a phase transition is the change from one phase to another as some control parameter goes through a critical point. The order parameter measures the transition from a symmetric phase (where it is zero) to a symmetry broken phase (where it is not zero). In equilibrium statistical mechanics, stable configurations or phases minimize the free energy. In active matter and in non-equilibrium systems, stable means resilient to disturbances, as in the well known theories of dynamical systems [56].

Familiar examples of phases in everyday life are the gaseous, liquid and solid phases (see Fig. 1.8). The particle configuration is quite different for each of those states. In the solid phase, particles are placed in a regular lattice. The liquid phase is a homogeneous state of the particles, but the particle density is high and there are short range correlations that decay at longer distances. In the gas phase, the density is lower and the particles may travel longer distances without experiencing interactions with each other. Typically, the temperature is a control parameter that induces changes from solid to liquid and from liquid to gas in substances such as water, as it increases past appropriate critical values.

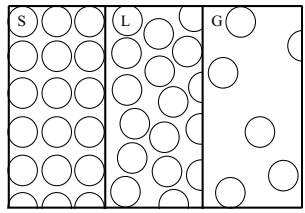

Figure 1.8: Particles in a solid, liquid and gaseous material.

The first classification of the transitions between phases of matter was proposed by Paul Ehrenfest in 1933. He recognized that some derivative of the free energy (with respect to temperature, volume, ...) becomes discontinuous at the critical value of the control parameter (e.g., the temperature) of the phase transition. Landau [57] introduced a theory of phase transitions based upon the notion of order parameter and on the smoothness of the free energy with respect to the order parameter and the control parameter of the phase transition. The order parameter is zero at the symmetric phase and nonzero at the less symmetric phases. Landau expanded the free energy in powers of the order parameter and considered that the coefficients in the expansion to be smooth functions of the control parameter. The condition that the derivative of the free energy with respect to the order parameter be zero yields the stable phase. If the free energy is invariant under sign change of the order parameter, then it is an even function thereof. The derivative of the free energy contains linear, cubic, quintic, ..., powers of the order parameter. At the critical control parameter, the coefficient of the linear term vanishes, and therefore, by a Taylor expansion, this coefficient is a linear function of the difference of the control parameter to its critical value, say $|\eta - \eta_c|$. Keeping up to cubic terms in the order parameter in the equation for the derivative of the free energy, the order parameters of the less symmetric phases are proportional to the square root of the absolute value of the difference of the control parameter to its critical value. Clearly, the nonzero order parameter exists only for control parameters above ($\eta > \eta_c$) or below ($\eta < \eta_c$) the critical point, whereas the symmetric phase with zero order parameter exists for any value of the control parameter. If the less symmetric phase exists for control parameter value for which the symmetric phase is unstable, then the phase transition is continuous. Otherwise, the less symmetric phase given by the vanishing of the free energy derivative is unstable, more terms in the order parameter expansion of the free energy needs to be kept and the phase transition is discontinuous [57]. Thus, depending on how is the change of the order parameter, we can distinguish between two different phase transitions,

- **First-order or discontinuous phase transition,** the order parameter changes discontinuously from zero to a non-zero value at a critical point.

- **Second-order or continuous phase transition**, the order parameter changes continuously from zero to a non-zero value (proportional to $|\eta - \eta_c|^{1/2}$) at the critical point. Close to the critical point, the behavior of the quantities describing the system are characterized by the so called critical exponents (for example $1/2$ in the previous

expression).

The Landau theory is very general and it can be extended to spatially dependent order parameters. In later years, it became clear that the assumptions about smoothness of the free energy near the critical point are not correct. By considering fluctuations [58], Ginzburg gave a criterion that establishes when the assumption that the coefficient of the free energy derivative is linear in $(\eta - \eta_c)$ fails. Then the Landau theory needs to be changed at a close neighborhood of the critical point and a renormalization group theory of phase transitions developed by K. G. Wilson and others replaces it [59]. However, the Landau theory gives a convenient frame of reference for ideas about equilibrium and non-equilibrium phase transitions. What kind of phase transitions and order parameters can be involved in active matter?

In many active systems, individuals move coordinately, which means that particles and their neighborhoods move with the same direction. This behavior arises when the agents after an interaction, align their velocities. The order parameter that measures the presence of swarming in these systems is the polarization

$$V = \frac{1}{N v_0} \left| \sum_{i=1}^{N} \vec{v}_i \right|, \tag{1.4}$$

where N is the number of particles, v_0 is the common speed of the particles and $\vec{v}_i$ the velocity of each particle. The polarization approximately equals zero when the particle velocities have no preferred orientation, and it is positive when a flock is formed. Thus, it is an order parameter in the sense of the Landau theory. If the particles move coherently in the same direction, V is going to be close to one. Possible control parameters of the flocking transition are the alignment noise (the larger the noise is, the greater the trend towards disorder) and the particle density (greater density favors particle polarization).

In the ordered state, particles are strongly correlated. The velocity between them is strongly correlated and behaviors as leader and follower can arise. It consist of the motion of the leader unit anticipating the motion of the follower unit. This behavior can be measured using a correlation function

$$c_{ij} = \langle \vec{v}_i(t) \cdot \vec{v}_j(t+\tau) \rangle, \tag{1.5}$$

where τ is the time delay of leader with respect to the motion of the follower, and the average is over the initial time. To characterize correlations after a sufficient time to form a flock, we can introduce the correlation function of the fluctuations [11]:

$$C(r) = \frac{\sum_{ij, i \neq j} \vec{u}_i \cdot \vec{u}_j \delta(r - r_{ij})}{c_0 \sum_{ij, i \neq j} \delta(r - r_{ij})}, \quad \vec{u}_i = \vec{v}_i - \frac{1}{N} \sum_{j=1}^{N} \vec{v}_j, \tag{1.6}$$

$r_{ij} = |\vec{x}_i - \vec{x}_j|$, such that $C(0) = 1$. The correlation function measures the average inner product of the velocity fluctuations of particles at distance r. A large value of $C(r)$ implies that the fluctuations are nearly parallel and thus strongly correlated. Conversely, the correlation function is negative if the fluctuations are anti-parallel. $C(r) = 0$ if the particles are

uncorrelated: $C(\xi) = 0$ defines the correlation distance ξ. For starling flocks, Cavagna et al. [11] have shown that $C(r) = r^{-\gamma} f(r/L)$, where L is the flock size and $\gamma \approx 0.19$ is a critical exponent. More precise definitions can be found in [60].

The previously defined parameters can be used to characterize flocking phenomena, and to compare theory, numerical simulations and experiments.

1.5 Linear stability and bifurcation analysis

The property of stability is necessary to understand the phase changes in complex systems. Typically, these systems have a simple *basic state*, which satisfies steady-state or time-dependent equations. The basic state is stable when sufficiently small disturbances decay with time, thereby leaving the system in the basic state. If any disturbance decays with time, the basic state is globally stable. The basic state is unstable when the amplitude of such disturbances grows with time, subsequently driving the system to a qualitatively different new state. Instabilities are responsible for many pattern formation processes in condensed and active matter systems that appear in nature and in complex media.

1.5.1 Some basic notions

Linear stability analysis allows as to measure quantitatively the stability of a system obtained by linearizing its governing equations about a fixed point, as it is now briefly explained. Let us consider a general n-dimensional dynamical system which evolves under the equation $\dot{\vec{x}}(t) = \vec{F}(\vec{x}(t))$. Fixed points or equilibrium solutions $\vec{x}^*$ satisfy $\vec{F}(\vec{x}^*) = \vec{0}$. To figure out whether a disturbance about a fixed point decays or grows, we linearize the equation $\dot{\vec{x}}(t) = \vec{F}(\vec{x}(t))$ about $\vec{x}^*$. The disturbance of $\varepsilon\vec{x} = \vec{x} - \vec{x}^*$ near the critical point evolves as

$$\frac{d}{dt}\varepsilon x_i(t) = \sum_{j=0}^{d} L_{ij}\varepsilon x_j(t) \tag{1.7}$$

If all the eigenvalues of the Jacobian matrix $L_{ij} = \partial F_i/\partial x_j$ (calculated at $\vec{x}^*$) have negative real part, the fixed point $\vec{x}^*$ is linearly stable, and $\varepsilon\vec{x}(t)$ decays to zero exponentially fast. The fixed point is unstable if there are eigenvalues with positive real part. Let us assume now that $\vec{F}(\vec{x}, \alpha)$ depends on a control parameter α and that all eigenvalues of L_{ij} have negative real parts for $\alpha < \alpha_c$, whereas one or more eigenvalues have zero real part for $\alpha = \alpha_c$ and acquire positive real part for $\alpha > \alpha_c$. Then the fixed point $\vec{x}^*$ loses stability at $\alpha = \alpha_c$. Depending on the type of instability at $\alpha = \alpha_c$, other solutions of the dynamical system may issue from the fixed point there. We say that these solutions *bifurcate* from $\vec{x}^*$ [61]. There are techniques that provide simpler equations for the amplitude of the bifurcating solutions, e.g., normal form theory [62]. Simple examples of bifurcation occur when a single real eigenvalue becomes zero at α_c. They include the saddle-node bifurcation in which two branches of stationary solutions that exist for $\alpha < \alpha_c$ merge at the critical value of the parameter and disappear

for $\alpha > \alpha_c$. If $\vec{x}^*$ exists for all values of α and the linearized dynamical system has a zero eigenvalue at α_c, the normal form of the resulting pitchfork bifurcation is the scalar equation [61]

$$\frac{dx}{dt} = ax - bx^3, \quad a = c(\alpha - \alpha_c). \tag{1.8}$$

Here the zero solution corresponds to the fixed point $\vec{x}^*$, and there are two other constant solutions $x^* = \pm\sqrt{a/b}$ for $ab > 0$. Let us assume $c > 0$ so that the zero solution becomes linearly unstable for $\alpha > \alpha_c$. The extra solutions bifurcate from $x^* = 0$ either for $\alpha > \alpha_c$ (if $b > 0$) or for $\alpha < \alpha_c$ ($cb < 0$). If $b > 0$ the extra solutions are stable (supercritical bifurcation), whereas they are unstable if $b < 0$ (subcritical bifurcation). Figure 1.9 depicts a supercritical pitchfork bifurcation for the one-dimensional dynamical system $\dot{x} = ax - x^3$. We can represent the supercritical normal form ($b = 1$) in terms of a potential

Figure 1.9: Shows $\dot{x}$ for different values of a and $b = 1$.

$$\frac{dx}{dt} = -\frac{d}{dx}U(x) \tag{1.9}$$

where U(x) is the potential. The local minimum of the potential is corresponded to the fixed stable point. In this case, the potential is

$$U(x) = -\frac{ax^2}{2} + \frac{x^4}{4}. \tag{1.10}$$

Note that this potential (See Fig 1.10) is the same as the free energy in the Landau theory for a second-order phase transition. From the structure of the potential, one immediately realises that for $a < 0$, $U(x)$ reaches its minimum at $x = 0$ (the symmetric phase). However, this state becomes unstable when $a > 0$ and two new minima appear at non-zero values of x symmetrically around $x = 0$ line. As soon as the system reaches one of this states, the $x \longrightarrow -x$ symmetry gets broken. This is the symmetry broken phase. Other simple bifurcations occur when two complex conjugate eigenvalues of the linearized equation about the fixed point acquire positive real parts for $\alpha > \alpha_c$. In this case, it is possible that time periodic solutions of the dynamical system bifurcate from the fixed point, which is called a Hopf bifurcation [61]. It turns out that the disordered state of a modified two-dimensional (2D) Vicsek model forms time-periodic flocks through a Hopf bifurcation. See Chapters 2 (pitchfork bifurcation) and 3 (Hopf bifurcation) for the analysis of a kinetic equation formulation of the VM.

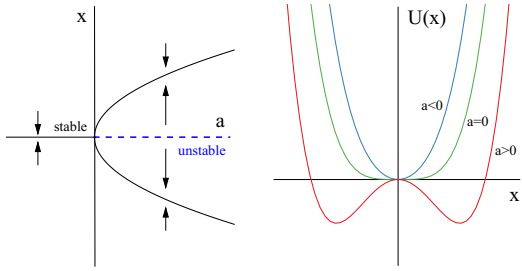

Figure 1.10: Bifurcation diagram, (supercritical pitchfork) and the potential for different values of a.

1.6 Thesis outline

The results obtained during the thesis will be organized in the body of the text divided as follows,

We will begin by studying in Chapter 2 the flocking in the two dimensional Vicsek model [15] by means of stability and bifurcation analysis of its kinetic equation [19]. The uniform distribution function corresponding to a disordered flock loses stability at a critical value of the control parameter (e.g., the alignment noise). At the critical value, one eigenvalue crosses from negative to positive values whereas there is always a zero eigenvalue corresponding to conservation of the total number of particles. There is a pitchfork bifurcation with one real and one complex mode. Our bifurcation study shows that the amplitude equations describing the flocking states consist of a scalar equation for the density disturbance, and a vector equation for the particle current density. The equation for the density disturbance is a continuity equation whereas the equation for the current density contains two different scalings: (i) a hyperbolic scaling in which both time and space scale in the same way; and (ii) a parabolic scaling in which time scales as space squared. In the hyperbolic time scale, the equations exhibit time dependent oscillations about a state of zero density disturbance and constant current density. These frequencies give rise to resonance phenomena that appear in direct simulations of the VM for values of the control parameter sufficiently close to the critical one.

Following the same research direction, in Chapter 3 we will study a two-dimensional modified Vicsek model and its corresponding kinetic equation. The mechanism of synchronization will consists of active particles being *conformist* when they align their velocities to the average velocity of their neighbors, or being *contrarian* when they move opposite to the average angle. Depending on the weights of conformist and contrarian or almost contrarian rules, the ordered states will appear as period-doubling, Hopf or pitchfork bifurcations.

The study of collective behavior in particles, allowed us to lead our research direction to the collective dynamics of tissue cells from a theoretical and experimental point of view. This phenomenon is known to be crucial in many relevant processes as wound healing,

morphogenesis or even the cancer cell invasion developing the metastasis. Within the context of epithelial cells, in Chapter 4, we will shed light on some fundamental aspects of tissue cell migration. In particular, we will describe the active vertex model and the different configurations analyzed: (i) A monolayer spreading on an empty space, and (ii) the collision of two different cell populations in an antagonistic migration assay. Indeed, we will apply topological data analysis on the images obtained from the experiments and simulations to study the defects flourished during these processes.

References

[1] O. A. Igoshin, A. Mogilner, R.D. Welch, D. Kaiser, and G. Oster, Pattern formation and traveling waves in myxobacteria: Theory and modeling, Proc. Natl. Acad. Sci. USA **98**, 14913 (2001).

[2] O.A. Igoshin, J.C. Neu, and G. Oster, Developmental waves in myxobacteria: A distinctive pattern formation mechanism, Phys. Rev. E **70**, 041911 (2004).

[3] R. Balagam, and O.A. Igoshin, Mechanism for Collective Cell Alignment in Myxococcus xanthus Bacteria, PLoS Comp. Biology **11**, e1004474 (2015).

[4] A. Creppy, F. Plouraboué, O. Praud, X. Druart, S. Cazin, H. Yu, and P. Degond, Symmetry-breaking phase transitions in highly concentrated semen. J. R. Soc. Interface **13**, 20160575 (2016).

[5] A. Cavagna, D. Conti, C. Creato, L. Del Castello, I. Giardina, T. S. Grigera, S. Melillo, L. Parisi, and M. Viale, Dynamic scaling in natural swarms, Nat. Phys. **13**, 914 (2017).

[6] A. Cavagna, I. Giardina, and T.S. Grigera, The physics of flocking: Correlation as a compass from experiments to theory, Phys. Rep. **728**, 1 (2018).

[7] J. Toner, Y. Tu, and S. Ramaswamy, Hydrodynamics and phases of flocks. Ann. Phys. **318**, 170 (2005).

[8] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic, Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study, Proc. Natl. Acad. Sci. USA **105**, 1232 (2008).

[9] M.C. Marchetti, J. F. Joanny, S. Ramaswamy, T.B. Liverpool, J. Prost, M. Rao, and R.A. Simha, Hydrodynamics of soft active matter, Rev. Mod. Phys. **85**, 1143 (2013).

[10] W. K. Potts, The chorus-line hypothesis of manoeuvre coordination in avian flocks, Nature, **309**, 344 (1984).

[11] A. Cavagna, A. Cimarelli, I. Giardina, G. Parisi, R. Santagati, F. Stefanini, and M. Viale, Scale-free correlations in starling flocks, Proc. Natl. Acad. Sci. USA **107**, 11865 (2010).

[12] C. Becco, N. Vandewalle, J. Delcourt, and P. Poncin, Experimental evidences of a structural and dynamical transition in fish school, Physica A: Statistical Mechanics and its Applications, **367**, 487, 2006.

[13] Z. Lin, C. Gao, M. Chen, X. Lin, and Q. He, Collective motion and dynamic self-assembly of colloid motors. Current opinion in colloid and interface science, **35**, 51 (2018).

[14] D.B. McMillan, and R.J. Harris, An atlas of comparative vertebrate histology, Academic Press, (2018).

[15] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, Novel type of phase transition in a system of self-driven particles, Phys. Rev. Lett. **75**, 1226 (1995).

[16] T. Vicsek, and A. Zafeiris, Collective motion, Phys. Rep. **517**, 71 (2012).

[17] T. Ihle, Large density expansion of a hydrodynamic theory for self-propelled particles, Eur. Phys. J. Spec. Top. **224**, 1303 (2015).

[18] T. Ihle, Chapman-Enskog expansion for the Vicsek model of self-propelled particles, J. Stat. Mech. (2016) 083205.

[19] T. Ihle, Kinetic theory of flocking: Derivation of hydrodynamic equations, Phys. Rev. E **83**, 030901(R) (2011).

[20] P. Rodríguez-Franco, A. Brugués, A. Marín-Llauradó, V. Conte, G. Solanas, E. Batlle, J. J. Fredberg, P. Roca-Cusachs, R. Sunyer, and X. Trepat, Long-lived force patterns and deformation waves at repulsive epithelial boundaries, Nature Mat. **16**, 1029 (2017).

[21] X. Serra-Picamal, V. Conte, R. Vicent, E. Anon, D.T. Tambe, E. Bazellieres, J.P. Butler, J.J. Fredberg, and X. Trepat, Mechanical waves during tissue expansion, Nature Phys. **8**, 628(2012).

[22] G. Y. Ouaknin and P. Z. Bar-Yoseph, Stochastic collective movement of cells and fingering morphology: no maverick cells, Biophys. J. **97**, 1811 (2009).

[23] R. Alert, C. Blanc-Mercader, and J. Casademunt, Active Fingering Instability in Tissue Spreading, Phys. Rev. Lett. **122**, 088104 (2019).

[24] N. Sepúlveda, L. Petitjean, O. Cochet, E. Grasland-Mongrain, P. Silberzan, and V. Hakim, Collective cell motion in an epithelial sheet can be quantitatively described by a stochastic interacting particle model, PLoS Comput Biol, **9**, e1002944 (2013).

[25] S. Porazinski, J. de Navascués, Y. Yako, W. Hill, M. R. Jones, R. Maddison, Y. Fujita, and C. Hogan, EphA2 Drives the Segregation of Ras-Transformed Epithelial Cells from Normal Neighbors, Current Biol. **26**, 3220 (2016).

[26] S. Moitrier, C. Blanch-Mercader, S. Garcia, K. Sliogeryte, T. Martin, J. Camonis, P. Marcq, P. Silberzan, and I. Bonnet, Collective stresses drive competition between monolayers of normal and Ras-transformed cells, Soft Matter **15**, 537 (2019).

[27] T. Angelini, E. Hannezo, X. Trepat, M. Marquez, J.J. Fredberg, and D.A. Weitz, Glass-like dynamics of collective cell migration. PNAS **108**, 4714(2011).

[28] J.-A. Park, J.-H. Kim, D. Bi, J.A. Mitchel, N. T. Qazvini, K. Tantisira, C. Y. Park, M. McGill, S.-H. Kim, B. Gweon, J. Notbohm, R. Steward Jr, S. Burger, S.H. Randell, A. T. Kho, D. T. Tambe, C. Hardin, S. A. Shore, E. Israel, D. A. Weitz, D. J. Tschumperlin, E. P. Henske, S.T. Weiss, M. L. Manning, J. P. Butler, J.M. Drazen, and J.J. Fredberg, Unjamming and cell shape in the asthmatic airway epithelium. Nature Mat. **14**, 1040 (2015).

[29] F. Giavazzi, C. Malinverno, S. Corallino, F. Ginelli, G. Scita, and R. Cerbino, Giant fluctuations and structural effects in a flocking epithelium, J. Phys. D: Appl. Phys. **50**, 384003 (2017).

[30] F. Giavazzi, M. Paoluzzi, M. Macchi, D. Bi, G. Scita, L. Manning, R. Cerbino, and C. Marchetti, Flocking Transition in Confluent Tissues. Soft Matter **14**, 3471 (2018).

[31] A. Palamidessi, C. Malinverno, E. Frittoli, S. Corallino, E. Barbieri, S. Sigismund, G. V. Beznoussenko, E. Martini, M. Garre, I. Ferrara, C. Tripodo, F. Ascione, E. A. Cavalcanti-Adam, Q. Li, P. P. Di Fiore, D. Parazzoli, F. Giavazzi, R. Cerbino, and G. Scita, Unjamming overcomes kinetic and proliferation arrest in terminally differentiated cells and promotes collective motility of carcinoma. Nat. Mater. **18**, 1252 (2019).

[32] E. Méhes and T. Vicsek, Collective motion of cells: from experiments to models, Integrative Biol. **6**, 831-854 (2014).

[33] C. Malinverno, S. Corallino, F. Giavazzi, M. Bergert, Q. Li, M. Leoni, A. Disanza, E. Frittoli, A. Oldani, E. Martini, T. Lendenmann, G. Deflorian, G.V. Beznoussenko, D. Poulikakos, K. H. Ong, M. Uroz, X. Trepat, D. Parazzoli, P. Maiuri, W. Yu, A. Ferrari, R. Cerbino, and G. Scita, Endocytic reawakening of motility in jammed epithelia, Nature Mat. **16**, 587 (2017).

[34] D. Bi, X. Yang, M.C. Marchetti, and L. Manning, Motility-Driven Glass and Jamming Transitions in Biological Tissues, Phys. Rev. X **6**, 021011 (2016).

[35] P. Friedl and D. Gilmour, Collective cell migration in morphogenesis, regeneration and cancer, Nature Rev. Cell Biol. **10**, 445 (2009).

[36] K. Goodwin, S. Mao, T. Guyomar, E. Miller, D. Radisky, A. Košmrlj, and C.M. Nelson, Smooth muscle differentiation shapes domain branches during mouse lung development, Development, **146**, 2019.

[37] A. Brugués, E. Anon, V. Conte, J.H. Veldhuis, M. Gupta, J. Colombelli, J.J. Muñoz, G.W. Brodland, B. Ladoux, and X. Trepat, Forces driving epithelial wound healing, Nature Phys. **10**, 683 (2014).

[38] M. Poujade, E. Grasland-Mongrain, A. Hertzog, J. Jouanneau, P. Chavrier, B. Ladoux, A. Buguin and P. Silberzan, Collective migration of an epithelial monolayer in response to a model wound, Proc. Natl. Acad. Sci. U. S. A. **104**, 15988 (2007).

[39] C. J. Weijer, Collective cell migration in development, J. Cell Sci. **122**, 3215 (2009).

[40] G. Forgacs, and S.A. Newman. (2005). Biological physics of the developing embryo, Cambridge University Press.

[41] J. Renkawitz, A. Kopf, J. Stopp, et al. Nuclear positioning facilitates amoeboid migration along the path of least resistance, Nature, **568**, 546 (2019).

[42] P. Rørth, Collective cell migration, Annual Review of Cell and Developmental, **25**, 407 (2009).

[43] V. Hakim and P. Silberzan, Collective cell migration: a physics perspective, Rep. Prog. Phys. **80**, 076601 (2017).

[44] D. L. Barton, S. Henkes, C. J. Weijer, and R. Sknepnek, Active Vertex Model for Cell-Resolution Description of Epithelial Tissue Mechanics, PLoS Comput. Biol. **13**, e1005569 (2017).

[45] R. Alert and X. Trepat, Physical Models of Collective Cell Migration, Ann. Rev. Cond. Matter Phys. **11**, 77 (2020).

[46] E. Moreno and K. Basler, dMyc transforms cells into super-competitors, Cell, **117**, 117 (2004).

[47] K. Ganesh, H. Basnet, Y. Kaygusuz, A. M. Laughney, L. He, R. Sharma, K. P. O'Rourke, V. P. Reuter, Y.-H. Huang, M. Turkekul, E. E. Er, I. Masilionis, K. Manova-Todorova, M. R. Weiser, L. B. Saltz, J. Garcia-Aguilar, R. Koche, S. W. Lowe, D. Pe'er, J. Shia, and J. Massagué, L1CAM defines the regenerative origin of metastasis-initiating cells in colorectal cancer, Nature Cancer **1**, 28 (2020).

[48] G. Baglietto, and E.V. Albano, Nature of the order-disorder transition in the Vicsek model for the collective motion of self-propelled particles, Phys. Rev. E, **80**, 050103 (2009).

[49] H. Chaté, F. Ginelli, G. Grégoire, and F. Raynaud, Collective motion of self-propelled particles interacting without cohesion, Phys. Rev. E, **77**, 046113 (2008).

[50] H. Chaté, F. Ginelli, G. Grégoire, F. Peruani, and F. Raynaud, Modeling collective motion: variations on the Vicsek model, The European Phys. Jour. B, **64**, 451 (2008).

[51] B. Szabo, G.J. Szöllösi, B. Gönci, Z. Jurányi,D. Selmeczi, and T. Vicsek, Phase transition in the collective migration of tissue cells: experiment and model, Phys. Rev. E, **74**, 061908 (2006).

[52] J. Toner and Y. Tu, Long-range order in a two-dimensional dynamical XY model: how birds fly together, Phys. Rev. Lett. **75**, 4326 (1995).

[53] J. Toner and Y. Tu, Flocks, herds, and schools: A quantitative theory of flocking, Phys. Rev. E, **58**, 4828 (1998).

[54] J.-B. Caussin, A. Solon, A. Peshkov, H. Chaté, T. Dauxois, J. Tailleur, V. Vitelli, and D. Bartolo, Emergent Spatial Structures in Flocking Models: A Dynamical System Insight. Phys. Rev. Lett. **112**, 148102 (2014).

[55] D.L. Weaire and S. Hutzler, The physics of foams. Oxford University Press (2001).

[56] S. H. Strogatz, Nonlinear dynamics and chaos, Perseus Books Publ. Cambridge, MA, 1994.

[57] L. Landau, in Collected Papers of L.D. Landau, Gordon and Breach (1967), p.193. [Phys. Z. Sowjet. **11**, 26 (1937); **11**. 545 (1937)].

[58] D. J. Amit and V. Martín Mayor, Field Theory, the Renormalization Group and Critical Phenomena, 3rd. rev. ed. World Sci. Singapore, 2005.

[59] K. G. Wilson, Renormalization group and critical phenomena. I Renormalization Group and the Kadanoff scaling picture, Phys. Rev. B **4**, 3174 (1971); Renormalization group and critical phenomena. II Phase-space cell analysis of critical behavior. Phys. Rev. B **4**, 3184 (1971). K. G. Wilson, The Renormalization Group and critical phenomena, Rev. Mod. Phys. **55**, 583 (1983) (Nobel Prize lecture).

[60] A. Attanasi, A. Cavagna, L. Del Castello, I. Giardina, S. Melillo, L. Parisi, O. Pohl, B. Rossaro, E. Shen, E. Silvestri, and M. Viale, Finite-Size Scaling as a Way to Probe Near-Criticality in Natural Swarms, Phys. Rev. Lett. **113**, 238102 (2014).

[61] G. Iooss and D.D. Joseph, Elementary Stability and Bifurcation Theory, 2nd ed. Springer, New York, 1990.

[62] J. Guckenheimer and P. Holmes, Nonlinear oscillations, dynamical systems, and bifurcations of vector fields, Springer, Berlin 1983.

Chapter 2

Flocking in the Vicsek Model

Contents

2.1 Introduction

Collective motion, or flocking, is a phenomenon observed in active systems with a large number of constituents such as bacteria colonies [1, 2, 3], spermatozoa [4], or larger-scale systems as insects [5, 6], birds [7, 8, 9, 10], animals [11], or even interacting robots [12]. These systems exhibit interesting changes in their behavior depending on internal or external influences. Such collective migrations can be predicted by simple models as the Vicsek model (VM) [13] or its variants [14, 15, 16, 17, 18], where particles move with equal speed and tend to align their velocities to an average of their neighbors. These organizations display phase transitions from disordered to coherent behavior depending on the parameters of the model. In particular, for the VM, when the alignment noise is sufficiently small or the particle density is high enough, particles move coherently as a swarm. Below a critical size of the box, flocking occurs as a continuous bifurcation from a disordered state, with uniform particle density, to an ordered state. For a box size larger than critical, the bifurcation is discontinuous and a variety of patterns are possible [20, 21].

As the VM is straightforward to simulate numerically, many variations thereof have contributed to our understanding of flocking [6, 14, 15, 18]. To delve deeper into flock

formation, many authors have derived continuum equations from the VM and its variants, often creating new models in the process (cf. the review papers Refs. [6, 14, 15, 18]). In most cases these active systems have many degrees of freedom, thereby complicating their theoretical and numerical analysis. In particular, the general form of the macroscopic transport equations is usually obtained taking advantage from the symmetries of the system. This means that the corresponding coefficients of the equations, have not explicit expressions and it can produce models with many free parameters that are difficult to obtain numerically. In addition, the complexity of some systems does not allow to obtain trivial symmetries.

Several authors have proposed kinetic theory equations based on the VM and then derived continuum equations from them. In a remarkable formulation, T. Ihle has derived several discrete-time kinetic equations that keep many features of the VM [16, 17, 19]. He then derived coupled continuum equations for the particle density and the momentum (or particle current) density by means of a Chapman-Enskog procedure valid near the transition to flocking [19]. These continuum equations contain terms that appear in the Toner-Tu macroscopic theory [7], and their coefficients have explicit expressions. However, Ihle's derivation introduces scaling *a posteriori* and it is not a systematic derivation based in bifurcation theory.

In this Chapter, we analyze flocking in the two-dimensional (2D) VM following Ihle's work [16] to derive a kinetic equation for a one-particle distribution function in the limit of infinitely many particles by assuming molecular chaos. The kinetic equation is discrete in space and time and it always has a simple uniform solution that corresponds to the disordered state of the system. We carry out a linear stability analysis of this state and study the possible bifurcations issuing from it. Direct numerical simulations confirm the theoretical studies.

2.2 The Vicsek Model

To set up a quantitative description of the behavior of flocking, VM appeared as a one of the pioneer models to display an explanation to order transitions. In nature, flocking are subjected to perturbations such as the weather or chemical concentrations, due to that, in the VM these perturbations are included in the direction of motion as a natural consequence of external or internal factors in the system.

The two dimensional VM model describes the dynamics of N particles represented by points and characterized by their position $\mathbf{x}_j$ and the direction of motion θ_j. Here, j is the particle index and $t = 0, \tau, 2\tau, \ldots$ is the time. All the particles move with the same constant speed (v_0) and tend to align locally their velocities to an average of those of their neighbors. The particles undergo discrete dynamics so that their positions are forwardly updated [1]

$$\mathbf{x}_j(t+\tau) = \mathbf{x}_j(t) + \tau \mathbf{v}_j(t+\tau), \tag{2.1}$$

the velocity of a particle $\mathbf{v}_j = v_0(\cos\theta_j, \sin\theta_j)$ is constructed to have an absolute value v_0

[1]Note that the scheme proposed by Vicsek et al. in [13] defines the speed as a backward difference, in contrast we use a forward difference.

and a direction of motion given by the angle θ_j. Particles align their direction of motion depending on their neighbors which are considered as all particles inside a circle of radius R_0 centered at particle $\mathbf{x}_j$ (which is included), see Fig. 2.1. In two dimensions, the angle of the direction of motion θ_j at time $t+\tau$, is obtained according to

$$\theta_j(t+\tau) = \mathrm{Arg}\left(\sum_{|\mathbf{x}_l-\mathbf{x}_j|<R_0} e^{i\theta_l(t)}\right) + \xi_j(t), \tag{2.2}$$

where we sum over all particles that, at time t, are inside a circle of radius R_0 centered

Figure 2.1: Sketch of the Vicsek dynamics. The particle in red aligns its direction towards the average direction of the local neighbors. Blue arrow is the new direction for the red particle.

at $\mathbf{x}_j$ (the *circle of influence* or *interaction circle*). $\xi_j(t)$ is a random number chosen with probability density $g(\xi)$. Typically, $g(\xi)$ is uniform inside an interval $(-\eta/2, \eta/2)$

$$g(\xi) = \begin{cases} \frac{1}{\eta}, & |\xi| < \frac{\eta}{2}, \\ 0, & \text{otherwise}, \end{cases} \tag{2.3}$$

where $0 \leq \eta \leq 2\pi$. To study the model, we choose a convenient nondimensionalization according to Table 2.1 and we set the velocity and time units as $v_0 = \tau = 1$. This consideration leads that model only depends on the (i) the density ρ_0 which establishes a relationship between the box size L and the number of particles N, $\rho_0 = N/L^2$, (ii) the noise strength η and (iii) the average number of particles inside the region of influence, $M = \pi R_0^2 \rho_0$ that remains an unchanged dimensionless parameter.

$\mathbf{x}$, R_0, L	$\mathbf{v}$	t	θ, ξ
$v_0\tau$	v_0	τ	$-$

Table 2.1: Units for nondimensionalizing the equations of the model.

2.2.1 Measuring the order

VM exhibits interesting changes in the behavior of the particles depending on internal (neighbouring interaction) or external influences. Eq. (2.2) contains two different alignment rules which affect directly on the particle's motion. The first term of the equation is an explicit polar term and just in the case that it overcomes the strength of the noisy rule, the system undergoes to an ordered state and the system has a coherent behavior. The way to measure these collective consensus is quantified by the complex order parameter

$$Z = W\, e^{i\Upsilon} = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j}, \tag{2.4}$$

whose amplitude $0 < W < 1$ (polarization) measures macroscopic coherence of the particles and Υ is their average phase.

For the VM, increasing the average number of particles inside the region of influence, $M = N\pi R_0^2/L^2$, favors flocking as more and more particles try to move together. Instead, increasing the alignment noise η tends to destroy flocks, as it dilutes the efficacy of the alignment rule. Then there are a critical value of M above which the polarization is $W > 0$ and below which $W = 0$ (in the limit as $N \to \infty$). The alignment noise also has a critical value, but now it has $W > 0$ below and $W = 0$ above threshold. Fig. 2.4 represents the polarization given by the modulus of the complex parameter Eq.(2.4). We observe that the values obtained by direct simulations of the VM tend to the predicted solution as we increase the density ρ_0 from 5 to 10 and when the noise amplitude increases, the system undergoes a phase transition from an ordered state, in which the particles move with the same direction, to an disordered state, in which the particles move randomly. At zero noise, particles move in

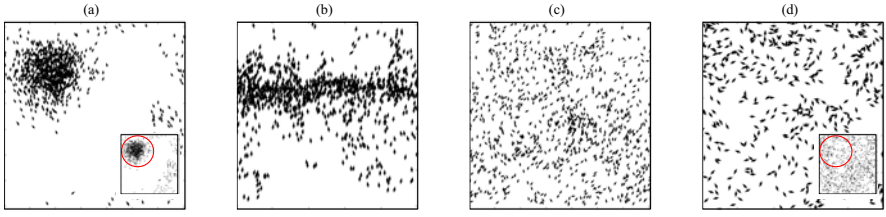

Figure 2.2: Four snapshots showing the different phases in the VM for different densities and noises. (a) Ordered motion for $M = 7$, $\rho_0 = 5$, $p = \eta = 0.8$. (b) Pattern formation for $M = 7$, $\rho_0 = 10$, $\eta = 0.3$. (c) Disordered motion $M = 7$, $\rho_0 = 5$, $\eta = 1.4$. (d) Random motion $M = 7$, $\rho_0 = 5$, $\eta = 3.7$.

an ordered way whereas, at maximum noise, particles move randomly, Fig. 2.2 can be helpful to visualize the order-disorder transition. It shows the velocity fields of the VM for $N = 1000$ particles and periodic boundary conditions. Varying the main parameters of the model, the different behaviors of the system flourished, (a) for small densities and noise, particles move in an ordered motion and in an unique cluster, (b) for box size higher than the critical, a variety of patterns are possible, for example, in this case the system moves in bands. In (c)-(d)

particles move randomly for higher densities.

2.3 Kinetic equation

In the limit of many particles, it is possible to derive a kinetic equation for the VM by assuming molecular chaos [16, 17, 19]. The kinetic approach presented in this Section assumes that all particles are independent and identically distributed before undergoing a collision described by Eqs. (2.1)-(2.2). This allows the definition of the N-particle probability density evolution equation

$$P(\theta^{(N)}, \mathbf{x}^{(N)} + \mathbf{v}^{(N)}, t+1) = \int_{[-\pi,\pi]^N} \int_{[-\pi,\pi]^N} P(\tilde{\theta}^{(N)}, \mathbf{x}^{(N)}, t) \times \left(\prod_{j=1}^{N} \hat{\delta}(\theta_j - \xi_j - \Phi_j(\tilde{\theta}^{(N)}, \mathbf{x}^{(N)}, t))\, d\xi_j\, d\theta_j \right). \quad (2.5)$$

Here, the notation $\theta^{(N)} = \theta_1, \ldots, \theta_N$ means the angles of the N particles and the same expressions are going to be used for the variables $\tilde{\theta}^{(N)}$, $\xi^{(N)}$, $\mathbf{x}^{(N)}$ and $\mathbf{v}^{(N)}$. The mean angle $\Phi_j = \mathrm{Arg}(\sum_{|\mathbf{x}_l - \mathbf{x}_j| < R_0} e^{i\tilde{\theta}_l})$ is defined in Eq. (2.2) and depends on the precollisional angles $(\tilde{\theta}^{(N)})$ of these particles that are inside the interaction circle R_0 centered in particle j. $\hat{\delta}(x) = \sum_{l=-\infty}^{\infty} \delta(x + 2\pi l)$ is a periodized delta function that incorporates the "collision rule" defined in Eq. (2.2), and the integrals over ξ_j and θ_j average over the noises and directions respectively and go from $-\pi$ to π.

Assuming that all particles are independent and identically distributed before undergoing a collision described by Eqs. (2.1) and (2.2) the probability distribution $P(\theta^{(N)}, \mathbf{x}^{(N)}, t)$ just before the collision step it can be approximated by a product of identical one-particle probability distributions. Then the N-particle probability density is product of N one-particle probability densities $f(\mathbf{x}_j, \theta_j, t)/N$, $j = 1, \ldots, N$, being $f(\mathbf{x}_j, \theta_j, t)$ the one-particle distribution function

$$P(\mathbf{x}_1, \theta_1, \ldots, \mathbf{x}_N, \theta_N, t) = \prod_{j=1}^{N} \frac{f(\mathbf{x}_j, \theta_j, t)}{N}, \quad (2.6)$$

Eq. (2.6) is the *molecular chaos assumption* first introduced by Boltzmann when deriving his transport equation [25]. Eq. (2.6) is appropriate for a time discrete model with unit time step if $r^* \ll 1$ in our non-dimensional units. For the VM, $r^* = 0$, and the molecular chaos assumption is very reasonable. Eq. (2.6) produces the formula

$$f(\theta, \mathbf{x}, t) = \int_{[-\pi,\pi]^N} \int_{[0,L]^{2N}} \sum_{j=1}^{N} \delta(\theta - \theta_j)\delta(\mathbf{x} - \mathbf{x}_j)\, P_N(\theta^{(N)}, \mathbf{x}^{(N)}, t) \prod_{i=1}^{N} (d\theta_i d\mathbf{x}_i), \quad (2.7)$$

where the $\mathbf{x}_j$ are integrated on the box $[0, L] \times [0, L]$. $\sum_{j=1}^{N} \delta(\theta - \theta_j)\delta(\mathbf{x} - \mathbf{x}_j)$ describes those

quantities depending on one position and time, and $P_N(\theta^{(N)}, \mathbf{x}^{(N)}, t) \prod_{i=1}^{N} d\theta_i \, d\mathbf{x}_i$ gives the section of those particles of the system within particle 1 is in the phase space element $d\theta_1 \, d\mathbf{x}_1$ around position $\mathbf{x}_1$ and angle θ_1 and at the same time particle 2 is in the phase space element $d\theta_2 \, d\mathbf{x}_2$ and so on. Once the one-particle distribution function has been calculated, the procedure to derive a kinetic equation for that function is to multiply the equation resulting from Eqs. (2.5) and (2.6) by $\sum_{j=1}^{N} \delta(\theta - \theta_j)\delta(\mathbf{x} - \mathbf{x}_j)$ (the microscopic one particle-density) and to integrate over all angles and positions. The result is [19]

$$f(\mathbf{x} + \mathbf{v}, \theta, t + 1) = C[f](\theta, \mathbf{x}, t), \quad \mathbf{v} = (\cos\theta, \sin\theta), \tag{2.8}$$

$$C[f] = \int_{-\pi}^{\pi} d\xi \, g(\xi) \sum_{n=1}^{N} \begin{pmatrix} N-1 \\ n-1 \end{pmatrix} \int \hat{\delta}(\theta - \xi - \Phi_1(\tilde{\theta}_1, \ldots, \tilde{\theta}_n)) \, f(\mathbf{x}, \tilde{\theta}_1, t) \left(1 - \frac{M_R(\mathbf{x}, t)}{N}\right)^{N-n} \times \prod_{i=2}^{n} \left[\int_{|\mathbf{x}_i - \mathbf{x}| < R_0} \frac{f(\mathbf{x}_i, \tilde{\theta}_i, t)}{N} d\tilde{\theta}_i d\mathbf{x}_i \right] d\tilde{\theta}_1, \tag{2.9}$$

$$\Phi_1(\tilde{\theta}_1, \ldots, \tilde{\theta}_n) = \mathrm{Arg}\left(\sum_{j=1}^{n} e^{i\tilde{\theta}_j} \right). \tag{2.10}$$

Now, let us explain in detail the terms in Eqs. (2.8), (2.9) and (2.10). n is the number of particles inside the interaction circle of radius R_0 about particle 1 (the latter included). $M_R(\mathbf{x}, t)$ defines the average number of particles inside an interaction circle about position $\mathbf{x}$, N is the total number of particles and ρ is the density respectively,

$$M_R(\mathbf{x}, t) = \int_{|\mathbf{x}' - \mathbf{x}| < R_0} \rho(\mathbf{x}', t) \, d\mathbf{x}', \quad \rho(\mathbf{x}, t) = \int_{-\pi}^{\pi} f(\mathbf{x}, \theta, t) \, d\theta, \quad \int \rho(\mathbf{x}, t) d\mathbf{x} = N. \tag{2.11}$$

The combinatorial factor in Eq. (2.9) counts the number of possible selections of neighbors excluding the latter $(n - 1)$ out of the $N - 1$ other particles. The factor $(1 - M_R/N)^{N-n}$ in Eq. (2.9) gives the probability that the particles $n + 1, \ldots, N$ are not within interaction distance of particle 1 and the factor $\prod_{i=1}^{n} \int_{|\mathbf{x}_i - \mathbf{x}| < R_0} f(\mathbf{x}_i, \theta_i, t) d\mathbf{x}_i$ is the probability that particles $2, \ldots, n$ be within interaction distance of particle 1 times their angular distribution, given that they are within the interaction distance. When we integrate Eq. (2.9) over θ, we find that the particle density immediately after collisions equals that before:

$$\int C[f](\theta, \mathbf{x}, t) \, d\theta = \rho(\mathbf{x}, t). \tag{2.12}$$

We may adopt two opposite approximations of the collision operator (2.9). For very diluted particle ensembles having small average density, $\rho_0 = N/L^2$, terms with $n \geq 2$ in Eq. (2.9)

provide negligible contributions. Then we get a binary collision operator

$$C_B[f] = \int_{-\pi}^{\pi} d\xi \frac{g(\xi)}{1+M_R}\left[\int_{-\pi}^{\pi} \hat{\delta}(\theta-\xi-\tilde{\theta}_1) f(\mathbf{x},\tilde{\theta}_1,t) d\tilde{\theta}_1 \right. \tag{2.13}$$
$$\left. + \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \hat{\delta}(\theta-\xi-\Phi_1(\tilde{\theta}_1,\tilde{\theta}_2))\, f(\mathbf{x},\tilde{\theta}_1,t) \times \left(\int_{|\mathbf{x}_2-\mathbf{x}|<R_0} f(\mathbf{x}_2,\tilde{\theta}_2,t) d\mathbf{x}_2\right) d\tilde{\theta}_2 d\tilde{\theta}_1 \right],$$

with $M_R = M_R(\mathbf{x},t)$, which has been normalized so that $\int_{-\pi}^{\pi} C_B[f] d\theta = \rho(\mathbf{x},t)$. Secondly, for larger densities and $n/N \ll 1$ as $N \to \infty$, the combinatorial factor times $(1-M_R/N)^{N-n}$ becomes

$$\frac{(N-1)!}{(n-1)!(N-n)!}\left(1-\frac{M_R}{N}\right)^{N-n} \sim \frac{N^{n-1}}{(n-1)!}\, e^{-M_R},$$

and (2.9) produces an Enskog-type collision operator [19]

$$C_E[f] = \int_{-\pi}^{\pi} d\xi\, g(\xi) e^{-M_R(\mathbf{x},t)} \sum_{n=1}^{\infty} \int_{[-\pi,\pi]^n} \frac{\hat{\delta}(\theta-\xi-\Phi_1(\tilde{\theta}_1,\ldots,\tilde{\theta}_n))}{(n-1)!} \tag{2.14}$$
$$\times f(\mathbf{x},\tilde{\theta}_1,t) \prod_{i=2}^{n}\left[\int_{|\mathbf{x}_i-\mathbf{x}|<R_0} f(\mathbf{x}_i,\tilde{\theta}_i,t) d\tilde{\theta}_i d\mathbf{x}_i\right] d\tilde{\theta}_1.$$

For active particles in a disordered state, the density $\rho(\mathbf{x},t)$ equals the constant average density, $\rho_0 = N/L^2$, and the uniform distribution function, $f_0 = \rho_0/(2\pi)$, is a fixed point of the collision operators:

$$C[f_0] = f_0, \quad C_B[f_0] = f_0, \quad C_E[f_0] = f_0. \tag{2.15}$$

Henceforth, we shall use the Enskog collision operator defined in Eq. (2.14).

2.4 Linear stability analysis

In this Section, we study the linear stability of the disordered solution having time-independent uniform particle density. Disorder is unstable if at least one eigenvalue has modulus larger than one. As a consequence of conservation of the number of particles, one is always an eigenvalue corresponding to a constant eigenfunction. To study the linear stability of the uniform distribution function, we insert $f = f_0 + \epsilon \tilde{f}(\theta,\mathbf{x},t)$, $\epsilon \ll 1$ into the kinetic equation and ignore quadratic terms, thereby obtaining Eq. (2.16). Here $\Phi_1(\tilde{\theta}_1,\ldots,\tilde{\theta}_n)$ and $M = \rho_0 \pi R_0^2$ have been defined in Eqs. (2.10) and (2.11) respectively. The separation of variables ansatz $\tilde{f}(\mathbf{x},\theta,t) = \tilde{F}(\mathbf{x},\theta)\, h(t)$ produces a discrete equation $h(t+1)/h(t) = Q$, where Q is the separation constant. Thus $h(t) = Q^t$. The equation for $\tilde{F}$ is an eigenvalue problem that yields Q. Moreover, $\tilde{F}$ is a periodic function of space and it can be written as a Fourier series expansion in plane waves, $e^{i\mathbf{K}\cdot\mathbf{x}}$, in which the components of the wave vectors are integer multiples of $2\pi/L$. In the limit as $L \to \infty$, the wave vectors $\mathbf{K}$ are real valued and

the Fourier series becomes a Fourier integral.

$$\tilde{f}(\theta, \mathbf{x}+\mathbf{v}, t+1) = \sum_{n=1}^{\infty} \frac{e^{-M}}{(n-1)!} \left(\frac{M}{2\pi}\right)^{n-1} \int_{-\pi}^{\pi} d\xi\, g(\xi) \int_{\tilde{\theta}} \hat{\delta}(\theta - \xi - \Phi_1) \times \left[\tilde{f}(\mathbf{x}, \tilde{\theta}_1, t) + \frac{n-1}{\pi R_0^2} \int_{|\mathbf{x}'-\mathbf{x}|<R_0} \tilde{f}(\mathbf{x}', \tilde{\theta}_1, t) d\mathbf{x}'\right] \times \prod_{l=1}^{n} d\tilde{\theta}_l - \frac{M}{2\pi} \frac{1}{\pi R_0^2} \int_{|\mathbf{x}'-\mathbf{x}|<R_0} \int_{-\pi}^{\pi} \tilde{f}(\theta', \mathbf{x}', t) d\mathbf{x}' d\theta'. \tag{2.16}$$

Setting $\tilde{F} = e^{i\mathbf{K}\cdot\mathbf{x}}\varphi(\theta; \mathbf{K})$, we are led to the separation of variables ansatz $\tilde{f} = Q^t e^{i\mathbf{K}\cdot\mathbf{x}}\varphi(\theta)$, where Q and $\varphi(\theta)$ are both functions of $\mathbf{K}$. This procedure of separation of variables is typically used in discussions of the Fourier-von Neumann stability of finite difference numerical methods for linear partial differential equations; see Ref. [26]. From Eq. (2.16), the integration of the plane wave on the disk of radius R_0 yields the eigenvalue problem for $\varphi(\theta)$

$$Q e^{i\mathbf{K}\cdot\mathbf{v}}\varphi - C^{(1)}[\varphi] = 0, \tag{2.17}$$

$$C^{(1)}[\varphi] = \frac{2J_1(|\mathbf{K}|R_0)}{|\mathbf{K}|R_0} \left[\sum_{n=1}^{\infty} \frac{e^{-M}}{(n-1)!} \left(\frac{M}{2\pi}\right)^{n-1} \left(n - 1 + \frac{|\mathbf{K}|R_0}{2J_1(|\mathbf{K}|R_0)}\right) \int_{-\pi}^{\pi} d\xi\, g(\xi) \int_{\tilde{\theta}} \hat{\delta}(\theta - \xi - \Phi_1)\varphi(\tilde{\theta}_1) \prod_{l=1}^{n} d\tilde{\theta}_l - \frac{M}{2\pi} \int_{-\pi}^{\pi} \varphi(\tilde{\theta}) d\tilde{\theta}\right]. \tag{2.18}$$

We have $C^{(1)}[1] = 1$, and therefore the uniform distribution $f_0 = \rho_0/(2\pi)$ solves Eq. (2.17) with $|\mathbf{K}| = 0$ and $Q = 1$. We now seek non-constant solutions of Eq. (2.17) by inserting the Fourier expansion $\varphi(\theta) = \sum_{j=-\infty}^{\infty} \varphi_j e^{ij\theta}$. We find

$$\sum_{j=-\infty}^{\infty} \left[Q(e^{i\mathbf{K}\cdot(\cos\theta, \sin\theta)})_j - C^{(1)}[\varphi]_j\right] e^{ij\theta} = 0, \tag{2.19}$$

from which we obtain the eigenvalue problem:

$$\sum_{l=-\infty}^{\infty} \{C^{(1)}[e^{ij\theta}]_j \delta_{jl} - Q(e^{i\mathbf{K}\cdot(\cos\theta, \sin\theta) + ij\theta})_l\}\varphi_l = 0. \tag{2.20}$$

Here the subscripts j and l indicate that $f(\theta)_j$ and $f(\theta)_l$ are the coefficients of the respective harmonics in the Fourier series of the function $f(\theta)$, and we have used $C^{(1)}[e^{il\theta}]_j = 0$ for

$j \neq l$ [19]. Equivalently, $1/Q$ are the eigenvalues of a matrix $\mathcal{M}_{jl}$:

$$\mathcal{M}_{jl} = \frac{(e^{i\mathbf{K}\cdot(\cos\theta,\sin\theta)+ij\theta})_l}{C^{(1)}[e^{ij\theta}]_j}, \quad (e^{i\mathbf{K}\cdot(\cos\theta,\sin\theta)+ij\theta})_j = (e^{i\mathbf{K}\cdot(\cos\theta,\sin\theta)})_0 = J_0(|\mathbf{K}|).$$

$$C^{(1)}[e^{ij\theta}]_j = \left(\int_{-\pi}^{\pi} e^{-ij\xi} g(\xi)\, d\xi\right) \sum_{n=1}^{\infty} \frac{M^{n-1}e^{-M}}{(n-1)!} \left[(n-1)\frac{2J_1(|\mathbf{K}|R_0)}{|\mathbf{K}|R_0} + 1\right] \times \int_{[-\pi,\pi]^n} e^{ij(\tilde{\theta}_1 - \Phi_1)} \prod_{l=1}^{n} \frac{d\tilde{\theta}_l}{2\pi}. \tag{2.21}$$

If $\mathbf{K} = K(0,1)$, the off-diagonal matrix elements are $J_{l-j}(K)/C^{(1)}[e^{ij\theta}]_j$.

2.4.1 Space independent eigenfunctions

In this Section, we study solutions that bifurcate from disorder with zero wave number, which correspond to bifurcations for box sizes below critical. For $|\mathbf{K}| = 0$, Eq. (2.20) produces the following eigenvalues and eigenfunctions:

$$Q_j = C^{(1)}[e^{ij\theta}]_j, \quad \varphi_j(\theta) = e^{ij\theta}, \quad (\varphi_j)_l = \delta_{lj}, \tag{2.22}$$

with $j, l = 1, 2, \ldots$. The disordered state is stable when $|Q_j| \leq 1$ for all j, and unstable if $|Q_j| > 1$ for some j. $\varphi_0 = 1$ is one eigenfunction corresponding to eigenvalue $Q_0 = 1$. The eigenvalue with largest modulus for $j \neq 0$ is Q_1, which, for large M, becomes [19]

$$Q_1 \sim \frac{\sqrt{\pi M}}{2} \int_{-\pi}^{\pi} e^{-i\xi} g(\xi) d\xi = \frac{\sqrt{\pi M}}{\eta} \sin\frac{\eta}{2}. \tag{2.23}$$

Other eigenvalues have moduli smaller than 1 in the limit as $M \to \infty$, as shown in Appendix 2.A.

2.4.2 Perturbation of eigenvalues for small off-diagonal elements

We do not know how to find the eigenvalues of the matrix $\mathcal{M}_{jl}$, given by Eq. (2.21), for general nonzero $\mathbf{K}$. However, the diagonal entries $\mathcal{M}_{jj}$ are proportional to $J_0(|\mathbf{K}|) = 1 + O(|\mathbf{K}|)$ (as $|\mathbf{K}| \to 0$), whereas the off-diagonal elements of $\mathcal{M}_{jl}$ with $j \neq l$ vanish for zero wave vector. Thus, for small $|\mathbf{K}|$, the off-diagonal elements of the matrix $\mathcal{M}_{l,j}$ of Eq. (2.21) are small compared to the diagonal elements. Assuming that the matrix $\mathcal{M}_{jl}$ is equal to the matrix of its diagonal elements plus a small perturbation, we can use regular perturbation theory to calculate its eigenvalues. The situation is analogous to the usual perturbation theory of eigenvalues in non-relativistic Quantum Mechanics as explained in Ref [27]. The first order correction to the eigenvalues of $\mathcal{M}_{jl}$ is given by the diagonal elements of the perturbation

matrix. However, this perturbation matrix is $\mathcal{M}_{jl}$ minus its diagonal, and therefore it has zero diagonal elements. Thus, we need to calculate the eigenvalues by using second-order perturbation theory for a perturbation matrix comprising the off-diagonal elements of $\mathcal{M}_{jl}$. We obtain [cf. Eq. (38.10) of Ref [27]]:

$$\frac{1}{Q_j} = \frac{J_0(|\mathbf{K}|)}{C^{(1)}[e^{ij\theta}]_j} - \sum_{l\neq j} \frac{\mathcal{M}_{l,j}\mathcal{M}_{j,l}}{\frac{J_0(|\mathbf{K}|)}{C^{(1)}[e^{ij\theta}]_l} - \frac{J_0(|\mathbf{K}|)}{C^{(1)}[e^{ij\theta}]_j}}, \tag{2.24}$$

which holds for $0 \leq |\mathbf{K}| < \gamma_{1,0}$ [$\gamma_{1,0} \approx 2.4048$ is the first zero of the Bessel function $J_0(x)$].

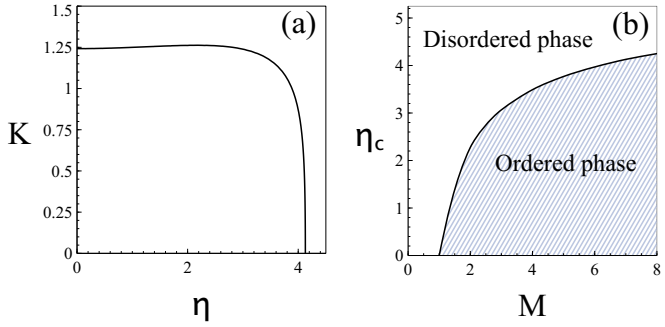

Figure 2.3: (a) Critical wave vector $K = |\mathbf{K}|$ versus η for $M = 7$, $\rho_0 = 5$. This curve is obtained by solving $|Q_1(K,\eta)| = 1$, where $Q_1(K,\eta)$ is given by Eq. (2.26). (b) At zero wave number, critical noise versus M is plotted.

We are interested in the eigenvalue close to $1/Q_1$, because Q_1 has the largest modulus for $|\mathbf{K}| = 0$. It is approximately given by

$$\frac{1}{Q_1} \approx \frac{J_0(|\mathbf{K}|)}{C^{(1)}[e^{i\theta}]_1} - \frac{\mathcal{M}_{2,1}\mathcal{M}_{1,2}}{\frac{J_0(|\mathbf{K}|)}{C^{(1)}[e^{i2\theta}]_2} - \frac{J_0(|\mathbf{K}|)}{C^{(1)}[e^{i\theta}]_1}}, \tag{2.25}$$

in which we have ignored higher order terms having $|l-1| > 1$. For $\mathbf{K} = K(0,1)$, these terms are proportional to $[J_{l-1}(K)]^2 = O(|\mathbf{K}|^{2(l-1)})$ (with $l > 2$). Thus, compared with the last term in Eq. (2.25), which is $O(|\mathbf{K}|^2)$, they can be ignored in the limit as $|\mathbf{K}| \to 0$. This also occurs for general $\mathbf{K}$. We have $(e^{i\mathbf{K}\cdot(\cos\theta,\sin\theta)})_0 = J_0(|\mathbf{K}|)$, and $(e^{i\mathbf{K}\cdot(\cos\theta,\sin\theta)+i\theta})_0 = ((\cos\theta + i\sin\theta)(e^{i\mathbf{K}\cdot(\cos\theta,\sin\theta)})_0$ produces

$$\mathcal{M}_{2,1} = -\frac{1}{C^{(1)}[e^{i2\theta}]_2}(i\frac{\partial}{\partial K_x} - \frac{\partial}{\partial K_y})J_0(|\mathbf{K}|) = -\frac{iK_x - K_y}{|\mathbf{K}|\,C^{(1)}[e^{i2\theta}]_2}J_1(|\mathbf{K}|) \Longrightarrow$$
$$\mathcal{M}_{1,2}\mathcal{M}_{2,1} = -\frac{[J_1(|\mathbf{K}|)]^2}{C^{(1)}[e^{i\theta}]_1 C^{(1)}[e^{i2\theta}]_2}.$$

Thus, Eq. (2.25) becomes

$$Q_1 \approx \frac{C^{(1)}[e^{i\theta}]_1}{J_0(|\mathbf{K}|) + \frac{[J_1(|\mathbf{K}|)]^2 C^{(1)}[e^{i\theta}]_1}{J_0(|\mathbf{K}|)\,(C^{(1)}[e^{i\theta}]_1 - C^{(1)}[e^{i2\theta}]_2)}}. \tag{2.26}$$

According to Eqs. (2.21) and (2.26), $Q_1 = Q_1(\eta, |\mathbf{K}|)$. Then the equation $Q_1 = 1$ may have different solution branches $\eta(|\mathbf{K}|)$ for nonzero wave number, depending on the parameters M and ρ_0. Fig. 2.3 (a) show the solution curve of $|Q_1(|\mathbf{K}|, \eta)| = 1$ for $\rho_0 = 5$. In this case, the uniform distribution becomes first unstable for the noise corresponding to zero wave number. Within our approximations, this justifies that the largest value of the multiplier Q_1 is attained at zero wave number. The case of critical eigenvalue $|Q_1(|\mathbf{K}|, \eta)| = 1$ for zero wave number, is shown in Fig. 2.3 (b) where the critical noise versus M is plotted.

2.5 Bifurcation theory

In this Section, we analyze flocking in the two-dimensional (2D) VM by using systematically bifurcation theory for its Enskog kinetic equation. As justified in Section 2.4, the largest multiplier is $Q_1 = 1$ corresponding to $\mathbf{K} = 0$. The solution of the linearized equation

$$\mathcal{L} f^{(1)} \equiv f^{(1)}(\theta, t+1, \mathbf{X}, T) - C^{(1)}[f^{(1)}](\theta, t, \mathbf{X}, T) = 0, \tag{2.27}$$

is

$$f^{(1)}(\theta, t, \mathbf{X}, T, \epsilon) = \frac{r(\mathbf{X}, T; \epsilon)}{2\pi} + A(\mathbf{X}, T; \epsilon) e^{i\theta} + \text{cc}, \tag{2.28}$$

$$\mathbf{X} = \epsilon \mathbf{x}, \quad T = \epsilon t. \tag{2.29}$$

Here cc means the complex conjugate of the preceding term. We do not need to include more terms in (2.28) because the other modes decay rapidly in the fast time scale t. The first term in Eq. (2.28) is a space dependent disturbance of the uniform density, whereas the complex amplitude of the second term corresponds to a vector current density, as we will show below in Eqs. (2.41)-(2.43). We anticipate crossover scalings and therefore we shall use the Chapman-Enskog method [22, 23, 24, 28]. The Chapman-Enskog ansatz is [22, 23, 24],

$$f(\theta, \mathbf{x}, t; \epsilon) = f_0 + \epsilon f^{(1)} + \sum_{j=2}^{\infty} \epsilon^j f^{(j)}(\theta, t; r, A, \overline{A}), \tag{2.30}$$

$$\frac{\partial r}{\partial T} = \mathcal{R}^{(0)}(r, A, \overline{A}) + \epsilon \mathcal{R}^{(1)}(r, A, \overline{A}) + O(\epsilon^2), \tag{2.31}$$

$$\frac{\partial A}{\partial T} = \mathcal{A}^{(0)}(r, A, \overline{A}) + \epsilon \mathcal{A}^{(1)}(r, A, \overline{A}) + O(\epsilon^2). \tag{2.32}$$

where $\overline{A}$ is the complex conjugate of A. We select a scaling $\eta = \eta_c + \epsilon^2\eta_2$, which is appropriate for the case of the pitchfork bifurcation that occurs for space independent solutions. We will explain later what happens for a different choice of scaling. Inserting Eqs. (2.30)-(2.32) into Eqs. (2.8) and (2.14), we obtain the following hierarchy of equations

$$\mathcal{L}f^{(2)} = C^{(2)}[f^{(1)}, f^{(1)}] - \mathbf{v}\cdot\nabla_X f^{(1)} - \frac{\mathcal{R}^{(0)}}{2\pi} - \mathcal{A}^{(0)}e^{i(\theta+\Omega t+\Omega)} + \text{cc}, \tag{2.33}$$

$$\mathcal{L}f^{(3)} = C^{(3)}[f^{(1)}, f^{(1)}, f^{(1)}] + 2C^{(2)}[f^{(1)}, f^{(2)}] - \mathbf{v}\cdot\nabla_X f^{(2)} - \frac{\mathcal{R}^{(1)}}{2\pi} - \mathcal{A}^{(1)}e^{i(\theta+\Omega t+\Omega)} + \text{cc} - \frac{1}{2}\left(\frac{\partial}{\partial T} + \mathbf{v}\cdot\nabla_X\right)^2 f^{(1)} + \eta_2\frac{\partial}{\partial\eta}C^{(1)}[f^{(1)}]. \tag{2.34}$$

etc. In these equations, we have the following terms

$$C^{(2)}[\varphi,\varphi] = \frac{\pi R_0^2}{2}\left[\sum_{n=2}^{\infty}\frac{ne^{-M}}{(n-2)!}\left(\frac{M}{2\pi}\right)^{n-2}\int_{-\pi}^{\pi}d\xi g(\xi)\int\hat{\delta}(\theta-\xi-\Phi_1)\varphi(\tilde{\theta}_1)\varphi(\tilde{\theta}_2)\prod_{l=1}^{n}d\tilde{\theta}_l\right.$$
$$-2\left(\int_{-\pi}^{\pi}\varphi(\theta_1)d\theta_1\right)\sum_{n=1}^{\infty}\frac{ne^{-M}}{(n-1)!}\left(\frac{M}{2\pi}\right)^{n-1}\int_{-\pi}^{\pi}d\xi g(\xi)\int\hat{\delta}(\theta-\xi-\Phi_1)\varphi(\tilde{\theta}_1)\prod_{l=1}^{n}d\tilde{\theta}_l$$
$$\left.+\frac{M}{2\pi}\left(\int_{-\pi}^{\pi}\varphi(\theta_1)d\theta_1\right)^2\right], \tag{2.35}$$

$$C^{(3)}[\varphi,\varphi,\varphi] = \frac{\pi^2 R_0^4}{6}\left[\sum_{n=3}^{\infty}\frac{ne^{-M}}{(n-3)!}\left(\frac{M}{2\pi}\right)^{n-3}\int_{-\pi}^{\pi}d\xi g(\xi)\int\hat{\delta}(\theta-\xi-\Phi_1)\varphi(\tilde{\theta}_1)\varphi(\tilde{\theta}_2)\varphi(\tilde{\theta}_3)\prod_{l=1}^{n}d\tilde{\theta}_l\right.$$
$$-3\left(\int_{-\pi}^{\pi}\varphi(\theta_3)d\theta_3\right)\sum_{n=2}^{\infty}\frac{ne^{-M}}{(n-2)!}\left(\frac{M}{2\pi}\right)^{n-2}\int_{-\pi}^{\pi}d\xi g(\xi)\int\hat{\delta}(\theta-\xi-\Phi_1)\varphi(\tilde{\theta}_1)\varphi(\tilde{\theta}_2)\prod_{l=1}^{n}d\tilde{\theta}_l$$
$$+3\left(\int_{-\pi}^{\pi}\varphi(\theta_2)d\theta_2\right)^2\sum_{n=1}^{\infty}\frac{ne^{-M}}{(n-1)!}\left(\frac{M}{2\pi}\right)^{n-1}\int_{-\pi}^{\pi}d\xi g(\xi)\int\hat{\delta}(\theta-\xi-\Phi_1)\varphi(\tilde{\theta}_1)\prod_{l=1}^{n}d\tilde{\theta}_l$$
$$\left.-\frac{M}{2\pi}\left(\int_{-\pi}^{\pi}\varphi(\theta_1)d\theta_1\right)^3\right]. \tag{2.36}$$

and so on. Note that $C_E[f_0+\epsilon\tilde{\rho}] = f_0+\epsilon\tilde{\rho}$ and $C^{(1)}[\tilde{\rho}] = \tilde{\rho}$ for constant $\tilde{\rho}$ imply $C^{(2)}[1,1] = C^{(3)}[1,1,1] = 0$, which can be checked from Eqs. (2.35)-(2.36). The solvability conditions for non-homogeneous equations of the hierarchy is that their right hand sides be orthogonal to the solutions of the homogeneous equation $\mathcal{L}\varphi = 0$, namely 1 and $e^{i\theta}$, using the scalar product

$$\langle f(\theta), g(\theta)\rangle = \int_{-\pi}^{\pi}\overline{f(\theta)}g(\theta)\,d\theta. \tag{2.37}$$

We now proceed to derive the amplitude equations. We insert Eq. (2.28) into Eq. (2.33) and impose that its right hand side be orthogonal to 1 and to $e^{i\theta}$, thereby obtaining

$$\begin{aligned}\mathcal{R}^{(0)} &= -2\pi \mathrm{Re}\left[\left(\frac{\partial}{\partial X} + i\frac{\partial}{\partial Y}\right)A\right],\\ \mathcal{A}^{(0)} &= \frac{1}{\pi}C^{(2)}[1, e^{i\theta}]_1 rA - \frac{1}{4\pi}\left(\frac{\partial}{\partial X} - i\frac{\partial}{\partial Y}\right)r.\end{aligned} \tag{2.38}$$

Then Eq. (2.33) has the solution

$$f^{(2)}(\theta, t, \mathbf{X}, T) = \left[\frac{A^2 C^{(2)}[e^{i\theta}, e^{i\theta}]_2}{1 - C^{(1)}[e^{i2\theta}]_2} - \frac{\left(\frac{\partial}{\partial X} - i\frac{\partial}{\partial Y}\right)A}{2(1 - C^{(1)}[e^{i2\theta}]_2)}\right] e^{i2\theta} + \mathrm{cc}. \tag{2.39}$$

We now insert Eqs. (2.28) and (2.39) into Eq. (2.34), and impose the solvability conditions to the resulting equation. The solvability conditions for this equation produce:

$$\begin{aligned}\mathcal{R}^{(1)} &= -\mathrm{Re}\left[C^{(2)}[1, e^{i\theta}]_1\left(\frac{\partial}{\partial X} + i\frac{\partial}{\partial Y}\right)rA\right],\\ \mathcal{A}^{(1)} &= [\Gamma(r) - \mu|A|^2]A + \delta\Delta_X A + \gamma_1\left(\frac{\partial}{\partial X} + i\frac{\partial}{\partial Y}\right)A^2 + \gamma_2\overline{A}\left(\frac{\partial}{\partial X} - i\frac{\partial}{\partial Y}\right)A\\ &+ \gamma_3 A\left(\frac{\partial}{\partial X} - i\frac{\partial}{\partial Y}\right)\overline{A} + \frac{\gamma_3}{8\pi^2}\left(\frac{\partial}{\partial X} - i\frac{\partial}{\partial Y}\right)r^2.\end{aligned} \tag{2.40}$$

In terms of the mean current density $\mathbf{w}$ defined as

$$A = \frac{w_x - i w_y}{2\pi}, \quad \mathbf{w} = (w_x, w_y), \tag{2.41}$$

Eqs. (2.31), (2.32), (2.38) and (2.40) can be written as

$$\frac{\partial r}{\partial T} + \nabla_X \cdot \left[\left(1 + \frac{\epsilon\gamma_3 r}{2\pi}\right)\mathbf{w}\right] = 0, \tag{2.42}$$

$$\begin{aligned}\frac{\partial \mathbf{w}}{\partial T} =& -\frac{1}{2}\nabla_X\left[\left(1 - \frac{\epsilon\gamma_3}{4\pi}r\right)r + \frac{\epsilon(2\gamma_1 - \gamma_2 - \gamma_3)}{2\pi}|\mathbf{w}|^2\right] + \epsilon\frac{2\gamma_1 + \gamma_2 - \gamma_3}{2\pi}(\mathbf{w}\cdot\nabla_X)\mathbf{w} +\\ &\epsilon\frac{2\gamma_1 - \gamma_2 + \gamma_3}{2\pi}\mathbf{w}(\nabla_X \cdot \mathbf{w}) + \epsilon\delta\nabla_X^2\mathbf{w} + \left[\frac{\gamma_3 r}{\pi} + \epsilon\left(\Gamma(r) - \frac{\mu|\mathbf{w}|^2}{4\pi^2}\right)\right]\mathbf{w}.\end{aligned} \tag{2.43}$$

The coefficients in Eqs. (2.42) and (2.43) are calculated in Appendix 2.B.

In particular, for $\epsilon = 0$, Eq. (2.42) is the continuity equation for a density variable r and a current density $\mathbf{w}$, which explains the name of the latter variable. The overall density of particles equals the average density N/L^2, which implies the following constraint for $r(\mathbf{X}, T)$:

$$\int r(\mathbf{X}, T)\, d\mathbf{X} = 0. \tag{2.44}$$

Remark 1. All the parameters in Eqs. (2.42)-(2.43) are real numbers. For $r = 0$ and $\gamma_3 = 0$, these equations are exactly equivalent to the amplitude equation (132) of Ref. [19].

Remark 2. For $r \neq 0$, Eqs. (2.42)-(2.43) are not equivalent to those in [19]. The differences are due to two inconsistencies in Ihle's approach [19]. Firstly, Ihle did not expand the eigenvalue $Q_1 = C^{(1)}[e^{i\theta}]_1$ in powers of ϵ. Instead, he introduced $(Q_1 - 1)\mathbf{w}$ as a linear term in the amplitude equation for $\mathbf{w}$ [cf. Eq. (2.43)]. This is equivalent to setting $\gamma_3 = 0$ and $\Gamma(r) = Q_1 - 1$ in Eq. (2.43). Thus, Ihle's equations do not contain the quadratic term proportional to $r\mathbf{w}$ appearing in Eq. (2.43). Secondly, we have used a consistent perturbation scheme in the parameter ϵ whereas the Chapman-Enskog procedure of Ref. [19] mixes different orders in ϵ. For instance, the tensors $\mathbf{\Omega}_1$ and ω_3 defined in Eqs. (117)-(119) and (121) of Ref. [19] are $O(\epsilon^2)$ whereas $\mathbf{\Omega}_j = O(\epsilon^3)$ for $j = 2, 4, 5$. However, all these tensors enter in the equation for the current density, Eq. (130) of Ref. [19], with equal footing. On the other hand, for uniform time-independent density, all terms in Ihle's Eq. (132) are of order ϵ^3 provided $\partial/\partial t = O(\epsilon^2)$, $\nabla = O(\epsilon)$, $\lambda - 1 = O(\epsilon^2)$ and $\mathbf{w} = O(\epsilon)$.

Remark 3. The equations of motion for the average velocity in bird flocking proposed by Toner and Tu do not contain the quadratic term $r\mathbf{w}$ in Eq. (2.43) [7, 8].

Remark 4. What happens if the noise scales differently with ϵ? We have chosen a parabolic scaling, $\eta = \eta_c + \epsilon^2\eta_2$. The only case that affects differently the outcome in Eqs. (2.42)-(2.43) is $\eta = \eta_c + \epsilon\eta_1$. Then $\mathcal{A}^{(0)}$ has to include an additional term $\eta_1 Q_\eta$ in Eq. (2.38). This means we have to write $r^* = r + \pi\eta_1 Q_\eta / C^{(2)}[1, e^{i\theta}]_1$ instead of r in Eqs. (2.42)-(2.43). In Eq. (2.43), we also have to replace the term $Q_{\eta\eta}\eta_1^2/2$ (where $Q_{\eta\eta} = \partial^2 Q_1/\partial\eta^2|_{\eta=\eta_c}$) instead of $Q_\eta\eta_2$ in $\Gamma(r)$. The resulting $\Gamma(r^*)$ is negative (at least for the numerical values $M = 7$, $\rho_0 = 5$ used in our simulations). We will not obtain a consistent stationary space independent solution of Eqs. (2.42)-(2.43) unless $r^* = O(\epsilon)$, which will take us back to the parabolic scaling of the noise.

2.6 Space-independent amplitude equation

For space independent A and $r = 0$, Eq. (2.32) is the typical pitchfork amplitude equation

$$\frac{\partial A}{\partial(\epsilon^2 t)} = (\eta_2 Q_\eta - \mu|A|^2)A. \tag{2.45}$$

Note that time scales as the square of space (diffusive scaling). As $\mu > 0$ and $Q_\eta < 0$, the stationary solution

$$A = \sqrt{\frac{\eta_2 Q_\eta}{\mu}}\, e^{i\Upsilon}, \quad \Upsilon \in \mathbb{R}, \tag{2.46}$$

is stable and it exists for $\eta < \eta_c$. In this region, the uniform distribution f_0 is unstable because $Q_1 > 1$. Thus the transition from incoherence to order is a supercritical pitchfork bifurcation, as depicted in Fig. 2.4. It represents the polarization given by the modulus of the complex

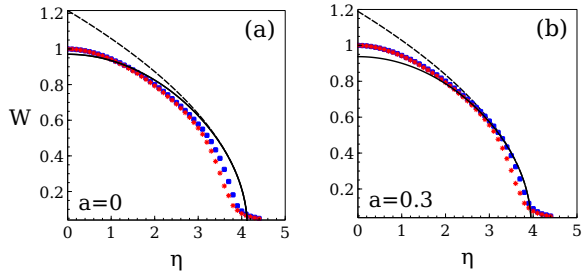

Figure 2.4: (a) Polarization W versus η for $M = 7$, $N = 1000$, and $\rho_0 = 10$ (blue squares) or $\rho_0 = 5$ (red asterisks). Dashed and solid lines correspond to (2.48) and (2.49), respectively. (b) Same graph with the critical noise shifted according to Eq. (2.66) with $a = (\epsilon/\rho_0)\sqrt{\langle \tilde{r}^2 \rangle/6} = 0.3$.

parameter (2.4). According to Eq. (2.28), the order parameter is

$$Z = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j} = \frac{1}{\rho_0 L^2}\int e^{i\theta} f(\theta, \mathbf{x}, t)\, d\theta d\mathbf{x} \sim \frac{2\pi\epsilon}{\rho_0 \epsilon^2 L^2}\int \overline{A(\mathbf{X}, T)}\, d\mathbf{X}. \tag{2.47}$$

For the uniform solution given by Eq. (2.46), the order parameter is

$$Z_0 \sim \frac{2\pi}{\rho_0}\sqrt{\frac{(\eta - \eta_c)Q_\eta}{\mu}}\, e^{-i\Upsilon}. \tag{2.48}$$

Near η_c, we have $Q_1 \sim 1 + Q_\eta(\eta - \eta_c)$. Then we can replace $Q_\eta(\eta - \eta_c)$ by $Q_1 - 1$ in Eq. (2.48), thereby getting a formula that holds for larger values of $|\eta - \eta_c|$ (cf. Fig. 2.4):

$$Z_0 \sim \frac{2\pi}{\rho_0}\sqrt{\frac{Q_1 - 1}{\mu}}\, e^{-i\Upsilon}. \tag{2.49}$$

We observe that the values obtained by direct simulations of the VM tend to the predicted solution as we increase the density ρ_0 from 5 to 10 in Fig. 2.4(a). The change in convexity of the curve near η_c is due to finite size effects. Moreover, Fig. 2.4(b) shows that the critical noise found in the simulations of the VM is shifted to $\eta_c - a^2$ with $a = 0.3$ for $\rho_0 = 10$.

2.7 Space-dependent amplitude equations

For nonzero r and space dependent A, we need to consider the space-dependent Eqs. (2.42)-(2.43). The leading order equations for $\epsilon = 0$ are

$$\frac{\partial r}{\partial T} + \nabla_X \cdot \mathbf{w} = 0, \tag{2.50}$$

$$\frac{\partial \mathbf{w}}{\partial T} = -\frac{1}{2}\nabla_X r + \frac{\gamma_3}{\pi} r\mathbf{w}. \tag{2.51}$$

In these equations, space and time scale in the same way as $\mathbf{X} = \epsilon\mathbf{x}$ and $T = \epsilon t$ (hyperbolic or convective scaling). The order parameter (2.47) can now be used to define a vector order parameter, $\mathbf{Z}$ =(ReZ,ImZ):

$$\mathbf{Z} \sim \frac{\epsilon}{\rho_0 L^2} \int \mathbf{w}(\mathbf{X}, T)\, dx\, dy, \tag{2.52}$$

because $\overline{A(\mathbf{X}, T)} = [w_x(\mathbf{X}, T) + i w_y(\mathbf{X}, T)]/(2\pi)$, according to Eq. (2.41). Ignoring the nonlinear term in Eq. (2.51) (small initial data), we eliminate the current density and obtain the linear wave equation with velocity $1/\sqrt{2}$. Then $\nabla_X \cdot \mathbf{w}$ obeys the same wave equation and $\nabla_X \times \mathbf{w}$ is independent of time. The overall constraint (2.44) holds for all time provided it does so initially. Space independent solutions of this system produce a current density that increases with time if $r > 0$ and decreases if $r < 0$. It seems that shock waves may form after a finite time.

2.7.1 Linearized 2D equations for convective scaling

Figs. 2.4 shows that, for the long times employed in direct simulations of the VM, there is agreement between the simulations and the uniform solution (2.46) of Eqs. (2.42)-(2.43),

$$r_0 = 0,\ \mathbf{w}_0 = 2\pi\sqrt{\frac{\eta_2 Q_\eta}{\mu}}\mathbf{e}_\Upsilon,\ \mathbf{e}_\Upsilon = (\cos\Upsilon, -\sin\Upsilon). \tag{2.53}$$

However, very close to η_c, the separation between the hyperbolic and parabolic time scales has appreciable effects. To uncover them, we linearize Eqs. (2.50)-(2.51) about Eq. (2.53), thereby obtaining

$$\frac{\partial \tilde{r}}{\partial T} + \nabla_X \cdot \tilde{\mathbf{w}} = 0, \tag{2.54}$$

$$\frac{\partial \tilde{\mathbf{w}}}{\partial T} = -\frac{1}{2}\nabla_X \tilde{r} + \frac{\gamma_3}{\pi}\mathbf{w}_0 \tilde{r}. \tag{2.55}$$

By differentiating Eq. (2.54) and eliminating $\tilde{\mathbf{w}}$ by means of Eq. (2.55), we find the wave equation:

$$\frac{\partial^2 \tilde{r}}{\partial T^2} = \frac{1}{2}\nabla_X^2 \tilde{r} - \frac{\gamma_3}{\pi}\,\mathbf{w}_0 \cdot \nabla_X \tilde{r}. \tag{2.56}$$

For periodic boundary conditions, we can solve this equation by writing $\tilde{r}(\mathbf{X}, T)$ as a Fourier series on the square box of size L. Then we can find $\tilde{\mathbf{w}}(\mathbf{X}, T)$ from Eq. (2.55). However, the gradient term produces a combination of factors exponentially increasing with time and factors exponentially decreasing with time. It is then hard to predict the long time behavior of the solutions. We can obtain an equivalent formulation by using the change of variable

$$\tilde{r} = e^{\gamma_3 \mathbf{w}_0 \cdot \mathbf{X}/\pi} R, \tag{2.57}$$

to eliminate the gradient term in Eq. (2.56), thereby producing the Klein-Gordon equation:

$$\frac{\partial^2 R}{\partial T^2} = \frac{1}{2}\nabla_X^2 R - \frac{\gamma_3^2|\mathbf{w}_0|^2}{2\pi^2}\, R. \tag{2.58}$$

For periodic boundary conditions, $R(\mathbf{X},T) = \sum_{n,m} R_{n,m}(T)e^{i\mathbf{k}_{n,m}\cdot\mathbf{X}}$, and the coefficients $R_{n,m}(T)$ solve the equation of a linear oscillator with frequency

$$\omega_{n,m} = \sqrt{\frac{1}{2}|\mathbf{k}_{n,m}|^2 + \frac{\gamma_3^2|\mathbf{w}_0|^2}{2\pi^2}},\ \mathbf{k}_{n,m} = \frac{2\pi}{\epsilon L}(n,m). \tag{2.59}$$

Note that the frequency $\omega_{n,m}$ mixes frequencies corresponding to the acoustic velocity $1/\sqrt{2}$ [cf. $|\mathbf{w}_0| = 0$ in Eq. (2.59)] with the fundamental mode of frequency $\gamma_3|\mathbf{w}_0|/(\sqrt{2}\pi)$ corresponding to $n = m = 0$. We now solve the equation for $R_{n,m}(T)$ and then reconstruct $\tilde{r}(\mathbf{X},T)$ and $\tilde{\mathbf{w}}(\mathbf{X},T)$ from Eqs (2.50)-(2.51) and (2.57). The results are

$$\tilde{r}(\mathbf{X},T) = \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\left[R_{n,m}(0)\cos(\omega_{n,m}T) + \frac{\dot{R}_{n,m}(0)}{\omega_{n,m}}\sin(\omega_{n,m}T)\right] e^{(i\mathbf{k}_{n,m}+\frac{\gamma_3\mathbf{w}_0}{\pi})\cdot\mathbf{X}}, \tag{2.60}$$

$$\begin{aligned}\tilde{\mathbf{w}}(\mathbf{X},T) = \frac{1}{2}\sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}&\left(i\mathbf{k}_{n,m} - \frac{\gamma_3\mathbf{w}_0}{\pi}\right)\frac{e^{(i\mathbf{k}_{n,m}+\frac{\gamma_3\mathbf{w}_0}{\pi})\cdot\mathbf{X}}}{\omega_{n,m}}\\ &\times\left[R_{n,m}(0)\sin(\omega_{n,m}T) - \frac{\dot{R}_{n,m}(0)}{\omega_{n,m}}\cos(\omega_{n,m}T)\right] + \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\mathbf{C}_{m,n}\, e^{i\mathbf{k}_{n,m}\cdot\mathbf{X}},\end{aligned} \tag{2.61}$$

$$\begin{aligned}R_{n,m}(0) = \frac{1}{\epsilon^2L^2}\int_0^{\epsilon L}\int_0^{\epsilon L} e^{-(i\mathbf{k}_{n,m}+\frac{\gamma_3\mathbf{w}_0}{\pi})\cdot\mathbf{X}}\,\tilde{r}(\mathbf{X},0)\,d\mathbf{X} =\\ \sum_{l=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}\frac{\tilde{r}_{l,j}(0)\,(e^{\frac{\epsilon\gamma_3 w_{0x}L}{\pi}}-1)(e^{\frac{\epsilon\gamma_3 w_{0y}L}{\pi}}-1)}{\left[\frac{\epsilon\gamma_3 w_{0x}L}{\pi} + i2\pi(l-n)\right]\left[\frac{\epsilon\gamma_3 w_{0y}L}{\pi} + i2\pi(j-m)\right]},\end{aligned} \tag{2.62}$$

$$\dot{R}_{n,m}(0) = -\frac{1}{\epsilon^2L^2}\int_0^{\epsilon L}\int_0^{\epsilon L} e^{-(i\mathbf{k}_{n,m}+\frac{\gamma_3\mathbf{w}_0}{\pi})\cdot\mathbf{X}}\left(i\mathbf{k}_{n,m} + \frac{\gamma_3\mathbf{w}_0}{\pi}\right)\cdot\tilde{\mathbf{w}}(\mathbf{X},0)\,dX\,dY, \tag{2.63}$$

$$\begin{aligned}\mathbf{C}_{n,m} = \tilde{\mathbf{w}}_{n,m}(0) + \frac{1}{2}\sum_{l=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}&\left(i\mathbf{k}_{l,j} - \frac{\gamma_3\mathbf{w}_0}{\pi}\right)\frac{e^{\frac{\epsilon\gamma_3 w_{0x}L}{\pi}}-1}{\frac{\epsilon\gamma_3 w_{0x}L}{\pi} + i2\pi(l-n)}\\ &\times\frac{e^{\frac{\epsilon\gamma_3 w_{0y}L}{\pi}}-1}{\frac{\epsilon\gamma_3 w_{0y}L}{\pi} + i2\pi(j-m)}\,\frac{\dot{R}_{l,j}(0)}{\omega_{l,j}^2}.\end{aligned} \tag{2.64}$$

For $\mathbf{w}(\mathbf{X},T) = \mathbf{w}_0 + \tilde{\mathbf{w}}(\mathbf{X},T)$, the order parameter (2.52) becomes

$$\begin{aligned}\mathbf{Z} &\sim \frac{\epsilon}{\rho_0}\mathbf{w}_0 + \frac{\epsilon}{\rho_0\epsilon^2L^2}\int_0^{\epsilon L}\int_0^{\epsilon L}\tilde{\mathbf{w}}(\mathbf{X},T)\,dX\,dY \\ &= \frac{\epsilon}{\rho_0}\mathbf{w}_0 + \frac{\epsilon}{\rho_0}\tilde{\mathbf{w}}_{0,0}(T).\end{aligned} \tag{2.65}$$

2.7.2 Shift in the critical noise

The oscillating density disturbance $\tilde{r}(\mathbf{X},T)$ will produce a nonzero value of the average of $r^2 = \tilde{r}^2$. Averaging Eq. (2.43) over space and time, and assuming that the average of a product is the product of averages, we have $\langle r\mathbf{w}\rangle \approx \langle r\rangle\langle\mathbf{w}\rangle = 0$ and $\langle r^2\mathbf{w}\rangle \approx \langle r^2\rangle\langle\mathbf{w}\rangle$. Then the time-independent and space-averaged part of the term r^2 in $\Gamma(r)$ gives a contribution to $|\mathbf{w}_0| = 2\pi\sqrt{\Gamma(r)/\mu}$, which yields the first term in Eq. (2.65):

$$\begin{aligned}\frac{\epsilon\mathbf{w}_0}{\rho_0} &\sim \frac{2\pi}{\rho_0}\sqrt{\frac{Q_\eta(\eta-\eta_c) - \frac{\epsilon^2\langle\tilde{r}^2(\mathbf{X},T)\rangle}{6\rho_0^2}}{\mu}}\,\mathbf{e}_\Upsilon \\ &\sim \frac{2\pi}{\rho_0}\sqrt{\frac{Q_1 - 1 - \frac{\epsilon^2\langle\tilde{r}^2(\mathbf{X},T)\rangle}{6\rho_0^2}}{\mu}}\,\mathbf{e}_\Upsilon.\end{aligned} \tag{2.66}$$

Here $\tilde{r}$ is given by Eq. (2.60) and, inserted into the average over time and space in Eq. (2.66), contributes to shift the bifurcation point to a smaller noise value η_c^*. Fig. 2.4 shows that simulation data are consistent with $\epsilon\sqrt{\langle\tilde{r}^2(\mathbf{X},T)\rangle/6} = a\rho_0$, with $a \approx 0.3$. Then $\eta_c^* \approx 3.95$. One first correction consists of using a more accurate formula instead of Eq. (2.23) for finite values of M. Using the same procedure as explained in Ref. [19] but keeping more terms in the expansions, the critical condition $Q_1 = 1$ for the noise becomes

$$\frac{\sqrt{\pi M}}{\eta}\left(1 - \frac{1}{8M} - \frac{7}{128M^2} - \frac{5}{128M^3}\right)\sin\frac{\eta}{2} = 1. \tag{2.67}$$

See Appendix 2.C. For $M = 7$, we obtain $\eta_c = 4.09$ ($a = 0.26$) instead of the theoretical value $\eta_c = 4.13$, as in Fig. 2.4(a). Taking into account that the second term in Eq. (2.65) is obtained from linearization about the first term (and is therefore assumed to be small compared with it), the polarization becomes $|\mathbf{Z}| \sim (\epsilon/\rho_0)\,|\mathbf{w}_0+\tilde{\mathbf{w}}_{0,0}| \sim (\epsilon/\rho_0)(|\mathbf{w}_0| + \mathbf{w}_0\cdot\tilde{\mathbf{w}}_{0,0}/|\mathbf{w}_0|)$, i.e.,

$$|\mathbf{Z}| \sim \frac{2\pi}{\rho_0}\sqrt{\frac{Q_1 - 1 - \frac{\epsilon^2\langle\tilde{r}^2(\mathbf{X},T)\rangle}{6\rho_0^2}}{\mu}} + \frac{\epsilon}{\rho_0}\mathbf{e}_\Upsilon\cdot\tilde{\mathbf{w}}_{0,0}(T). \tag{2.68}$$

In Appendix 2.D, we estimate a value for the shift after some uncontrolled approximations that take advantage of the linearized equations (2.54)-(2.56). The result $a \approx 0.01$ is a small improvement that agrees better with the numerically estimated shift of the bifurcation value.

Better agreement should be achieved by numerically solving the full nonlinear equations (2.42)-(2.43) and finding a more precise value of the time and space average $\langle r^2 \rangle$ in the formula for $|\mathbf{w}_0|$. The shift in critical noise was noticed earlier in Ref. [29] but no explanation thereof was given there.

2.7.3 Oscillatory correction to the polarization: resonances in the Vicsek model

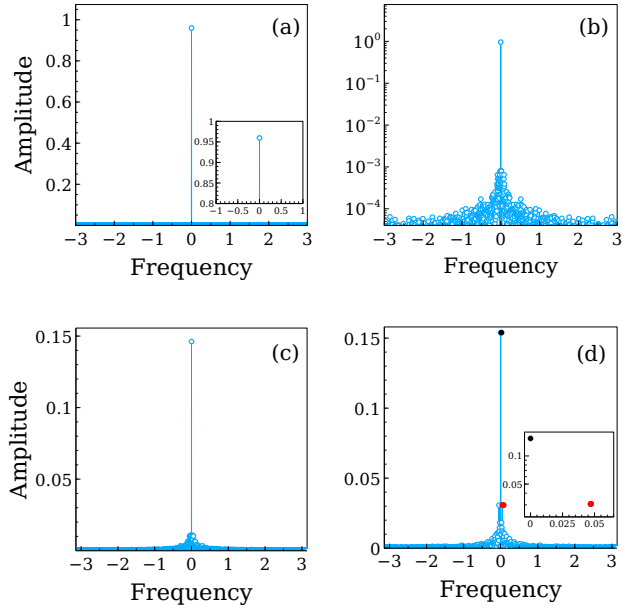

Figure 2.5: Discrete Fourier transform of the polarization $W(t)$ obtained from single simulations of the VM (without ensemble averaging) for: (a), (b) and (c) Vicsek alignment rule of Eq. (2.2); (d) alignment rule of Eq. (2.71), which includes harmonic forcing of frequency $\omega = \Omega_Z \approx 0.05$ with $h = 5$. Parameter values are $M = 7$, $N = 1000$, $\rho_0 = 5$, and (a)-(b) $\eta = 0.8$, (c)-(d) $\eta = 3.7$. Panel (b) is the same as Panel (a) in logarithmic scale. The inset of Panel (a) shows that, for this single simulation, the amplitude of the zero-frequency mode coincides with the ensemble averaged value of W in Fig. 2.4. In Panel (d), filled circles indicate the zero (black dot) and resonant (red dot) frequencies, both peaks are highlighted in the Inset.

Eq. (2.68) contains a bounded oscillatory function of T [cf. Eq. (2.61)]. Thus the polarization is a function of the time t whose lowest angular frequency is Ω_Z:

$$\Omega_Z \sim 2\gamma_3 \sqrt{\frac{Q_1 - 1 - \frac{\epsilon^2 \langle \tilde{r}^2(\mathbf{X},T) \rangle}{6\rho_0^2}}{2\mu}} \sim \frac{W}{2\sqrt{2}}. \tag{2.69}$$

Here W is given by Eq. (2.66). The other frequencies given by Eq. (2.59) are now

$$\epsilon\,\omega_{n,m} \sim \sqrt{\Omega_Z^2 + \frac{2\pi^2}{L^2}(n^2+m^2)}\,. \tag{2.70}$$

How can we confirm the existence of these oscillation frequencies? One possibility is to modify the alignment rule from Eq. (2.2) to

$$\theta_i(t+\tau) = \mathrm{Arg}\left(\sum_{|\mathbf{x}_j-\mathbf{x}_i|<R_0} e^{i\theta_j(t)}\right) + \xi_i(t) + h\cos(\omega t), \tag{2.71}$$

move the forcing frequency until it resonates with one of the frequencies of Eq. (2.70), and simulate the resulting forced VM. However, we need to explore a region of η sufficiently close to η_c. Otherwise, the parameter $\epsilon = \sqrt{(\eta-\eta_c)/\eta_2}$ is so large that there is no separation between the hyperbolic and parabolic scalings. In this later case, say for $\eta = 0.8$, the discrete Fourier transform of the polarization $W(t)$, shown in Fig. 2.5(a), has a single peak at zero frequency and a small, seemingly flat, background [the amplitudes of the transform at nonzero frequency are all smaller than 0.001, cf. Fig. 2.5(b)]. For the transform depicted in Figs. 2.5(a) and 2.5(b), the peak height at zero frequency coincides with the value displayed in Fig. 2.4, which has been obtained as an ensemble average over many realizations of the stochastic process given by the VM. As η increases towards η_c, one simulation of the VM shows that $W(t)$ still has a large peak at zero frequency, but there is a small mound about it [cf. Fig. 2.5(c)]. If we repeat the simulations of the VM with the modified rule Eq. (2.71), Fig. 2.5(d) shows that the mound is higher and that there is a small resonant peak at $\omega = \Omega_Z$. This effect is very small because, for η close to η_c, the polarization and, consequently, the frequency given by Eq. (2.69), are very small, and the alignment noise is large.

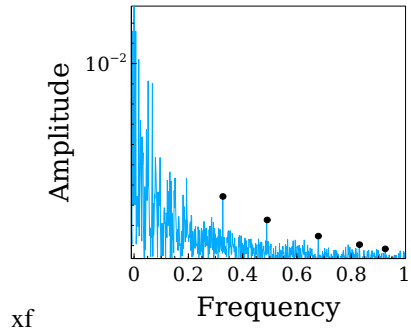

Figure 2.6: Same as Fig. 2.5 but with the forcing frequency of Eq. (2.70) with $n = m = 2$ and $h = 5$.

We can also excite higher frequencies by setting the forcing frequency to be one of those in Eq. (2.70). For instance, if we set ω in Eq. (2.71) equal to the frequency given by

Eq. (2.70) with $n = m = 2$, the nonlinearity of the amplitude equations excites several nearby frequencies. Thus, the peaks appearing in Fig. 2.6 are close to frequencies of the modes (i) $(n, m) = (1, 0),\ (0, 1)$, (ii) $(1, 1)$, (iii) $(2, 0),\ (0, 2)$, (iv) $(2, 2)$, and (v) $(3, 0),\ (0, 3)$, given by Eq. (2.70).

2.8 Conclusions

We have studied flocking in the standard Vicsek model by analyzing the bifurcation of the uniform distribution function to solutions of the associated kinetic equation that have nonzero polarization. Our linear stability analysis of the uniform distribution is limited to small wave numbers. Within this constraint, linear instability first occurs at a real eigenvalue with zero wave number. The picture of flocking that emerges from the bifurcation analysis is intricate.

We have shown that the amplitude equations for the bifurcating modes near the critical value η_c of the noise η are equivalent to coupled equations for a disturbance of the number density and a current density. The equation for the density disturbance is a continuity equation whereas the equation for the current density contains two different scalings: (i) a hyperbolic scaling in which both time and space scale as $|\eta - \eta_c|^{-1/2}$; and (ii) a parabolic scaling in which time scales as $|\eta - \eta_c|^{-1}$, therefore, as space squared.

Space-independent solutions of the amplitude equations obey the usual equation for a supercritical pitchfork bifurcation on the longer parabolic time scale. Stable stationary solutions of this equation produce a polarization proportional to $|\eta - \eta_c|^{1/2}$ as $\eta \to \eta_c$, which is depicted in Fig. 2.4. Compared to direct simulations of the VM, ensemble averages of the polarization are similar to the bifurcation predictions, except for a shift of the critical noise to smaller values and a round off very close to η_c.

Space-dependent patterns near the flocking bifurcation satisfy the full amplitude equations. On the longer parabolic time scale, the solutions of the latter should be close to the space-independent stationary solutions. Then we can study the linearization of the amplitude equations about such solutions. The leading order approximation of the linearized equations on the hyperbolic scaling is equivalent to a Klein-Gordon equation whose solutions for periodic boundary conditions contain infinitely many frequencies of oscillation. Thus, the emerging picture of the flocking bifurcation in the VM is that of an almost uniform polarization with small superimposed Klein-Gordon oscillations that occur on the faster hyperbolic scaling. These oscillations produce a nonzero average of the square of the density disturbance resulting in a shift in the critical noise, which may explain the observed one in direct simulations of the VM.

To confirm this picture of flocking, we have modified the Vicsek alignment rule by adding a harmonic forcing term to the average alignment direction and to the noise. We have then simulated the resulting model looking for resonances between the forcing frequency and one of the Klein-Gordon frequencies. The discrete Fourier transform of the resulting time-dependent polarization shows a large peak at zero frequency (with amplitude equal to that of the stable

stationary solution) and much smaller peaks at other frequencies. For large $|\eta - \eta_c|$, the parabolic and hyperbolic scalings are no longer separated and dissipation effects move the polarization to its stationary value. Close to the critical noise, both scalings are separated and the resonant peak emerges from the background in the discrete Fourier transform obtained from a single direct simulation of the VM. This effect is masked by the large size of the noise near its critical value but it is clearly observable in single simulations of the VM, as shown in Fig. 2.5(d).

One caveat should be indicated here. If the box size is sufficiently large, the flocking transition is discontinuous and band-like patterns have been observed in numerical simulations of the VM [20]. Our for a mode with zero wave number is not applicable to these large box sizes. However, a future extension of our linear stability analysis of the uniform distribution to larger wave numbers may describe the resulting flocking transition, at least near the critical box size at which flocking becomes a discontinuous bifurcation.

Appendix 2.A Eigenvalues for $|\mathbf{K}| = 0$ in the limit as $n \to \infty$

The eigenvalues Q_j for zero wave number are given by Eqs. (2.21)-(2.22). To find them, we need to calculate integrals of the form

$$\mathcal{J}(n,j) = \int_{[-\pi,\pi]^n} e^{ij(\theta_1-\Phi_1)} \prod_{l=1}^{n} \frac{d\theta_l}{2\pi}, \tag{2.72}$$

in the limit as $n \to \infty$. We have

$$e^{ij(\theta_1-\Phi_1)} = \left(\frac{1 + Le^{i(\theta_1-\beta)}}{\sqrt{(1 + Le^{i(\theta_1-\beta)})(1 + Le^{i(\beta-\theta_1)})}} \right)^j = \left(\frac{1 + Le^{i(\theta_1-\beta)}}{1 + Le^{i(\beta-\theta_1)}} \right)^{j/2}, \tag{2.73}$$

where

$$Le^{i\beta} = \frac{\sum_{l=2}^{n} e^{i\theta_l}}{|\sum_{l=2}^{n} e^{i\theta_l}|}. \tag{2.74}$$

In the limit as $n \to \infty$, the central limit theorem implies that we can replace $n-1$ integrals (with $n - 1 \sim n$) in $\mathcal{J}(n,j)$ by [19]

$$\begin{aligned} \mathcal{J}(n,j) &= \tfrac{1}{2\pi} \int_0^\infty \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{Le^{-L^2/n}}{\pi n} \left(\tfrac{1+Le^{i(\theta_1-\beta)}}{1+Le^{i(\beta-\theta_1)}} \right)^{j/2} dL\, d\theta_1 d\beta \\ &= \int_0^\infty \int_{-\pi}^{\pi} \frac{Le^{-L^2/n}}{\pi n} \left(\tfrac{1+Le^{i\theta}}{1+Le^{-i\theta}} \right)^{j/2} dL\, d\theta. \end{aligned} \tag{2.75}$$

We find different approximations for odd and even j. For odd j, we split the L-integral in sub-integrals over $(0,\Lambda)$ and (Λ,∞), with $\Lambda \gg 1$ fixed in the limit as $n \to \infty$. We can approximate $1/n = 0$ in the first sub-integral and expand the fraction in powers of $1/L$ in the

second sub-integral. The result is

$$\mathcal{J}(n,j) = \frac{1}{n}\int_0^\Lambda \int_{-\pi}^{\pi} \frac{L}{\pi}\left(\frac{1+Le^{i\theta}}{1+Le^{-i\theta}}\right)^{j/2} dL\, d\theta$$
$$+ \int_\Lambda^\infty \int_{-\pi}^{\pi} \frac{Le^{ij\theta - L^2/n}}{\pi n}\left\{1 - \frac{ij}{L}\sin\theta + \frac{j}{2L^2}[e^{i2\theta} - 1\right.$$
$$-(j-2)\sin^2\theta] + \frac{j}{2L^3}\left[e^{i\theta} - e^{i3\theta} - \frac{j-2}{2}(e^{i\theta} - e^{-i\theta})\right.$$
$$\left.\left.\times\left(e^{i2\theta} - 1 + \frac{j-4}{12}(e^{i\theta} - e^{-i\theta})^2\right)\right]\right\} dL\, d\theta. \tag{2.76}$$

For $j = 3$, the first integral is $-3\ln\pi/(2\pi n)$ and the second integral is $\sqrt{\pi}n^{-3/2}/8$. For $j = 1$, the second integral yields $n^{-1}\int_\Lambda^\infty e^{-L^2/n}dL = (1/2)\sqrt{\pi/n}$. Thus, for odd j,

$$\mathcal{J}(n,j) = \frac{1}{2}\sqrt{\frac{\pi}{n}}\delta_{j1} - \frac{3\ln\pi}{2\pi n}\delta_{j3} + O(n^{-1}), \tag{2.77}$$

as $n \to \infty$. For even j, the integrals over θ can be transformed into integrals over the unit circle on the complex plane and calculated by the residue theorem.

$$\int_{-\pi}^{\pi}\left(\frac{1+Le^{i\theta}}{1+Le^{-i\theta}}\right)^{j/2}\frac{d\theta}{\pi} = \int_{|z|=1} z^{\frac{j}{2}-1}\left(L + \frac{1-L^2}{z+L}\right)^{j/2}\frac{dz}{i\pi}.$$

For $j = 2$, the residue theorem yields

$$\int_{-\pi}^{\pi}\frac{1+Le^{i\theta}}{1+Le^{-i\theta}}\frac{d\theta}{\pi} = 2(1-L^2)\,\Theta(1-L^2),$$

because the pole $z = -L$ is outside the unit circle if $|L| > 1$. Here $\Theta(x)$ is the Heaviside unit step function. Then, as $n \to \infty$, we get

$$\mathcal{J}(n,2) = \frac{2}{n}\int_0^1 (1-L^2)Le^{-L^2/n}dL = \frac{1}{2n} + O(\frac{1}{n^2}). \tag{2.78}$$

Similarly, for $j = 4$, the residue theorem yields

$$\int_{-\pi}^{\pi}\left(\frac{1+Le^{i\theta}}{1+Le^{-i\theta}}\right)^{2}\frac{d\theta}{\pi} = 2(1-L^2)(1-3L^2)\,\Theta(1-L^2).$$

Then, as $n \to \infty$, we get

$$\begin{aligned}\mathcal{J}(n,4) &= \frac{2}{n}\int_0^1 (1-L^2)(1-3L^2)Le^{-L^2/n}dL \\ &= \frac{1}{12n^2} + O(\frac{1}{n^3}).\end{aligned} \tag{2.79}$$

According to Eq. (2.21),

$$C^{(1)}[e^{ij\theta}]_j = \left(\int_{-\pi}^{\pi} e^{-ij\xi} g(\xi)\, d\xi\right) e^{-M} \sum_{n=1}^{\infty} \frac{M^{n-1}}{(n-1)!} \times \left[(n-1)\frac{2J_1(|\mathbf{K}|R_0)}{|\mathbf{K}|R_0} + 1\right] \mathcal{J}(n,j). \tag{2.80}$$

For large M, we expand $(n-1)\mathcal{J}(n,j)$ and $\mathcal{J}(n,j)$ about M in this expression, thereby getting

$$C^{(1)}[e^{ij\theta}]_j \sim \left(\int_{-\pi}^{\pi} e^{-ij\xi} g(\xi)\, d\xi\right)\left[(M-1)\frac{2J_1(|\mathbf{K}|R_0)}{|\mathbf{K}|R_0} + 1\right] \mathcal{J}(M,j). \tag{2.81}$$

Then the functions $\mathcal{J}(n,j)$ produce Eqs. (2.23) and

$$\begin{aligned} Q_2 &= C^{(2)}[e^{i2\theta}]_2 \sim \frac{1}{2}\int_{-\pi}^{\pi} e^{-i2\xi} g(\xi) d\xi, \\ Q_3 &= C^{(2)}[e^{i3\theta}]_3 \sim -\frac{3\ln\pi}{2\pi}\int_{-\pi}^{\pi} e^{-i3\xi} g(\xi) d\xi, \\ Q_4 &= C^{(2)}[e^{i4\theta}]_4 \sim \frac{1}{12\,M}\int_{-\pi}^{\pi} e^{-i4\xi} g(\xi) d\xi. \end{aligned} \tag{2.82}$$

Other eigenvalues tend to zero as $M \to \infty$. As $|\int_{-\pi}^{\pi} e^{-ij\xi} g(\xi) d\xi| \le \int_{-\pi}^{\pi} g(\xi) d\xi = 1$, the multipliers Q_j with $j > 1$ have moduli smaller than 1 in the limit as $M \gg 1$.

Appendix 2.B Coefficients in the amplitude equations

We calculate the coefficients in the amplitude equations by identifying them with others computed in [19]. Using Ihle's notation, we obtain $C^{(1)}[e^{i2\theta}]_2 = Q_2 = p$, $C^{(2)}[e^{i\theta}, e^{i\theta}]_2 = 2\pi q$, $C^{(2)}[e^{-i\theta}, e^{i2\theta}]_1 = \pi S$, $C^{(3)}[e^{i\theta}, e^{i\theta}, e^{-i\theta}]_1 = 4\pi^2\Gamma$, [19]. In the limit as as $M \gg 1$, these identifications allow us to obtain the coefficients in (2.42)-(2.43):

$$Q_\eta \sim -\frac{\sqrt{\pi M}}{2\eta_c}\left(\frac{2}{\sqrt{\pi M}} - \cos\frac{\eta_c}{2}\right), \tag{2.83}$$

$$\delta \sim \frac{2\gamma_0 - 1}{8}, \qquad \gamma_0 = \frac{1}{1 - \frac{1}{\sqrt{\pi M}}\cos\frac{\eta_c}{2}}, \tag{2.84}$$

$$\mu \sim \frac{\pi^4 R_0^4/M}{1 - \frac{1}{\sqrt{\pi M}}\cos\frac{\eta_c}{2}}, \quad \gamma_2 \sim \frac{\gamma_0 \pi^2 R_0^2}{4M}. \tag{2.85}$$

To calculate the other coefficients appearing in the amplitude equations, we note that, as $M \gg 1$,

$$C^{(2)}[1, e^{i\theta}]_1 = \frac{\pi M}{\rho_0}\frac{\partial C^{(1)}[e^{i\theta}]_1}{\partial M} \sim \frac{\pi C^{(1)}[e^{i\theta}]_1}{2\rho_0} = \frac{\pi}{2\rho_0}, \tag{2.86}$$

$$C^{(3)}[1, 1, e^{i\theta}]_1 = \frac{2\pi^4 R_0^4}{3}\frac{\partial^2 C^{(1)}[e^{i\theta}]_1}{\partial M^2} \sim -\frac{\pi^2 C^{(1)}[e^{i\theta}]_1}{6\rho_0^2} = -\frac{\pi^2}{6\rho_0^2}. \tag{2.87}$$

We have calculated these coefficients at the critical noise η_c, where $C^{(1)}[e^{i\theta}]_1 = 1$. Eqs. (2.86) and (2.87) yield the remaining coefficients:

$$\Gamma(r) \sim \eta_2 Q_\eta - \frac{r^2}{6\rho_0^2}, \quad \gamma_3 \sim \frac{\pi^2 R_0^2}{2M} \tag{2.88}$$

$$\gamma_1 \sim \pi^2 R_0^2\left(1 + \frac{1}{8M} - \gamma_0\right). \tag{2.89}$$

Appendix 2.C Fourier coefficients $R_{n,m}(0)$

Here we give examples of initial conditions used to calculate the solutions of Eqs. (2.54)-(2.55). A simple initial condition is to set $\tilde{\mathbf{w}}(\mathbf{X}, 0) = \mathbf{0}$. Then Eqs. (2.63) and (2.64) yield

$$\dot{R}_{n,m}(0) = 0, \quad \mathbf{C}_{n,m} = \mathbf{0}, \tag{2.90}$$

whereas Eq. (2.61) gives

$$\begin{aligned}\epsilon\tilde{\mathbf{w}}_{0,0}(T) = {} & (e^{\frac{\epsilon\gamma_3 w_{0x}L}{\pi}} - 1)(e^{\frac{\epsilon\gamma_3 w_{0y}L}{\pi}} - 1) \\ & \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\frac{\left(i\frac{2\pi}{L}(n,m) - \frac{\epsilon\gamma_3\mathbf{w}_0}{\pi}\right)R_{n,m}(0)\sin(\omega_{n,m}T)}{\omega_{n,m}\left(\frac{\epsilon\gamma_3 w_{0x}L}{\pi} + i2\pi n\right)\left(\frac{\epsilon\gamma_3 w_{0y}L}{\pi} + i2\pi m\right)}.\end{aligned} \tag{2.91}$$

We now have to calculate $R_{n,m}(0)$. One possibility is to have a function $\tilde{r}(\mathbf{X}, 0)$ with finitely many harmonics. For example, harmonics $(\pm 1, 0)$ and $(0, \pm 1)$. We obtain

$$\begin{aligned}R_{n,m}(0) = {} & 2(e^{\frac{\epsilon\gamma_3 w_{0x}L}{\pi}} - 1)(e^{\frac{\epsilon\gamma_3 w_{0y}L}{\pi}} - 1)\left\{\frac{(\epsilon\gamma_3 w_{0x}L - i2\pi^2 n)\mathrm{Re}\tilde{r}_{1,0}(0) + 2\pi^2\mathrm{Im}\tilde{r}_{1,0}(0)}{\left[\left(\frac{\epsilon\gamma_3 w_{0x}L}{\pi} - i2\pi n\right)^2 + 4\pi^2\right](\epsilon w_{0y}L - i2\pi^2 m)}\right. \\ & \left. + \frac{(\epsilon\gamma_3 w_{0y}L - i2\pi^2 m)\mathrm{Re}\tilde{r}_{0,1}(0) + 2\pi^2\mathrm{Im}\tilde{r}_{0,1}(0)}{\left[\left(\frac{\epsilon\gamma_3 w_{0y}L}{\pi} - i2\pi m\right)^2 + 4\pi^2\right](\epsilon\gamma_3 w_{0x}L - i2\pi^2 n)}\right\}.\end{aligned} \tag{2.92}$$

Another simple possibility is the initial condition $\tilde{r}_{n,m}(0) = (-1)^{n+m}(1-\delta_{n0}\delta_{m0})$. Then Eq. (2.62) yields

$$\begin{aligned} R_{n,m}(0) &= (-1)^{n+m} e^{\frac{\epsilon\gamma_3 w_{0x} L}{2\pi}} e^{\frac{\epsilon\gamma_3 w_{0y} L}{2\pi}} \\ &+ \frac{e^{\frac{\epsilon\gamma_3 w_{0x} L}{\pi}} - 1}{2n\pi + i\frac{\epsilon\gamma_3 w_{0x} L}{\pi}} \frac{e^{\frac{\epsilon\gamma_3 w_{0y} L}{\pi}} - 1}{2m\pi + i\frac{\epsilon\gamma_3 w_{0y} L}{\pi}}. \end{aligned} \tag{2.93}$$

Appendix 2.D Calculation of the shift in the critical noise

According to Eq. (42) of Ref. [19], we have to approximate better the sum

$$S(M) = e^{-M} \sum_{n=0}^{\infty} \frac{M^n}{n!} h(n), \tag{2.94}$$

where $h(n)$ has a maximum at $n = M$. Expanding this function about its maximum and keeping four terms in the expansion, we obtain

$$S(M) \sim h(M) + \frac{M}{2} h''(M) + \frac{M}{6} h'''(M) + \frac{(1+3M)M}{24} h^{(4)}(M), \tag{2.95}$$

in which we have summed the corresponding series. To write Eq. (2.23), we took into account only the first term of Eq. (2.95) with $h(n) = \sqrt{n}$. For this function, we get

$$S(M) \sim \sqrt{M}\left(1 - \frac{1}{8M} + \frac{1}{16M^2} - \frac{5(1+3M)}{128M^3}\right), \tag{2.96}$$

which produces Eq. (2.67). For $M = 7$, keeping more term in the expansion of $h(n)$ does not change appreciably the numerical value of η_c. Other corrections could come from calculating the term of order $n^{-3/2}$ in Eq. (2.77) because $h(n)$ is proportional to $n\,\mathcal{J}(n,1)$. We now calculate $\langle r^2(\mathbf{X},T)\rangle$. We can use the Parseval equality and Eq. (2.60) to get

$$\langle \tilde{r}^2(\mathbf{X},T)\rangle = \lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \sum_{n,m} |\tilde{r}_{n,m}(T)|^2 dT = \frac{1}{2} \sum_{n,m} \left[|R_{n,m}(0)|^2 + \frac{|\dot{R}_{n,m}(0)|^2}{\omega_{n,m}^2} \right] g(\mathbf{w}_0), \tag{2.97}$$

$$\begin{aligned} g(\mathbf{w}_0) &= \delta_{\mathbf{w}_0,\mathbf{0}} + \delta_{w_{0x},0}(1-\delta_{w_{0y},0}) \frac{1 + e^{\frac{\gamma_3}{\pi} w_{0y}\epsilon L}}{2\frac{\gamma_3}{\pi} w_{0y}\epsilon L} + \delta_{w_{0y},0}(1-\delta_{w_{0x},0}) \frac{1 + e^{\frac{\gamma_3}{\pi} w_{0x}\epsilon L}}{2\frac{\gamma_3}{\pi} w_{0x}\epsilon L} \\ &+ \frac{1}{4}(1-\delta_{w_{0x},0}) \times (1-\delta_{w_{0y},0}) \frac{(1 + e^{\frac{\gamma_3}{\pi} w_{0y}\epsilon L})(1 + e^{\frac{\gamma_3}{\pi} w_{0x}\epsilon L})}{\left(\frac{\gamma_3}{\pi}\epsilon L\right)^2 w_{0x} w_{0y}}. \end{aligned} \tag{2.98}$$

To get these expressions, we have used that: (i) the averages of $\cos^2(\omega_{n,m}T)$ and of $\sin^2(\omega_{n,m}T)$ are both $1/2$, that the average of $\cos(\omega_{n,m}T)\cos(\omega_{l,j}T)$ is zero unless $n = l$,

$m = j$, etc.; (ii) the integrals

$$\frac{1}{\epsilon^2 L^2}\int_0^{\epsilon L}\int_0^{\epsilon L} e^{(\frac{\gamma_2}{\pi}\mathbf{w}_0 - i\mathbf{k}_{n-l,m-j})\cdot\mathbf{X}} d\mathbf{X} = \frac{(e^{\frac{\gamma_2}{\pi}w_{0x}\epsilon L}-1)\,(e^{\frac{\gamma_2}{\pi}w_{0y}\epsilon L}-1)}{[\frac{\gamma_3}{\pi}w_{0x}\epsilon L - i2\pi(n-l)][\frac{\gamma_3}{\pi}w_{0y}\epsilon L - i2\pi(m-j)]}, \tag{2.99}$$

(iii) the sums

$$\sum_{n=-\infty}^{\infty}\frac{1}{\frac{\gamma_3^2}{\pi^2}w_{0x}^2\epsilon^2 L^2 + 4\pi^2(n-l)^2} = \frac{\pi\coth\frac{\gamma_3 w_{0x}\epsilon L}{2\pi}}{2\gamma_3 w_{0x}\epsilon L}. \tag{2.100}$$

In Eq. (2.97), we can use again the Parseval equality and then the Schwarz inequality to obtain

$$\begin{aligned}\sum_{n,m}|R_{n,m}(0)|^2 &= \frac{1}{\epsilon^2 L^2}\int_0^{\epsilon L}\int_0^{\epsilon L}\tilde{r}^2(\mathbf{X},0)e^{-2\frac{\gamma_3}{\pi}\mathbf{w}_0\cdot\mathbf{X}}d\mathbf{X}\\ &\le \sqrt{\langle\tilde{r}(\mathbf{X},0)^4\rangle\frac{1}{\epsilon^2 L^2}\int_0^{\epsilon L}\int_0^{\epsilon L}e^{-4\frac{\gamma_3}{\pi}\mathbf{w}_0\cdot\mathbf{X}}d\mathbf{X}}.\end{aligned} \tag{2.101}$$

We can calculate $r^4 = \epsilon^4\langle\tilde{r}^4\rangle$ by means of the grand canonical ensemble as [25]:

$$\begin{aligned}\langle(N-\langle N\rangle)^4\rangle &= \left(z\frac{\partial}{\partial z}\right)^4\ln\mathcal{Q}(v,\beta) + 3\langle(N-\langle N\rangle)^2\rangle^2\\ &= \frac{1}{\beta^4}\frac{\partial^4}{\partial\mu^4}\ln\mathcal{Q}(v,\beta) + 3\langle(N-\langle N\rangle)^2\rangle^2,\end{aligned} \tag{2.102}$$

where $z = e^{\beta\mu}$ is the fugacity, and $v = 1/\rho_0 = V/N$ ($V = L^2$), β and μ are the specific area, the reciprocal of temperature in enegy units and the chemical potential, respectively. In the grand canonical ensemble, $\ln\mathcal{Q}(v,\beta) = \beta V P$, where $P(v)$ is the pressure,

$$\mathcal{Q}(v,\beta) = \sum_{N=0}^{\infty} z^N Q_N(v,\beta), \tag{2.103}$$

and $Q_N(v,\beta)$ is the partition function of the canonical ensemble. The average number of particles in a volume V is the density of particles and $\langle(N-\langle N\rangle)^2\rangle$ is the fluctuation of the density. In terms of the grand partition function, they are

$$\langle N\rangle = \frac{\sum_{N=0}^{\infty}N z^N Q_N(v,\beta)}{\sum_{N=0}^{\infty}z^N Q_N(v,\beta)} = \left(z\frac{\partial}{\partial z}\right)\ln\mathcal{Q}(v,\beta), \tag{2.104}$$

$$\langle(N-\langle N\rangle)^2\rangle = \left(z\frac{\partial}{\partial z}\right)^2\ln\mathcal{Q}(v,\beta). \tag{2.105}$$

Following Ref. [25], we now define pressure and chemical potential in terms of the free energy per particle $a(v)$ (we ignore the temperature dependence):

$$A(N,V,T) = N\,a(v), \quad P = -\frac{\partial a(v)}{\partial v}, \qquad \mu = a(v) + Pv. \tag{2.106}$$

Using Eq. (2.106), Eq. (2.104) becomes

$$\begin{aligned}\langle N\rangle &= \left(z\frac{\partial}{\partial z}\right)\ln \mathcal{Q}(v,\beta) = \frac{1}{\beta}\,\frac{\partial}{\partial \mu}\ln \mathcal{Q}(v,\beta)\\ &= V\,\frac{\partial P}{\partial \mu} = V\,\frac{\frac{\partial P}{\partial v}}{\frac{\partial \mu}{\partial v}} = V\frac{\frac{\partial P}{\partial v}}{\frac{\partial a(v)}{\partial v} + P + v\frac{\partial P}{\partial v}} = \frac{V}{v},\end{aligned} \tag{2.107}$$

which is indeed the average number of particles. Then $1/v = \langle N\rangle/V = \rho_0$. Similarly, Eq. (2.105) yields

$$\langle (N - \langle N\rangle)^2\rangle = \frac{V}{\beta}\,\frac{\partial}{\partial \mu}\frac{1}{v} = -\frac{V}{v^2\beta}\,\frac{1}{\frac{\partial \mu}{\partial v}} = \frac{\langle N\rangle}{-\beta v^2\frac{\partial P}{\partial v}}. \tag{2.108}$$

This is Eq. (7.43) of Ref. [25]. The particles can be thought of as belonging to an ideal gas at the initial time, therefore $P = \rho_0/\beta$ and $-v^2\partial P/\partial v = \partial P/\partial \rho_0 = 1/\beta$. Eq. (2.108) becomes $\langle (N - \langle N\rangle)^2\rangle = \langle N\rangle$, and therefore

$$\langle r^2\rangle = \frac{\rho_0}{V} = \frac{\rho_0^2}{N} \Longrightarrow \sqrt{\langle r^2\rangle} = \frac{\rho_0}{\sqrt{N}}, \tag{2.109}$$

where now N is the total number of particles in the box.

Using Eq. (2.108) and the ideal gas assumption in Eq. (2.108), Eq. (2.102) becomes

$$\begin{aligned}&\langle (N - \langle N\rangle)^4\rangle = \langle N\rangle + 3\langle N\rangle^2 \Longrightarrow\\ &\langle r^4\rangle = \frac{3\rho_0^4}{N^2}\left(1 + \frac{1}{3N}\right).\end{aligned} \tag{2.110}$$

Then, near the bifurcation point where $\mathbf{w}_0\epsilon L \ll 1$, Eq. (2.101) produces

$$\epsilon^2 \sum_{n,m} |R_{n,m}(0)|^2 \le \frac{\rho_0^2\sqrt{3}}{N}\sqrt{1 + \frac{1}{3N}}. \tag{2.111}$$

Let us assume that Eq. (2.111) is an equality and that $\dot{R}_{n,m}(0) = 0$, which is the case if the initial current density is zero. Then Eq. (2.97) gives

$$\langle r^2(\mathbf{X},T)\rangle \approx \frac{\sqrt{3}}{2N}\rho_0^2\sqrt{1 + \frac{1}{3N}}, \tag{2.112}$$

and the shift in Eq. (2.66) is

$$a^2 = \frac{\langle r^2(\mathbf{X}, T)\rangle}{6\rho_0^2} \approx \frac{\sqrt{3}}{12N}\sqrt{1 + \frac{1}{3N}}. \tag{2.113}$$

This produces $a = 0.01$ which is about a factor 30 smaller than the shift in the critical noise measured from direct numerical simulations of the VM.

References

[1] O. A. Igoshin, A. Mogilner, R.D. Welch, D. Kaiser, and G. Oster, Pattern formation and traveling waves in myxobacteria: Theory and modeling, Proc. Natl. Acad. Sci. USA **98**, 14913 (2001),

[2] O.A. Igoshin, J.C. Neu, and G. Oster, Developmental waves in myxobacteria: A distinctive pattern formation mechanism, Phys. Rev. E **70**, 041911 (2004).

[3] R. Balagam and O.A. Igoshin, Mechanism for Collective Cell Alignment in Myxococcus xanthus Bacteria, PLoS Comp. Biology **11**(8), e1004474 (2015).

[4] A. Creppy, F. Plouraboué, O. Praud, X. Druart, S. Cazin, H. Yu, and P. Degond, Symmetry-breaking phase transitions in highly concentrated semen, J. R. Soc. Interface **13**, 20160575 (2016).

[5] A. Cavagna, D. Conti, C. Creato, L. Del Castello, I. Giardina, T. S. Grigera, S. Melillo, L. Parisi, and M. Viale, Dynamic scaling in natural swarms, Nat. Phys. **13**, 914 (2017).

[6] A. Cavagna, I. Giardina, and T.S. Grigera, The physics of flocking: Correlation as a compass from experiments to theory, Phys. Rep. **728**, 1 (2018).

[7] J. Toner, and Y. Tu, Long-Range Order in a Two-Dimensional Dynamical XY Model: How Birds Fly Together, Phys. Rev. Lett. **75**, 4326 (1995).

[8] J. Toner, Y. Tu, and S. Ramaswamy, Hydrodynamics and phases of flocks, Ann. Phys. **318**, 170 (2005).

[9] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic, Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study, Proc. Natl. Acad. Sci. U.S.A. **105**, 1232 (2008).

[10] A. Attanasi, A. Cavagna, L. Del Castello, I. Giardina, S. Melillo, L. Parisi, O. Pohl, B. Rossaro, E. Shen, E. Silvestri, and M. Viale, Finite-Size Scaling as a Way to Probe Near-Criticality in Natural Swarms, Phys. Rev. Lett. **113**, 238102 (2014).

[11] I. D. Couzin, J. Krause, N. R. Franks, and S. A. Levin, Effective leadership and decision-making in animal groups on the move, Nature (London) **433**, 513 (2005).

[12] M. Rubenstein, A. Cornejo, and R. Nagpal, Programmable self-assembly in a thousand-robot swarm, Science **345**, 795 (2014).

[13] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, Novel type of phase transition in a system of self-driven particles, Phys. Rev. Lett. **75**, 1226 (1995).

[14] S. Ramaswamy, The Mechanics and Statistics of Active Matter, Annu. Rev. Condens. Matter Phys. **1**, 323 (2010).

[15] M.C. Marchetti, J.F. Joanny, S. Ramaswamy, T.B. Liverpool, J. Prost, M. Rao and R.A. Simha, Hydrodynamics of soft active matter, Rev. Mod. Phys. **85**, 1144 (2013).

[16] T. Ihle, Kinetic theory of flocking: Derivation of hydrodynamic equations, Phys. Rev. E **83**, 030901(R) (2011).

[17] T. Ihle, Large density expansion of a hydrodynamic theory for self-propelled particles, Eur. Phys. J. Spec. Top. **224**, 1303 (2015).

[18] T. Vicsek and A. Zafeiris, Collective motion, Phys. Rep. **517**, 71 (2012).

[19] T. Ihle, Chapman-Enskog expansion for the Vicsek model of self-propelled particles, J. Stat. Mech.: Theor. Expts. **(2016)**, 083205.

[20] G. Grégoire and H. Chaté, Onset of Collective and Cohesive Motion, Phys. Rev. Lett. **92**, 025702 (2004).

[21] A.P. Solon, H. Chaté, and J. Tailleur, From Phase to Microphase Separation in Flocking Models: The Essential Role of Nonequilibrium Fluctuations, Phys. Rev. Lett. **114**, 068101 (2015).

[22] L. L. Bonilla, Chapman-Enskog method and synchronization of globally coupled oscillators, Phys. Rev. E **62**, 4862 (2000).

[23] J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, and R. Spigler, The Kuramoto model: a simple paradigm for synchronization phenomena, Rev. mod. Phys. **77**, 137 (2005).

[24] L. L. Bonilla and S. W. Teitsworth, *Nonlinear wave methods for charge transport.* (Wiley-VCH, 2010).

[25] K. Huang, *Statistical Mechanics*, 2nd ed (Wiley, New York 1987).

[26] R. Haberman, *Applied partial differential equations with Fourier series and boundary value problems, 5th ed.* (Pearson, Boston, MA 2013).

[27] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics (Non-relativistic Theory)*, 3rd ed. (Pergamon Press, New York, 1977)

[28] J. C. Neu, *Singular Perturbations in the Physical Sciences*, Graduate Studies in Mathematics, Vol. 147 (American Math- ematical Society, Providence, RI, 2015).

[29] Y.-L. Chou, R. Wolfe, and T. Ihle, Kinetic theory for systems of self-propelled particles with metric-free interactions, Phys. Rev. E **86**, 021120 (2012).

Chapter 3

Contrarian compulsions in the Vicsek model

Contents

3.1 Introduction

The ability to convert free energy into systematic motion characterizes active matter [1, 2, 3, 4, 5]. Thus collective behavior of active particles is a central aspect of subject. As we have presented in Chapter 2 collective motion of active particles can be predicted by simple models as the Vicsek model (VM) [6] or its variants [1, 2, 3, 7, 8, 9, 10, 11]. There are examples of more complex individual behavior in active particles. For example, instead of moving linearly, *E. coli* bacteria swim clockwise in circular trajectories near walls [12, 13]. Simpler artificial micro-swimmers can also be induced to move circularly about solid surfaces [14] and autonomous motion of active colloids is reviewed in Ref. [15]. Recently, Liebchen and Levis have proposed a model of chiral active matter to explain this motion [16]. In two dimensions, each particle i moves with constant speed v and is parallel to its polarization vector. The angular velocity of the latter is the sum of three terms: a constant rotation, a white noise, and a Kuramoto coupling between the angles of particles inside the circle of influence of the particle i [16]. This model exhibits a flocking transition to one or several clusters comprising particles rotating in synchrony (microflocking) and a variety of patterns.

Traditionally, collective motion is considered to be a profitable state for the entities of the system. It gives them an easier way to look for food and their internal communication allows them a bigger protection against a predator. However, recently studies show that there exist more intricate individual interests between the entities inside the organization remaining them together but not with a collective purpose. For instance in some insect species, the cannibalism is a nutritional behavior that makes them to move in a coordinated migration. In Ref. [17] study the mechanisms involved in the behavior of desert locusts. When these insects are in a swarm, they try to bite each other but with the risk to be bitten and although this behavior should makes them to be separated, it produces a collective migration. In particular, Ref. [18] proposes a self propelled particle model that characterizes the behavior of a system with switches in their motion observed in locusts and starlings.

Following the approach presented above, in this Chapter we explore a different mechanism to attain synchronous rotation in small clusters. We consider a two-dimensional (2D) modified VM with forward update. Active particles may be *conformist* and align their velocities to the average velocity of their neighbors with probability $1-p$, or be *contrarian* and move opposite to the average angle with probability $p \in [0,1]$. This choice makes the VM similar to a Kuramoto model of phase synchronization [19, 20] with conformist and contrarian oscillators [21]. In opinion formation models [22, 23], our VM may mimic the conflict between reaching consensus with others and keeping a contrarian opinion. Experiments on imitative behavior in emergency escape of human crowds have shown that sometimes avoiding the majority is the best survival strategy [24]. Our VM is a step towards exploring this kind of behavior.

3.2 The modified Vicsek model

In this Section, we are going to present a modified Vicsek model in which active particles may align their velocity with the local average direction of motion or with the (almost) opposite direction. Flocking behavior is going to depend on the probability p of (almost) contrarian compulsions compared with that of conformist alignment according to the Vicsek rule.

In two spatial dimensions, the angle of a particle j is updated according to a modified angular noise rule and the particles undergo discrete dynamics so that their positions are forwardly updated

$$\theta_j(t+\tau) = \mathrm{Arg}\left(\sum_{|\mathbf{x}_l-\mathbf{x}_j|<R_0} e^{i\theta_l(t)}\right) + \xi_j(t), \tag{3.1}$$

$$\mathbf{x}_j(t+\tau) = \mathbf{x}_j(t) + \tau \mathbf{v}_j(t+\tau). \tag{3.2}$$

Here, the direction of motion is defined with $\mathbf{v}_j = v_0(\cos\theta_j, \sin\theta_j)$, $j = 1,\ldots,N$. In Eq. (3.1), neighbors are all particles inside a circle of radius R_0 centered at particle $\mathbf{x}_j$ which is included in the sum. At each time, $\xi_j(t)$ are independent identically distributed (i.i.d.)

random alignment noises, selected with probability density

$$g(\xi) = p\,\delta(\xi - \xi_0) + \frac{1-p}{\eta}\chi_{[-\eta/2,\eta/2]}(\xi), \quad 0 \le p \le 1. \tag{3.3}$$

Here $\xi_0 \in (-\pi/2, \pi]$, and $\chi_{[-\eta/2,\eta/2]}(\xi)$ is 1 for $-\eta/2 < \xi < \eta/2$ and 0 otherwise, see Fig. 3.1. The parameter η measures the width of the alignment noise being, a tolerance to failure in the alignment rule. In particular, we distinct two cases,

- If $\xi_0 = \pi$, particles experience contrarian compulsions with probability p, and conformist compulsions toward the average angle of their neighbors with probability $(1-p)$.

- If $-\pi/2 < \xi_0 < \pi$, particles experience a large deflection instead of a perfectly contrarian alignment. We speak of *almost contrarian compulsions* because the ordered phases are similar rotating wave phases for all angles of deflection in $(-\pi/2, \pi)$. The numerical method used to implement Eq. (3.3) is described in Appendix 3.A.

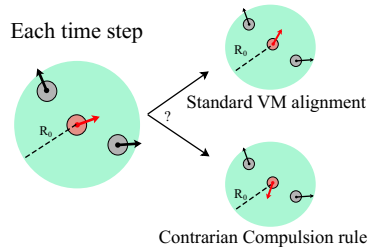

Figure 3.1: Sketch of the Standard VM vs. Modified VM. The particle in red can align its direction towards the average direction of their local neighbors plus a noise chosen between $-\eta/2 < \xi < \eta/2$ or experience a contrarian compulsion and $\xi = \xi_0$.

Collective consensus is quantified by the same complex order parameter defined in Eq. (2.4). As we have been explained, for the standard VM ($p = 0$), increasing the average number of particles M, favors flocking as more and more particles try to move together. Instead, increasing the alignment noise η tends to destroy flocks, as it dilutes the efficacy of the alignment rule. Then there is a critical value of M and η above which the polarization is $W > 0$ and below which $W = 0$ (in the limit as $N \to \infty$). For p closer to 1, increasing M also favors flocking because more and more particles interact and get to move together. However, we shall see that the alignment noise η may favor flocking because it dilutes the strength of the conformist rule and it allows the contrarian or almost contrarian rule to form clusters of particles moving synchronously. We could replace a similar density $\chi_{[-\tilde{\eta}/2,\tilde{\eta}/2]}(\xi - \xi_0)/\tilde{\eta}$ instead of $\delta(\xi - \xi_0)$ in Eq. (3.3), thus adding a more realistic tolerance $\tilde{\eta}$ to the choice of deflection angle ξ_0. However, the results would be qualitatively similar to those reported here.

3.3 Kinetic equation and Linear stability

In this Section, we present the macroscopic evolution equation derived from the microscopic rules of the modified VM following the procedure derived in Section 2.3. The Enskog-type collision operator equation for the modified VM is going to be build following the same assumptions presented in Chapter 2 but with the new alignment rule. Thus, $g(\xi)$ is going to be considered as the modified alignment rule defined in Eq. (3.3). To remember the equation, the Enskog-type collision operator is

$$C_E[f] = \int_{-\pi}^{\pi} d\xi\, g(\xi) e^{-M_R(\mathbf{x},t)} \sum_{n=1}^{\infty} \int_{[-\pi,\pi]^n} \frac{\hat{\delta}(\theta - \xi - \Phi_1(\tilde{\theta}_1,\ldots,\tilde{\theta}_n))}{(n-1)!} \times f(\mathbf{x},\tilde{\theta}_1,t) \prod_{i=2}^{n} \left[\int_{|\mathbf{x}_i - \mathbf{x}|<R_0} f(\mathbf{x}_i,\tilde{\theta}_i,t) d\tilde{\theta}_i d\mathbf{x}_i \right] d\tilde{\theta}_1. \tag{3.4}$$

which we shall use henceforth. For active particles in a disordered state, the density $\rho(\mathbf{x},t)$ equals the constant average density, $\rho_0 = N/L^2$, and the uniform distribution function, $f_0 = \rho_0/(2\pi)$, is a fixed point of the collision operators:

$$C[f_0] = f_0, \quad C_E[f_0] = f_0. \tag{3.5}$$

3.3.1 Eigenvalue problem

To study the linear stability of the uniform distribution, let us consider the same study carried out in Section 2.3 except for the probability density $g(\xi)$ defined in Eq. (3.3). We have not succeeded in finding the eigenvalues of the matrix $\mathcal{M}_{jl}(Q)$ for arbitrary $\mathbf{K}$. However, for small $|\mathbf{K}|$, the off-diagonal elements of the matrix are small compared to its diagonal elements. Then we can use second-order perturbation theory to derive the following formula (cf. Appendix 3.A):

$$Q_1(|\mathbf{K}|,\eta) \approx \frac{C^{(1)}[e^{i\theta}]_1}{J_0(|\mathbf{K}|)} \left[1 - \frac{[J_1(|\mathbf{K}|)/J_0(|\mathbf{K}|)]^2}{1 - C^{(1)}[e^{i2\theta}]_2 / C^{(1)}[e^{i\theta}]_1} \right], \tag{3.6}$$

in which $C^{(1)}[e^{ij\theta}]_j$ are given by

$$C^{(1)}[e^{ij\theta}]_j = \left(\int_{-\pi}^{\pi} e^{-ij\xi} g(\xi)\, d\xi \right) \sum_{n=1}^{\infty} \frac{M^{n-1} e^{-M}}{(n-1)!} \left[(n-1) \frac{2J_1(|\mathbf{K}|R_0)}{|\mathbf{K}|R_0} + 1 \right] \int_{[-\pi,\pi]^n} e^{ij(\theta_1 - \Phi_1)} \prod_{l=1}^{n} \frac{d\theta_l}{2\pi}. \tag{3.7}$$

The uniform distribution becomes unstable when one eigenvalue Q moves outside the unit circle in the complex plane. It turns out that Q_1 has the largest modulus for small $|\mathbf{K}|$, as we argue below. Depending on the parameters M and ρ_0, the equation $|Q_1(|\mathbf{K}|,\eta)| = 1$ may have

different solution branches $\eta(|\mathbf{K}|)$ for $|\mathbf{K}| > 1$. However, for these branches, the off-diagonal elements of the matrix $\mathcal{M}_{jl}(Q)$ are no longer small compared to its diagonal elements, the regular perturbation theory is no longer valid, and we ignore them. In all cases, the uniform

Figure 3.2: Critical wave vector $K = |\mathbf{K}|$ versus η for $M = 7$, $p = 0.1$, $\xi_0 = 3\pi/4$, $\rho_0 = 10$.

distribution becomes first unstable for the noise corresponding to zero wave number. Fig. 3.2 shows the solution curve of $|Q_1(|\mathbf{K}|, \eta)| = 1$ for $p = 0.1$, $\xi_0 = 3\pi/4$ and $\rho_0 = 10$. A similar curve is found for the contrarian case, $\xi_0 = \pi$. In all cases, the uniform distribution becomes first unstable for the noise corresponding to zero wave number. Within our approximations, this justifies that the largest value of the multiplier Q_1 is attained at zero wave number. For $|\mathbf{K}| = 0$, in the limit as $M \gg 1$, we have the eigenvalues

$$Q_1 \sim \frac{\sqrt{\pi M}}{2} \int_{-\pi}^{\pi} e^{-i\xi} g(\xi) d\xi = \frac{\sqrt{\pi M}}{2} \left(p\, e^{-i\xi_0} + \frac{2(1-p)}{\eta} \sin \frac{\eta}{2} \right), \tag{3.8}$$

$$Q_2 \sim \frac{1}{2} \int_{-\pi}^{\pi} e^{-i2\xi} g(\xi) d\xi, \tag{3.9}$$

$$Q_3 \sim -\frac{3 \ln \pi}{2\pi} \int_{-\pi}^{\pi} e^{-i3\xi} g(\xi) d\xi, \tag{3.10}$$

$$Q_4 \sim \frac{1}{12M} \int_{-\pi}^{\pi} e^{-i4\xi} g(\xi) d\xi, \tag{3.11}$$

and so on, with eigenfunctions $\varphi_j(\theta) = e^{ij\theta}$ (cf. Ref. [9] for Q_1 and Q_2 and the standard VM). As $|\int_{-\pi}^{\pi} e^{-ij\xi} g(\xi) d\xi| \leq \int_{-\pi}^{\pi} g(\xi) d\xi = 1$, $|Q_j| \leq 1$ for $j > 1$ in the limit as $M \gg 1$. For $|\mathbf{K}| = 0$, the eigenvalue with largest modulus is therefore Q_1, which is the only one that can exit the unit circle in the complex plane. The other modes have $|Q_j| < 1$ and decay as $t \to \infty$.

3.3.2 Phase diagrams

This Section discusses the phase diagrams depending on the parameters of the system. Phase diagrams give a variety of details among the phases as well as quantitative information on phase distribution in a specific system. Indeed, it is important to understand the behavior of phase transformations as well as phase relations and their stability.

3.3.2.1 Noises p vs η

Fig. 3.3 depicts the stability regions of the disordered state in the parameter space (p, η) at zero wave number.

In particular, the eigenvalue Q_1 given by Eq. (3.8) is real if $g(\xi)$ is even. For $p > 0$ and

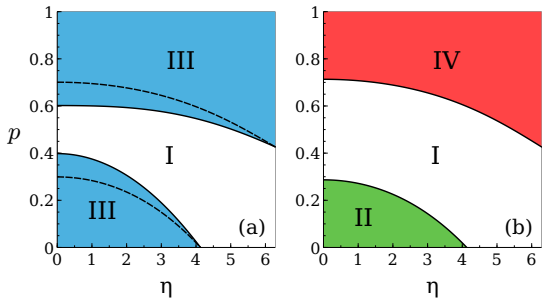

Figure 3.3: Phase diagram of p versus η obtained from kinetic theory. Stable phases at each region are: I (incoherent), II (stationary), III (rotating wave), IV (period 2). Parameter values are $M = \pi R_0^2 N/L^2 = 7$, and: (a) $\xi_0 = 3\pi/4$ (solid line) and $\xi = 9\pi/10$ (dashed line); (b) $\xi_0 = \pi$. For $p = 0$, the lines separating Regions I and III and I and II in the lower half of these panels intersect at the same point, $\eta_c \approx 4.13$, which is the critical noise of the standard VM. Increasing the noise decreases order when η crosses the lower critical lines and increases it when η crosses the upper critical lines. There is a narrow interval of p values where phase I is stable for any η. This interval disappears for large enough M.

$0 < \xi_0 < \pi$ (Fig. 3.3(a)), the noise density is no longer even, the eigenvalue Q_1 is complex, and the order-disorder phase transition occurs with $|Q_1| = 1$. In this case, the eigenfunction is a rotating wave, $\tilde{f}(\theta, t) = e^{i(\theta+\Omega t)}$, $\Omega = \text{Arg}\, Q_1$ and at the critical lines separating the Regions I and III in Fig.3.3(a) there are supercritical Hopf bifurcations. For $\xi_0 = \pi$ (Fig. 3.3(b)), there are supercritical pitchfork bifurcations at the critical line I-II ($Q_1 = 1$) and supercritical period doubling bifurcations at the critical line I-IV ($Q_1 = -1$). Note that increasing the angle ξ_0 makes Region I larger. Fig. 3.3(a) shows the critical lines separating Region I (stable incoherent motion) from rotating wave phases (RWPs) in Region III, namely as ξ_0 approaches the value π, the lower critical line moves down toward the line separating Regions I and II in and $\Omega = \text{Arg}\, Q_1$ approaches zero. Meanwhile, the upper critical line in Fig. 3.3(a) moves up towards the line separating Regions I and IV in Fig. 3.3(b) and Ω approaches -1. As the phase diagrams and numerical simulations of the VM are similar for $\pi/2 < \xi_0 < \pi$, we have selected $\xi_0 = 3\pi/4$ to present our results for RWPs. By an abuse of language, we will call this the case of *almost contrarian compulsions* (even though $\xi_0 = 3\pi/4$ is not close to contrarian compulsions with $\xi_0 = \pi$).

Fig. 3.4 depicts the frequency Ω as a function of alignment noise on the line $\eta(p)$ that separates Regions I and III in the lower part of Fig. 3.3(a) (for $\xi_0 = 3\pi/4$). Note that $\Omega = 0$ at the critical value η_c of the standard VM corresponding to $p = 0$.

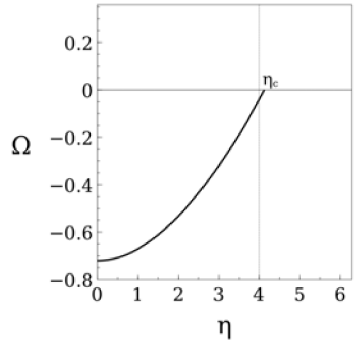

Figure 3.4: Frequency $\Omega = \mathrm{Arg}\, Q_1$ versus alignment noise η along the critical line $p(\eta)$. Parameter values are $\xi_0 = 3\pi/4$ and $M = 7$.

3.3.2.2 Average number of neighbors M vs p

The other parameter appearing in Eq. (3.8) is the average number of neighboring particles, $M = \rho_0 \pi R_0^2$. This parameter changes with the radius of the influence region or the average number density. Fig. 3.5 shows the phase diagram of p versus M for a typical value of alignment noise, $\eta = \pi/3$. Increasing the average number of particles inside the circle of influence favors swarming phases, and therefore Regions II, III and IV (polarized phases) grow at the expense of Region I (zero polarization). For $\xi_0 = 3\pi/4$, Fig. 3.5(a) shows that Region I disappears for M larger than a critical value of about 9. When the average number of neighbors is larger than this critical value, $M_*(\xi_0)$, particles are always polarized. $M_*(\xi_0)$ increases with the angle ξ_0 for a fixed value of the noise strength η and numerical simulations suggest that $M_*(\pi) = +\infty$.

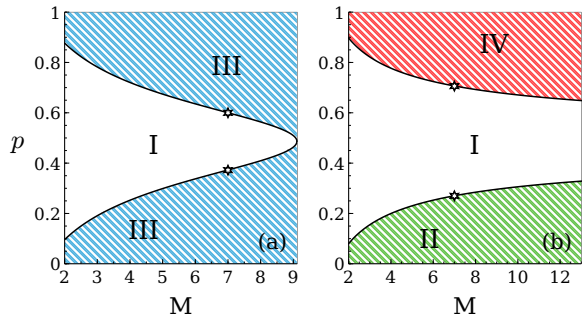

Figure 3.5: Phase diagram of p versus M obtained from kinetic theory. Stable phases at each region are: I (incoherent), II (stationary), III (rotating wave), IV (period 2). Parameter values are $\eta = \pi/3$, and: (a) $\xi_0 = 3\pi/4$ (solid line) and (b) $\xi_0 = \pi$. The value $M = 7$ used in Fig. 3.4 is marked by a star.

3.4 Bifurcation theory

The case of critical eigenvalue $Q_1 = 1$ (pitchfork bifurcation) has been discussed in Chapter 2. Here we shall use the same method to analyze the bifurcations at zero wave number when the critical eigenvalues are $Q_1 = -1$ (period doubling bifurcation) and complex $Q_1 = e^{i\Omega}$ (Hopf bifurcation). We shall use the alignment noise η as a bifurcation parameter and comment on the change in our results had the average particle density been used instead. The solution of the linearized equation

$$\mathcal{L}f^{(1)} \equiv f^{(1)}(\theta, t+1, \mathbf{X}, T) - C^{(1)}[f^{(1)}](\theta, t, \mathbf{X}, T) = 0, \tag{3.12}$$

is

$$f^{(1)}(\theta, t, \mathbf{X}, T, \epsilon) = \frac{r(\mathbf{X}, T; \epsilon)}{2\pi} + A(\mathbf{X}, T; \epsilon) e^{i(\theta + \Omega t)} + \text{cc}, \tag{3.13}$$

$$\Omega = \text{Arg}\, Q_1, \quad \mathbf{X} = \epsilon \mathbf{x}, \quad T = \epsilon t. \tag{3.14}$$

Here cc means the complex conjugate of the preceding term. We do not need to include more terms in Eq. (3.13) because the other modes decay rapidly in the fast time scale t. The argument of the eigenvalue Q_1 is $\Omega = 0$, (pitchfork bifurcation), $\Omega = \pi$ (period doubling bifurcation) and $\Omega \in (-\frac{\pi}{2}, 0)$ (Hopf bifurcation). The two latter cases for $\Omega \neq 0$ can be treated together. Bifurcation theory is quite different in the cases of uniform and space dependent density. In particular, the case of uniform density can be treated by using multiple scales for all bifurcation types: pitchfork (I-II), Hopf (transition I-III) and period doubling (I-IV); see Appendix 3.A. Here we consider space dependent density. We anticipate crossover scalings and therefore we shall use the Chapman-Enskog method [25]. In all three bifurcation cases, the hierarchy of equations resulting from the Chapman-Enskog ansatz [20, 25, 26],

$$f(\theta, \mathbf{x}, t; \epsilon) = f_0 + \epsilon f^{(1)} + \sum_{j=2}^{\infty} \epsilon^j f^{(j)}(\theta, t; r, A, \overline{A}), \tag{3.15}$$

$$\frac{\partial r}{\partial T} = \mathcal{R}^{(0)}(r, A, \overline{A}) + \epsilon \mathcal{R}^{(1)}(r, A, \overline{A}) + O(\epsilon^2), \tag{3.16}$$

$$\frac{\partial A}{\partial T} = \mathcal{A}^{(0)}(r, A, \overline{A}) + \epsilon \mathcal{A}^{(1)}(r, A, \overline{A}) + O(\epsilon^2), \tag{3.17}$$

with $\eta = \eta_c + \epsilon^2 \eta_2$ and $\overline{A}$ the complex conjugate of A, is

$$\mathcal{L}f^{(2)} = C^{(2)}[f^{(1)}, f^{(1)}] - \mathbf{v}\cdot\nabla_X f^{(1)}|_{t+1} - \frac{\mathcal{R}^{(0)}}{2\pi} - \mathcal{A}^{(0)} e^{i(\theta + \Omega t + \Omega)} + \text{cc}, \tag{3.18}$$

$$\mathcal{L}f^{(3)} = C^{(3)}[f^{(1)}, f^{(1)}, f^{(1)}] + 2C^{(2)}[f^{(1)}, f^{(2)}] - \left(\frac{\partial}{\partial T} + \mathbf{v}\cdot\nabla_X\right) f^{(2)}|_{t+1}$$
$$- \frac{\mathcal{R}^{(1)}}{2\pi} - \mathcal{A}^{(1)} e^{i(\theta + \Omega t + \Omega)} + \text{cc} - \frac{1}{2}\left(\frac{\partial}{\partial T} + \mathbf{v}\cdot\nabla_X\right)^2 f^{(1)}|_{t+1} + \eta_2 \frac{\partial}{\partial \eta} C^{(1)}[f^{(1)}], \tag{3.19}$$

etc. In these equations, $C^{(1)}[\varphi]$ is given by Eq. (2.18) with $\mathbf{K} = 0$, and we have

$$C^{(2)}[\varphi,\varphi] = \frac{\pi R_0^2}{2}\left[\sum_{n=2}^{\infty}\frac{ne^{-M}}{(n-2)!}\left(\frac{M}{2\pi}\right)^{n-2}\int_{-\pi}^{\pi}d\xi g(\xi)\int\hat{\delta}(\theta-\xi-\Phi_1)\varphi(\tilde{\theta}_1)\varphi(\tilde{\theta}_2)\prod_{l=1}^{n}d\tilde{\theta}_l\right.$$
$$-2\left(\int_{-\pi}^{\pi}\varphi(\theta_1)d\theta_1\right)\sum_{n=1}^{\infty}\frac{ne^{-M}}{(n-1)!}\left(\frac{M}{2\pi}\right)^{n-1}\int_{-\pi}^{\pi}d\xi g(\xi)\int\hat{\delta}(\theta-\xi-\Phi_1)\varphi(\tilde{\theta}_1)\prod_{l=1}^{n}d\tilde{\theta}_l$$
$$\left.+\frac{M}{2\pi}\left(\int_{-\pi}^{\pi}\varphi(\theta_1)d\theta_1\right)^2,\right. \tag{3.20}$$

$$C^{(3)}[\varphi,\varphi,\varphi] = \frac{\pi^2 R_0^4}{6}\left[\sum_{n=3}^{\infty}\frac{ne^{-M}}{(n-3)!}\left(\frac{M}{2\pi}\right)^{n-3}\int_{-\pi}^{\pi}d\xi g(\xi)\int\hat{\delta}(\theta-\xi-\Phi_1)\varphi(\tilde{\theta}_1)\varphi(\tilde{\theta}_2)\varphi(\tilde{\theta}_3)\prod_{l=1}^{n}d\tilde{\theta}_l\right.$$
$$-3\left(\int_{-\pi}^{\pi}\varphi(\theta_3)d\theta_3\right)\sum_{n=2}^{\infty}\frac{ne^{-M}}{(n-2)!}\left(\frac{M}{2\pi}\right)^{n-2}\int_{-\pi}^{\pi}d\xi g(\xi)\int\hat{\delta}(\theta-\xi-\Phi_1)\varphi(\tilde{\theta}_1)\varphi(\tilde{\theta}_2)\prod_{l=1}^{n}d\tilde{\theta}_l$$
$$+3\left(\int_{-\pi}^{\pi}\varphi(\theta_2)d\theta_2\right)^2\sum_{n=1}^{\infty}\frac{ne^{-M}}{(n-1)!}\left(\frac{M}{2\pi}\right)^{n-1}\int_{-\pi}^{\pi}d\xi g(\xi)\int\hat{\delta}(\theta-\xi-\Phi_1)\varphi(\tilde{\theta}_1)\prod_{l=1}^{n}d\tilde{\theta}_l$$
$$\left.-\frac{M}{2\pi}\left(\int_{-\pi}^{\pi}\varphi(\theta_1)d\theta_1\right)^3\right]. \tag{3.21}$$

and so on. Note that $C_E[f_0+\epsilon\tilde{\rho}] = f_0+\epsilon\tilde{\rho}$ and $C^{(1)}[\tilde{\rho}] = \tilde{\rho}$ for constant $\tilde{\rho}$ imply $C^{(2)}[1,1] = C^{(3)}[1,1,1] = 0$, which can be checked from Eqs. (3.20)-(3.21). The solvability conditions for non-homogeneous equations of the hierarchy is that their right hand sides be orthogonal to the solutions of the homogeneous equation $\mathcal{L}\varphi = 0$, namely 1 and $e^{i(\theta+\Omega t)}$, using the scalar product

$$\langle f(\theta,t), g(\theta,t)\rangle = \lim_{S\to\infty}\frac{1}{S}\sum_{t=0}^{S}\int_{-\pi}^{\pi}\overline{f(\theta,t)}g(\theta,t)\,d\theta. \tag{3.22}$$

In the next two Sections, we consider separately the three bifurcation types.

3.5 Hopf and period doubling bifurcations

For $0 < p \leq 1$ and $0 < \xi_0 < \pi$, we have $Q_1 = e^{i\Omega}$, with $-\pi/2 \leq \Omega < 0$ in Eq. (3.14) at critical noise value η_c located on either the upper or the lower critical lines separating Regions III from Region I in Fig. 3.3(a). Similarly, for $\xi_0 = \pi$, $p > p_0 \approx 2/\sqrt{\pi M}$ [p_0 corresponds to $\eta_c = 2\pi$, $Q_1 = -1$ in Eq. (3.8) for $\xi_0 = \pi$], we have $\Omega = \pi$ at the critical noise η_c separating Regions I and IV in Fig. 3.3(b). Setting a nonzero Ω, we can treat these two cases simultaneously. The solvability conditions for Eq. (3.18) are that its right hand side should be orthogonal to the solutions of the homogeneous equation $\mathcal{L}\varphi(\theta) = 0$, which are 1 and $e^{i(\theta+\Omega t)}$. Using the

scalar product (3.22), these conditions yield

$$\mathcal{R}^{(0)} = 0, \quad \mathcal{A}^{(0)} = \frac{C^{(2)}[1, e^{i\theta}]_1}{\pi} r A e^{-i\Omega} \sim \frac{rA}{2\rho_0}. \tag{3.23}$$

Eq. (3.18) has the solution

$$f^{(2)}(\theta, t, \mathbf{X}, T) = \left[\frac{A^2 e^{i2(\theta+\Omega t)} C^{(2)}[e^{i\theta}, e^{i\theta}]_2}{e^{i2\Omega} - C^{(1)}[e^{i2\theta}]_2} - \frac{e^{i(2\theta+\Omega t)} \left(\frac{\partial}{\partial X} i \frac{\partial}{\partial Y}\right) A}{2(e^{i\Omega} - C^{(1)}[e^{i2\theta}]_2)} \right.$$
$$\left. - \frac{e^{i\Omega(t+1)} \left(\frac{\partial}{\partial X} + i \frac{\partial}{\partial Y}\right) A}{2(e^{i\Omega} - 1)} + \frac{e^{i\theta} \left(\frac{\partial}{\partial X} - i \frac{\partial}{\partial Y}\right) r}{2(e^{i\Omega} - 1)} \right] + \text{cc}. \tag{3.24}$$

Inserting Eqs. (3.13) and (3.24) in Eq. (3.19) and using the solvability conditions, we find $\mathcal{R}^{(1)} = 0$ and $\mathcal{A}^{(1)}$. Then, up to terms of order ϵ^2, the amplitude equations are

$$\frac{\partial r}{\partial T} = 0, \tag{3.25}$$

$$\frac{\partial A}{\partial T} = \left[\frac{r}{2\rho_0} + \epsilon \left(\eta_2 Q_\eta - \frac{r^2 e^{i2\Omega}}{8\rho_0^2} - \mu |A|^2 \right) \right] A + \epsilon \delta \nabla_X^2 A, \tag{3.26}$$

in the limit as $M \gg 1$. In the same limit we also have

$$Q_\eta = \frac{\partial}{\partial \eta} \left(\ln C^{(1)}[e^{i\theta}]_1 \right) \Big|_{\eta_c} \sim \frac{\sqrt{\pi M}}{\eta_c^2} (1 - p) e^{-i\Omega} \left(\frac{\eta_c}{2} \cos \frac{\eta_c}{2} - \sin \frac{\eta_c}{2} \right), \tag{3.27}$$

and we prove in Appendix 3.B that δ and μ both have positive real parts. According to Eq. (3.25), r is independent of T. Conservation of the number of particles implies

$$\int r(\mathbf{X})\, d\mathbf{X} = 0. \tag{3.28}$$

3.5.1 Complex Ginzburg-Landau equation for rescaled r

For $\epsilon = 0$, the solution of Eq. (3.26) yields an $|A|$ that increases in time for $r(\mathbf{X}) > 0$ and it decreases for $r(\mathbf{X}) < 0$. This indicates that a dominant balance occurs only if we assume $r(\mathbf{X}) = O(\epsilon)$. Then Eq. (3.26) becomes a modified complex Ginzburg-Landau equation (CGLE) with a diffusive scaling for the time:

$$\frac{\partial A}{\partial (\epsilon^2 t)} = \delta \nabla_X^2 A + \left[\eta_2 Q_\eta + \frac{r(\mathbf{X})}{2\rho_0} - \mu |A|^2 \right] A. \tag{3.29}$$

If the density is kept uniform, $r = 0$, Eq. (3.29) is the usual CGLE. In this case, it has the rotating wave solution

$$A(\epsilon^2 t) = \sqrt{\frac{\eta_2 \mathrm{Re} Q_\eta}{\mathrm{Re}\mu}}\, e^{i\eta_2 \varphi \epsilon^2 (t-t_0)}, \tag{3.30}$$

$$\varphi = \mathrm{Im} Q_\eta - \mathrm{Im}\mu\, \frac{\mathrm{Re} Q_\eta}{\mathrm{Re}\mu}. \tag{3.31}$$

As proven in Appendix 3.B, $\mathrm{Re}\mu > 0$, and therefore the phases issuing forth from Hopf and period doubling bifurcations are both supercritical: they exist only for $\eta < \eta_c$ (where the uniform distribution is unstable) and are linearly stable against space-independent disturbances. The polarization corresponding to the bifurcating solutions given by Eqs. (3.30)-(3.31) is the modulus of the complex parameter,

$$Z = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j} = \frac{1}{N}\int e^{i\theta} f(\theta, \mathbf{x}, t) d\theta d\mathbf{x} \sim \frac{2\pi}{\rho_0}\sqrt{\frac{(\eta-\eta_c)\mathrm{Re} Q_\eta}{\mathrm{Re}\mu}}\, e^{i\theta_0 - i[\Omega + (\eta-\eta_c)\varphi]t}. \tag{3.32}$$

Near η_c, we have $Q_1 \sim e^{i\Omega}[1 + Q_\eta(\eta - \eta_c)]$. Then we can replace $e^{-i\Omega}Q_1 - 1$ instead of $Q_\eta(\eta - \eta_c)$ in Eq. (3.32), thereby obtaining a formula that holds for larger values of $|\eta - \eta_c|$,

$$Z \sim \frac{2\pi}{\rho_0}\sqrt{\frac{\mathrm{Re}(e^{-i\Omega}Q_1) - 1}{\mathrm{Re}\mu}}\, e^{i\theta_0 - i(\Omega+\psi)t}, \tag{3.33}$$

$$\psi = \mathrm{Im}(e^{-i\Omega}Q_1) - \frac{\mathrm{Re}(e^{-i\Omega}Q_1) - 1}{\mathrm{Re}\mu}\,\mathrm{Im}\mu. \tag{3.34}$$

The bifurcating solutions given by Eqs. (3.30)-(3.31) are uniform in space. However, Eq. (3.26) is a nonlinear reaction-diffusion equation with a diffusion coefficient whose real part is positive according to Eq. (3.68). For a nonuniform particle density, Eq. (3.28) (conservation of the number of particles) strongly suggests the formation of ordered clusters. Let us imagine that $r(\mathbf{X}) = s^2\rho_0 \mathrm{sign}(X - L/2)$, $s^2 > \eta_2 \mathrm{Re} Q_\eta$. Then $W = 0$ for $X > L/2$, whereas for $X < L/2$ we have $W > 0$, given approximately by Eq. (3.30) with $\eta_2 \mathrm{Re} Q_\eta + s^2$ instead of $\eta_2 \mathrm{Re} Q_\eta$. The phase of the complex order parameter satisfies the integrated form of the Burgers equation [27]

$$\frac{\partial \Upsilon}{\partial \sigma} = \kappa \nabla_X^2 \Upsilon - \frac{1}{2}|\nabla_X \Upsilon|^2, \tag{3.35}$$

$$\Upsilon = 2\left(\frac{\mathrm{Im}\delta}{\mathrm{Re}\delta} - \frac{\mathrm{Im}\mu}{\mathrm{Re}\mu}\right)\mathrm{Arg}(A), \tag{3.36}$$

$$\sigma = \epsilon^2 \mathrm{Re}\delta\, t, \quad \kappa = 1 + \frac{\mathrm{Im}\delta\, \mathrm{Im}\mu}{\mathrm{Re}\delta\, \mathrm{Re}\mu}. \tag{3.37}$$

Near the bifurcation line I-III in Fig. 3.3(a), $\kappa > 0$. For the geometry we are considering, Υ depends only on the coordinate X and on σ. Then $k = \partial\Upsilon/\partial X$ satisfies the Burgers equation proper. Assuming that $L \gg 1$, $k \sim k_2 > 0$ at $X = 0$ and $k = 0$ at $X = L/2$, k is the shock wave solution given by Eq. (4.23) of Ref. [28]:

$$\frac{\partial\Upsilon}{\partial X} = \frac{k_2}{1 + \exp\left[\frac{k_2}{2\kappa}(X - \frac{k_2}{2}\sigma - X_0)\right]}. \tag{3.38}$$

This solution represents a planar wave front moving to the right with velocity $k_2/2$. The front encroaches an unpolarized region with zero wave number and leaves a polarized region with wave number k_2 on its wake. The region behind the wave front is a cluster rotating with angular velocity Ω and local wave number proportional to k_2. This simple example illustrates how a non-constant density may produce inhomogeneous ordered clusters of the RWP.

Eq. (3.30) and related plane wave solutions of the CGLE for $r(\mathbf{X}) = 0$ become unstable provided $\kappa < 0$ in Eq. (3.35). Equivalently, $b_3 < b_1$, $b_1 = \mathrm{Im}\delta/\mathrm{Re}\delta$ and $b_3 = -\mathrm{Re}\mu/\mathrm{Im}\mu$ (Newell's criterion). Close to the line $b_1 = b_3$, one can derive the Kuramoto-Sivashinsky equation, [29, 30, 31]. Phase turbulence consisting of disordered cellular structures is then possible. Solutions of the CGLE with periodic boundary conditions also include spiral waves and other defects (having $W = 0$ at one point in their cores), as well as phases of defect turbulence in which defects are created and annihilated in pairs [29]. For the period 2 solution, Eq. (3.13) with $\Omega = \pi$, δ and μ are real. Then Eq. (3.29) has vortex solutions with nonzero rotation number and a vortex gas evolves as indicated in Refs. [29, 30, 31, 32, 33].

3.5.2 Complex Ginzburg-Landau equation for random density disturbances

Let us assume now that $\rho(\mathbf{X}) - \rho_0 = \epsilon r(\mathbf{X})$ is a zero-mean random Gaussian process with standard deviation $\sigma_r = \epsilon\tilde{\sigma}_r$. Then the average value $\langle rA\rangle \approx \langle r\rangle\langle A\rangle = 0$, and the mean amplitude, $\tilde{A} = \langle A\rangle$, satisfies the approximate equation:

$$\frac{\partial\tilde{A}}{\partial(\epsilon^2 t)} = \delta\,\nabla_X^2\tilde{A} + \left[\eta_2 Q_\eta - \frac{\tilde{\sigma}_r^2}{8\rho_0^2} - \mu|\tilde{A}|^2\right]\tilde{A}. \tag{3.39}$$

Here we have made $\langle r^2 A\rangle \approx \tilde{\sigma}_r^2\tilde{A}$, $\langle|A|^2 A\rangle \approx |\tilde{A}|^2\tilde{A}$. Now the uniform solution is Eq. (3.30) in which $\eta_2\mathrm{Re}Q_\eta$ is replaced by $[\eta_2\mathrm{Re}Q_\eta - \tilde{\sigma}_r^2/(8\rho_0^2)]$. Then $\mathrm{Re}(e^{-i\Omega}Q_1) - 1$ in Eq. (3.33) for the order parameter is replaced by $[\mathrm{Re}(e^{-i\Omega}Q_1) - 1 - \sigma_r^2/(8\rho_0^2)]$, with the result

$$W = |Z| \sim \frac{2\pi}{\rho_0}\sqrt{\frac{\mathrm{Re}(e^{-i\Omega}Q_1) - 1 - \frac{\sigma_r^2}{8\rho_0^2}}{\mathrm{Re}\mu}}. \tag{3.40}$$

Equating to zero this last quantity, we find the critical value of the noise, η_c, which gets shifted to a smaller value. How large is $\sigma_r^2 = \langle r^2 \rangle$? We know that the particles are randomly placed in the box at the initial time. The fluctuation of the density is

$$\langle (\rho - \rho_0)^2 \rangle = \frac{\rho_0^2}{N} \frac{k_B T}{-v^2 \frac{\partial P}{\partial v}} = \frac{\rho_0^2}{N}, \tag{3.41}$$

where k_B, T, P, and $v = 1/\rho_0$ are the Boltzmann constant, temperature, pressure and specific volume, respectively [cf. Eq. (7.43) in Ref. [34]]. The particles can be thought of as belonging to an ideal gas at the initial time, therefore $P = \rho_0 k_B T$ and $-v^2 \partial P/\partial v = \partial P/\partial \rho_0 = k_B T$. Then we have $\langle (\rho - \rho_0)^2 \rangle = \rho_0^2/N$, as written in Eq. (3.41), and Eq. (3.40) becomes

$$W = |Z| \sim \frac{2\pi}{\rho_0} \sqrt{\frac{\mathrm{Re}(e^{-i\Omega} Q_1) - 1 - \frac{1}{8N}}{\mathrm{Re}\mu}}. \tag{3.42}$$

The shift in the bifurcation point indicated by Eq. (3.42) vanishes in the limit as $N \to \infty$.

3.5.3 Average particle density as bifurcation parameter

What happens if we select the average particle density ρ_0 as bifurcation parameter instead of the alignment noise η?

Firstly, we have to replace $\rho_0 = \rho_{0c} + \epsilon^2 \rho_2$ instead of $\eta = \eta_c + \epsilon^2 \eta_2$ and $f^{(2)} + \rho_2$ instead of $f^{(2)}$ in Eq. (3.15). Then a term $\rho_2 \pi R_0^2 \partial C^{(1)}[f^{(1)}]/\partial M$ replaces the last term in the right hand side of Eq. (3.19), to which we have to add a term $2\rho_2 C^{(2)}[1, f^{(1)}]$. For all bifurcation types, $\rho_2 Q_\rho$ replaces $\eta_2 Q_\eta$ in the amplitude equations. Q_ρ is defined as the derivative with respect to ρ_0 instead of the derivative with respect to η in the first line of Eq. (3.27). Secondly, for Hopf or period doubling bifurcations, we replace $r(\mathbf{X}) + 2\pi\rho_2$ instead of $r(\mathbf{X})$ in Eqs. (3.28) and (3.29). The remaining considerations are the same provided we make these replacements. One obvious change is that $\mathrm{Re}Q_\rho > 0$ (cf. Fig 3.5), whereas $\mathrm{Re}Q_\eta < 0$ on the lower critical lines of Fig 3.3, and $\mathrm{Re}Q_\eta > 0$ on the upper critical lines of the same figure. Thus, the disordered phase is stable for $\rho_0 < \rho_{0c}$ and unstable for $\rho_0 > \rho_{0c}$, whereas the situation is the opposite for the polarized random wave and stationary phases on the lower sectors of Fig 3.3 if we use the alignment noise as a bifurcation parameter.

3.6 Pitchfork bifurcation

For $p = 0$ (standard VM) or $p > 0$ and $\xi_0 = \pi$, corresponding to the lower sector of the phase diagram in Fig. 3.3(b), the critical condition is $Q_1 = 1$, and therefore $\Omega = 0$ in Eq. (3.14). The corresponding bifurcation has been analyzed in Chapter 2 for the case $p = 0$. The same procedure based on the solvability conditions for Eqs. (3.18) and (3.19) produces nonzero $R^{(0)}$ and $R^{(1)}$ and the density disturbance is no longer time independent. The amplitude equations are equivalent to the following system for the density disturbance and a current

density $\mathbf{w} = 2\pi(\mathrm{Re}A, -\mathrm{Im}A)$:

$$\frac{\partial r}{\partial T} + \nabla_X \cdot \left[\left(1 + \frac{\epsilon\gamma_3 r}{2\pi}\right)\mathbf{w}\right] = 0, \tag{3.43}$$

$$\frac{\partial \mathbf{w}}{\partial T} = -\frac{1}{2}\nabla_X\left[\left(1 - \frac{\epsilon\gamma_3}{4\pi}r\right)r + \frac{\epsilon(2\gamma_1 - \gamma_2 - \gamma_3)}{2\pi}|\mathbf{w}|^2\right] + \epsilon\frac{2\gamma_1 - \gamma_2 + \gamma_3}{2\pi}\mathbf{w}(\nabla_X \cdot \mathbf{w})$$
$$+\epsilon\frac{2\gamma_1 + \gamma_2 - \gamma_3}{2\pi}(\mathbf{w} \cdot \nabla_X)\mathbf{w} + \epsilon\delta\nabla_X^2\mathbf{w} + \left[\frac{\gamma_3 r}{\pi} + \epsilon\left(\eta_2 Q_\eta - \frac{r^2}{6\rho_0^2} - \frac{\mu|\mathbf{w}|^2}{4\pi^2}\right)\right]\mathbf{w}. \tag{3.44}$$

The coefficients appearing in these equations are all real valued and listed in Appendix 3.B. For $r = 0$ and $\gamma_3 = 0$, these equations are exactly equivalent to the amplitude equation (132) of Ref. [9]. As in the case of Eq. (3.26), δ and μ are both positive. For $\epsilon = 0$, Eq. (3.43) is the continuity equation for a density variable r and a current density $\mathbf{w}$, which explains the name of the latter variable. The overall density of particles is N/L^2, which implies the following constraint for $r(\mathbf{X}, T)$:

$$\int r(\mathbf{X}, T)\, d\mathbf{X} = 0. \tag{3.45}$$

3.6.1 Space independent A and r=0: diffusive scaling

For space independent A, $r = 0$, and Eq. (3.44) is the typical pitchfork amplitude equation with diffusive scaling. It has a stationary solution with modulus

$$|\mathbf{w}_0| = 2\pi\sqrt{\frac{\eta_2 Q_\eta}{\mu}}, \tag{3.46}$$

which is stable, and it exists for $\eta < \eta_c$, $\mu > 0$, and $Q_\eta < 0$. For $\eta < \eta_c$, the uniform distribution f_0 is unstable as $Q_1 > 1$ there. Thus the transition from incoherence to order is a supercritical pitchfork bifurcation. The polarization corresponding to this solution is

$$W_0 \sim \frac{2\pi}{\rho_0}\sqrt{\frac{(\eta - \eta_c)Q_\eta}{\mu}}. \tag{3.47}$$

Near η_c, we have $Q_1 \sim 1 + Q_\eta(\eta - \eta_c)$. Then we can replace $Q_1 - 1$ instead of $Q_\eta(\eta - \eta_c)$ in Eq. (3.47), thereby obtaining a formula that holds for larger values of $|\eta - \eta_c|$:

$$W_0 \sim \frac{2\pi}{\rho_0}\sqrt{\frac{Q_1 - 1}{\mu}}. \tag{3.48}$$

3.7 Results of simulations of the modified VM

We have performed numerical simulations of our modified VM in the different regions of Fig. 3.3 indicated by the linear stability analysis of the kinetic equation given in Section 3.3.

These regions and the predictions of bifurcation theory in Sections 3.5 and 3.6 are as follows:

(I) Incoherent motion with polarization $W = 0$.

(II) Stationary coherent motion with $W > 0$ for $\xi_0 = \pi$ and small p.

(III) Rotating wave phase with coherent motion (swarm rotation [40]), such that in Eq. (2.4), $W > 0$, $\Upsilon = \Omega t + \Upsilon_0$ for $\xi_0 < \pi$: the order parameter $Z(t)$ describes a circle about the origin at constant angular velocity as in Fig. 3.8(a); see also Fig. 3.9(a).

(IV) Coherent period 2 motion as in Figs. 3.8(b) and Fig. 3.9(b).

At the critical lines separating the regions of Figs. 3.3(a) and 3.3(b), we have found the following supercritical bifurcations from incoherent motion with uniform particle density: pitchfork (I to II), Hopf (I-III), and period doubling (I-IV).

Pitchfork bifurcation. Figs. 2.4 presented in Chapter 2 and 3.6(a) compare the polarization obtained from direct numerical simulations of the VM with the theoretical curves of Eqs. (3.47) and (3.48) for $p = 0$ and $p = 0.1$ (transition I-II in Fig. 3.3), respectively. The polarization shown in these figures is an ensemble average over 10 replicas of the stochastic process. In Fig. 2.4, we observe that the results of numerical simulations tend to the uniform solution predicted by bifurcation theory as we increase the density ρ_0. This indicates that the simulations produce solutions that are closer to be independent of space. However in Fig. 3.6(a), the results of simulations depart appreciably from the theoretical prediction, which indicates that flocking is not homogeneous. Although in both cases the bifurcation is pitchfork, there is a striking difference between the polarization curves for the standard and modified VM. For the standard VM, the maximum polarization ($W = 1$) is reached for zero alignment noise $\eta = 0$. However, for $p = 0.1$ [transition I-II in Fig. 3.3(b)], the polarization shown in Fig. 3.6(a) is maximal for a nonzero value of η. A similar behavior also occurs for RWPs, as explained later in this Section. Fig. 3.7 shows three snapshots of 1000 particles for the VM with contrarian compulsions for $M = 7$, $\rho_0 = 3$, $p = \eta = 0.1$. The insets show the location of all the particles in three time instants. We observe that there are small clusters that move and persist in time, with dynamics as shown in Fig. 3.7. Had different realizations of the stochastic process shown one large cluster and a number of freely moving particles, the result of ensemble averaging would be close to that of Eq. (3.48) for homogeneous flocking. Ensemble averages of heterogeneous flocking such as that depicted in Fig. 3.7 give a polarization that differs markedly from Eq. (3.48), as shown in Fig. 3.6(a).

Hopf bifurcation. Figs. 3.8(a) and 3.9(a) show that the complex order parameter of the RWP in the lower region III of Fig. 3.3(a) is close to the uniform values of Eqs. (3.32)-(3.33) predicted by bifurcation theory. The agreement between simulations of the VM and uniform solutions predicted by bifurcation theory is not as good for the period 2 phase, as shown in Figs. 3.8(b) and 3.9(b).

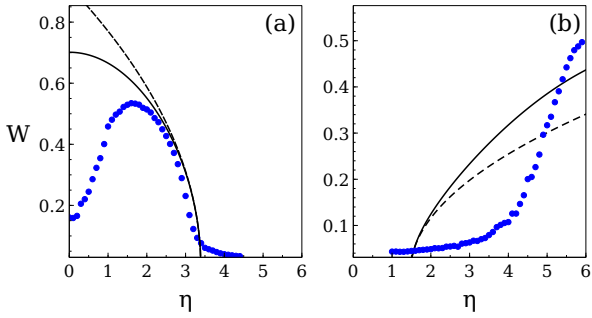

Figure 3.6: Polarization W versus η for $\xi_0 = \pi$, $M = 7$, $\rho_0 = 10$, $N = 1000$, and (a) lower sector in Fig. 3.3(b) ($p = 0.1$); (b) upper sector in Fig. 3.3(b) ($p = 0.7$). Circles are numerical simulation data whereas solid and dashed lines correspond to kinetic theory approximations of Eqs. (3.48) or (3.47), respectively. In (b), flocking is heterogeneous for any η and results from simulations are quite different from theoretical lines corresponding to the uniform amplitude.

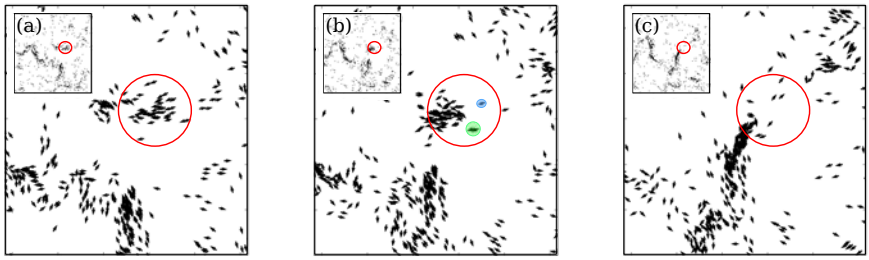

Figure 3.7: Three snapshots (a), (b) and (c) showing heterogeneous clustering for particles with contrarian compulsions. The insets show the position of the red circle on the complete box. A cluster of particles leaves the red circle, joins another set of particles and forms an elongated cluster after some time. Close-by particles may have one direction and its contrary, as the marked particles in panel (b). Parameter values are $\xi_0 = \pi$, $M = 7$, $\rho_0 = 3$, $p = \eta = 0.1$, $N = 1000$.

Fig. 3.10 describes the transition I-III in the lower sector of Fig. 3.3(a) for almost contrarian compulsions. Solid and dashed lines correspond to spatially uniform coherent phases calculated from kinetic theory. Departure of simulation data from these lines indicates heterogeneous RWPs. Observe the dispersion of simulation data in Fig. 3.8(a). As the average density $\rho_0 = N/L^2$ increases, the phases become more uniform and simulation data approach theoretical predictions. It is interesting to note that simulations of the VM with backward update produce polarizations closest to the theoretical curves except very close to the bifurcation point (which is due to finite size effects).

Fig. 3.11(a) shows that, for small values of p and η, forward update with almost contrarian compulsions produces small clusters and many seemingly free particles. Compared with the same VM and parameters but with backward update, Fig. 3.11(b) shows one large cluster

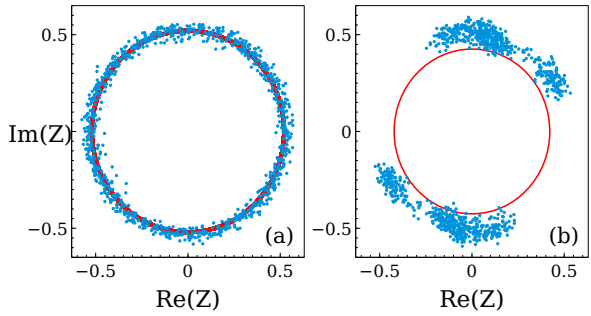

Figure 3.8: Complex flocking order parameter $Z(t)$ for $M = 7$, $\rho_0 = 10$, $N = 1000$, and: (a) $\xi_0 = 3\pi/4$, $p = 0.1$, $\eta = 2.5$, marked by an arrow in Fig. 3.10(a). The order parameter for the RWP traces a circle (theory: solid line), implying a nonzero angular velocity. (b) $\xi_0 = \pi$, $p = 0.7$, $\eta = 5.9$. The order parameter alternates between points with a phase difference of π, marking abrupt oscillations in flocking (theory: two opposite points on the circle). Sampling time is 1 and transients have been eliminated.

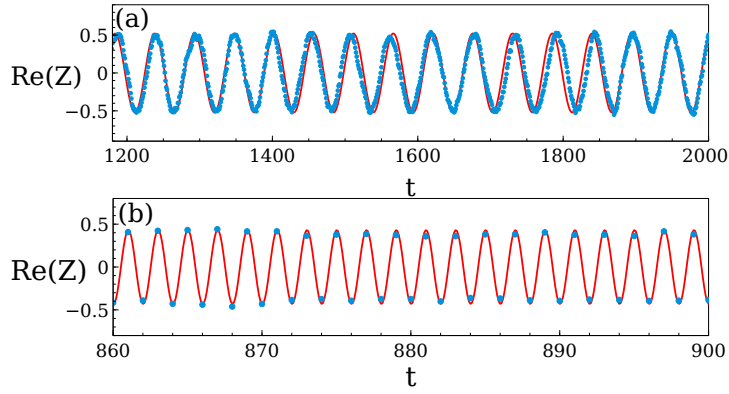

Figure 3.9: Real part of $Z(t)$ corresponding to: (a) Fig. 3.8(a); and (b) Fig. 3.8(b). Numerical simulations (circles) are compared to spatially uniform phases from kinetic theory (solid lines) given by Eq. (3.33) with $\psi = 0$.

and a small number of free particles. When averaging over many replicas of the stochastic process, backward update produces a polarization consistent with the homogeneous flocking prediction of Eq. (3.33). However, the persistence of several clusters in the case of forward update produces a polarization that does not correspond to homogeneous flocking.

For the VM of Eqs. (3.1)-(3.3) with forward update, there is an appreciable shift of the simulation data to smaller values of the noise η near the bifurcation point. This shift decreases as the average density ρ_0 increases but it does not disappear, as the comparison between Figs. 3.10(a) and 3.10(b) indicates. The same shift also occurs in the standard VM: Even for an average density as large as $\rho_0 = 20.4$, the critical noise η_c is different in direct

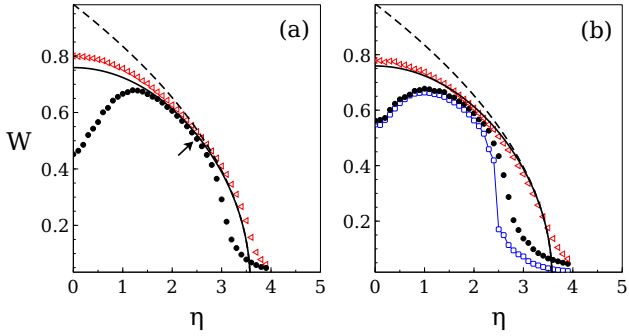

Figure 3.10: Polarization W versus η for the lower sector III of Fig. 3.3(a). Simulation data obtained by ensemble averages over 10 replicas of the VM for $\xi_0 = 3\pi/4$, $p = 0.1$, $M = 7$, and (a) $\rho_0 = 10$, $N = 1000$; (b) $\rho_0 = 3$, $N = 1000$ (filled circles) and $N = 5000$ (blue circles). Empty red triangles arise from simulating the VM with backward update instead of Eq. (3.2). Solid and dashed lines correspond to kinetic theory approximations of Eqs. (3.33) and (3.32), respectively. As ρ_0 increases, simulation data approach the line of the spatially uniform RWP with order parameter given by Eq. (3.33). Departure from this line indicates a nonuniform RWP. Near η_c, simulation data depart from the theoretical curve due to finite size effects. Panel (b) shows that the transition I-III becomes discontinuous for Eqs. (3.1)-(3.2) at larger box sizes ($L = \sqrt{N/\rho_0}$).

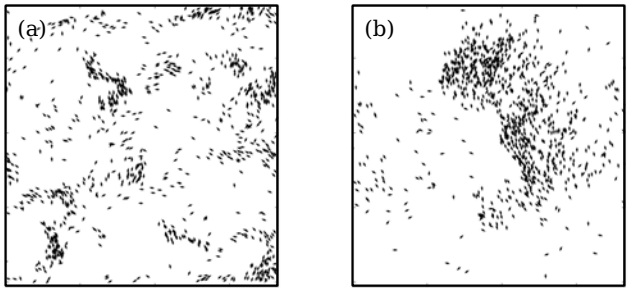

Figure 3.11: Two snapshots showing clustering for particles with almost contrarian compulsions for (a) forward, and (b) backward update. With forward update, clustering is heterogeneous, whereas with backward update, nonzero polarization is attained by one large cluster and a gas of particles outside it. Parameter values are $\xi_0 = 3\pi/4$, $M = 7$, $\rho_0 = 3$, $p = \eta = 0.1$, $N = 1000$. Values of the polarization W are: (a) 0.57, (b) 0.78.

simulations and theory, as noted in the different rescaling in Fig. 3 of Ref. [7]. Fig. 3.10(b) shows that W departs the solid line and the bifurcation becomes discontinuous for box size beyond a critical value. This change also occurs in the standard VM, except that the phase transition is from incoherence to flocking with nonzero average velocity [35, 36, 37, 41]. Fig. 10 of Ref. [36] shows that the critical length decreases as the average particle density

increases for the standard VM. The density $\rho_0 = 3$ considered in Fig. 3.10(b) is three times larger than the largest one in Fig. 10 of Ref. [36] ($\rho_0 = 1$ in our nondimensional units), which is why we observe a discontinuous transition for a length as small as $L = \sqrt{5000/3} = 40.8$.

Period doubling bifurcation. Strikingly, when the probability of contrarian movement is sufficiently high, increasing the alignment noise favors time periodic, spatially heterogeneous, flocking. Increasing η (tolerance to failure in particle alignment) in the upper part of Region I in Fig. 3.3 favors forming clusters. For $p > p_0 \approx 2/\sqrt{\pi M}$ [cf. Eq. (3.8)] and small η, preponderance of contrarian over conformist motion ensures incoherence of motion. The upper lines in Fig. 3.3 suggest transition from incoherent motion to a degree of flocking as η surpasses a critical value. Figs. 3.6(b) and 3.8(b) for contrarian compulsions show that the polarization W calculated from simulations departs markedly from the theoretical line corresponding to uniform density. Figs. 3.8(b) and 3.9(b) illustrate that the complex order parameter alternates between numbers with arguments differing by π. Thus, the ordered phase is periodic in time with period 2 although, as shown in Figs. 3.8(b) and 3.9(b), the amplitude of the oscillation has an envelope that wanders in a certain region of the complex plane. The dispersion of points in Fig. 3.8(b) is due to the formation of clusters with varying size that change in time and produce a nonuniform density. The persistence of heterogeneous clustering yields ensemble averaged polarizations that differ from the prediction of Eq. (3.33) with $\Omega = \pi$ for spatially homogeneous flocking. Sufficient tolerance to failures in the alignment of the conformist particles keeps nonuniform flocking at the expense of back and forth motion of the flocks between opposite average phases of the order parameter. Active particles in exotic phases perform rotations and oscillations, not just translations as in the standard VM.

Optimal noise. Figs. 3.6(a) and 3.10 indicate that the polarization W increases with alignment noise until it reaches a maximum: *To attain maximum flocking we need an optimum degree of alignment noise*, for both Hopf and pitchfork bifurcations in the lower sectors of Figs. 3.3(a) and 3.3(b). For larger noise, W decreases and the values obtained from simulations approach the theoretical curve for uniform particle density. With low probability of contrarian motion and small η, particle clusters form, move coherently and change size (heterogeneous flocking), as shown in Figs. 3.7 and 3.11(a).

3.8 Conclusions

We have proposed a modified Vicsek model in which active particles may align their velocity with the local average direction of motion or with the (almost) opposite direction. Flocking behavior depends on the probability p of (almost) contrarian compulsions compared with that of conformist alignment according to the Vicsek rule. Besides disordered and ordered phases with almost zero and finite stationary polarization, respectively, numerical simulations have shown that our VM exhibits novel exotic phases (stable solutions) with time periodic polarization order parameter. In the ordered phases, flocking is heterogeneous in space

and reminiscent of micro-flocking in Ref. [16]. Strikingly, in the presence of contrarian compulsions, increasing the Vicsek alignment noise may favor order in two ways. For small p, there is a nonzero optimal noise value for which polarization is maximal and reflects a trend to homogeneous flocking. When contrarian compulsions are prevalent (p closer to 1), increasing the alignment noise may transform incoherent particle motion to a phase displaying time periodic polarization with period 2. If we relax the contrarian rule so that particles may select deflection by some fixed large angle ξ_0 ($\pi/2 < \xi_0 < \pi$) measured counterclockwise from the average direction, the flocking order parameter may oscillate periodically in time (with period different from 2, which occurs for $\xi_0 = \pi$). Active particles perform rotations or oscillations besides the collective translation characterizing the ordered phase of the standard VM. Rotation allows active particles to explore larger regions of space and may be advantageous in emergency escape of a crowd from a confined region with several exits [24].

To interpret and understand the results of our numerical simulations, we have analyze near the transition to flocking a kinetic theory proposed by T. Ihle for the standard VM [11]. The exotic phases appear as Hopf and period doubling bifurcations from the uniform distribution function at a critical value of the noise. Our bifurcation calculations follow those we have developed for the standard VM [26]. We find that the equation for the complex amplitude of the bifurcating solution is a modified complex Ginzburg-Landau equation (CGLE). This equation has a uniform rotating wave solution whose polarization has critical exponent $1/2$ at the bifurcation point. Deviation of the numerically obtained polarization from this theoretical curve indicates heterogeneous flocking.

Stable phases with time dependent order parameter may have appeared in earlier work. For example, Chaté *et al*. introduced a Vicsek-like model for apolar nematic active particles can move with equal probability along their orientation θ_i or along the contrarian orientation $\theta_i + \pi$ [42]. At one time step later, the orientation is chosen as the first eigenvector of a tensor order parameter plus disorder noise. They observed a continuous transition similar to the Kosterlitz-Thouless transition, characterized by large spatial fluctuations of the time-averaged order parameter [42]. This is different from the discontinuous transition observed for the standard (polar) VM if the box size is sufficiently large [35]. In the case of our modified VM with contrarian or almost contrarian compulsions, the continuous bifurcation becomes discontinuous for sufficiently large box sizes. Clustering in our VM with almost contrarian compulsions is reminiscent of microflocking in Ref. [16]. In our case, the probability of deflecting a large angle from the conformist mean direction is responsible for the rotation inside heterogeneous flocks, and we do not need to impose an external common angular velocity to achieve rotating clusters. Menzel [43] studied a similar model to that in Ref. [16] for two different populations of particles, that had Kuramoto coupling but without the constant rotation. He found a variety of behaviors including clustering and stripe patterns, but not clusters of synchronously rotating particles. Lastly, Chepizhko *et al*. considered a similar model without constant rotation and for a single particle species, which interacts with obstacles that could be fixed or diffusing in space, [44]. For this quite different system, they observed

that the time-averaged order parameter exhibited a maximum for an optimal noise strength, a phenomenon similar to that presented in our Figs. 3.6(a) and 3.10.

Appendix 3.A Methods

Numerical methods. At each time step, we have updated the position and velocity of N active particles using Eqs. (3.1)-(3.2). To select the values ξ_i, we first choose a number between 0 and 1 randomly with uniform probability. If this number is p or smaller, we set $\xi = \xi_0$. Otherwise, ξ is selected randomly in the interval $(-\eta/2, \eta/2)$.

Regular perturbation theory for the eigenvalues of $\mathcal{M}_{ij}$. Assume that the matrix $\mathcal{M}(Q) = \mathcal{D}(Q) + \varepsilon\mathcal{N}(Q)$, where $\varepsilon\mathcal{N}_{ij}(Q)$ are the off-diagonal terms of $\mathcal{M}(Q)$ and $\varepsilon \ll 1$. The eigenvalues, $\lambda(\varepsilon) = \sum_{j=0}^{\infty}\lambda^{(j)}\varepsilon^j$, and eigenfunctions $\psi(\varepsilon) = \sum_{j=0}^{\infty}\psi^{(j)}\varepsilon^j$ of $\mathcal{M}(Q)$ can be expanded in powers of the scaling parameter ε and inserted in the eigenvalue equation for $\mathcal{M}(Q)$. As $\varepsilon \to 0$, we obtain the hierarchy of linear equations

$$(\mathcal{D} - \lambda^{(0)}\mathcal{I})\psi^{(0)} = 0, \tag{3.49}$$

$$(\mathcal{D} - \lambda^{(0)}\mathcal{I})\psi^{(1)} = (\lambda^{(1)}\mathcal{I} - \mathcal{N})\psi^{(0)}, \tag{3.50}$$

$$(\mathcal{D} - \lambda^{(0)}\mathcal{I})\psi^{(2)} = (\lambda^{(1)}\mathcal{I} - \mathcal{N})\psi^{(1)} + \lambda^{(2)}\psi^{(0)}, \tag{3.51}$$

etc. The first equation says that $\lambda^{(0)}$ are the diagonal elements of $\mathcal{M}$, $\lambda_j^{(0)} = QJ_0(|\mathbf{K}|) - C^{(1)}[e^{ij\theta}]_j$ ($j = 1, \ldots$). For the other non-homogeneous linear equations to have solutions, their right hand sides have to be orthogonal to the eigenvectors $\psi^{(0)}$. The jth eigenvector has components δ_{lj}. The orthogonality condition for Eq. (3.49) produces $\lambda_j^{(1)} = \mathcal{N}_{jj} = 0$. Eq. (3.50) becomes

$$(\lambda_i^{(0)} - \lambda_j^{(0)})\psi_i^{(1)} = -\mathcal{N}_{ij} \Longrightarrow \psi_i^{(1)} = \frac{\mathcal{N}_{ij}}{\lambda_j^{(0)} - \lambda_i^{(0)}}, \tag{3.52}$$

for $i \neq j$ and $\psi_j^{(1)} = 0$. We now insert Eq. (3.52) into Eq. (3.51) and use the orthogonality condition to obtain

$$\varepsilon^2\lambda_j^{(2)} = \varepsilon^2\sum_{l,l\neq j}\frac{\mathcal{N}_{jl}\,\mathcal{N}_{lj}}{\lambda_j^{(0)} - \lambda_l^{(0)}} = -Q^2\sum_{l,l\neq j}\frac{(e^{i\mathbf{K}\cdot\mathbf{v}+i(j-l)\theta})_0(e^{i\mathbf{K}\cdot\mathbf{v}+i(l-j)\theta})_0}{C^{(1)}[e^{ij\theta}]_j - C^{(1)}[e^{il\theta}]_l}, \tag{3.53}$$

with $\mathbf{v} = (\cos\theta, \sin\theta)$. In this expression, we have factors:

$$\begin{aligned}(e^{i\mathbf{K}\cdot\mathbf{v}+in\theta})_0 &= \frac{1}{2\pi}\int_{-\pi}^{\pi}e^{i|\mathbf{K}|\cos(\theta-\mathrm{Arg}\mathbf{K})+in\theta}d\theta\frac{e^{in(\mathrm{Arg}\mathbf{K}+\pi/2)}}{2\pi}\int_{-\pi}^{\pi}e^{i|\mathbf{K}|\sin\zeta-in\zeta}d\zeta \\ &= e^{in(\mathrm{Arg}\mathbf{K}+\pi/2)}J_n(|\mathbf{K}|),\end{aligned} \tag{3.54}$$

where we have changed variable $\theta = \mathrm{Arg}\mathbf{K} + \pi/2 - \zeta$, shifted the limits of integration and used the integral representation for the Bessel function $J_n(x)$ in Ref. [45]. Inserting Eq. (3.54) in Eq. (3.53) and using $J_{-n}(|\mathbf{K}|) = (-1)^n J_n(|\mathbf{K}|)$, we get the eigenvalues

$$\lambda_j = QJ_0(|\mathbf{K}|) - C^{(1)}[e^{ij\theta}]_j - Q^2 \sum_{l,l\neq j} \frac{(-1)^{l-j}[J_{l-j}(|\mathbf{K}|)]^2}{C^{(1)}[e^{ij\theta}]_j - C^{(1)}[e^{il\theta}]_l}, \tag{3.55}$$

up to terms of order ε^3. The last term on the right hand side of Eq. (3.55) is smaller than the other two and we can solve the equation $\lambda_j = 0$ by iteration, thereby finding the approximate solution

$$Q_j = \frac{C^{(1)}[e^{ij\theta}]_j}{J_0(|\mathbf{K}|)} \left[1 + \sum_{l,l\neq j} \frac{(-1)^{l-j} \left[\frac{J_{l-j}(|\mathbf{K}|)}{J_0(|\mathbf{K}|)} \right]^2}{1 - \frac{C^{(1)}[e^{il\theta}]_l}{C^{(1)}[e^{ij\theta}]_j}} \right]. \tag{3.56}$$

For $j = 1$, the largest term in this sum has index $l = 2$ and is of order $|\mathbf{K}|^2$, whereas all other terms are $O(|\mathbf{K}|^4)$ and smaller as $|\mathbf{K}| \to 0$. Ignoring them, Eq. (3.56) yields

$$Q_1 = \frac{C^{(1)}[e^{i\theta}]_1}{J_0(|\mathbf{K}|)} \left[1 - \frac{\left[\frac{J_1(|\mathbf{K}|)}{J_0(|\mathbf{K}|)} \right]^2}{1 - \frac{C^{(1)}[e^{i2\theta}]_2}{C^{(1)}[e^{i\theta}]_1}} \right], \tag{3.57}$$

which is Eq. (3.6). For small $|\mathbf{K}|$, this equation is equivalent to

$$Q_1 \approx \frac{C^{(1)}[e^{i\theta}]_1}{J_0(|\mathbf{K}|) + \frac{[J_1(|\mathbf{K}|)]^2 C^{(1)}[e^{i\theta}]_1}{J_0(|\mathbf{K}|)\,(C^{(1)}[e^{i\theta}]_1 - C^{(1)}[e^{i2\theta}]_2)}}, \tag{3.58}$$

which is Eq. (26) of Ref. [26].

Bifurcation theory using multiple scales. For Hopf and period doubling bifurcations, the scaling of space and time is diffusive and we can use the multiple scales ansatz [25, 46]

$$f(\theta, \mathbf{x}, t; \epsilon) = f_0 + \sum_{j=1}^{\infty} \epsilon^j f^{(j)}(\theta, t, \mathbf{X}, T), \tag{3.59}$$

with $\mathbf{X} = \epsilon\mathbf{x}$, $T = \epsilon^2 t$, $\eta = \eta_c + \epsilon^2\eta_2$, $\rho = \rho_0 + \epsilon^2 r(\mathbf{X})$. Inserting this ansatz into Eq. (4.8), we get a hierarchy of linear equations that have to be solved recursively:

$$\mathcal{L}f^{(1)} \equiv f^{(1)}(\theta, t+1) - C^{(1)}[f^{(1)}](\theta, t) = 0, \tag{3.60}$$

$$\mathcal{L}f^{(2)} = -\mathbf{v}\cdot\nabla_X f^{(1)}(\theta, t+1) + C^{(2)}[f^{(1)}, f^{(1)}](\theta, t), \tag{3.61}$$

$$\mathcal{L}f^{(3)} = -\left(\frac{\partial}{\partial T} + \frac{1}{2}(\mathbf{v}\cdot\nabla_X)^2\right)f^{(1)}(\theta,t+1) - \mathbf{v}\cdot\nabla_X f^{(2)}(\theta,t+1) + \eta_2\frac{\partial C^{(1)}[f^{(1)}(\theta,t)]}{\partial\eta}\Big|_{\eta_c}$$
$$+ r\frac{\partial C^{(1)}[f^{(1)}(\theta,t)]}{\partial\rho}\Big|_{\eta_c} + 2C^{(2)}[f^{(1)},f^{(2)}](\theta,t) + C^{(3)}[f^{(1)},f^{(1)},f^{(1)}](\theta,t), \tag{3.62}$$

etc. In these equations, we have omitted that the $f^{(j)}$ also depend on $\mathbf{X}$ and T. $C^{(2)}[f,g]$ and $C^{(3)}[f,g,h]$ are quadratic and cubic functionals, respectively, resulting from the expansion of $C_E[f_0+\epsilon\tilde{f}]$ in powers of ϵ; cf. Eqs. (3.20)-(3.21). The solution of Eq. (3.60) is Eq. (3.13) with $T=\epsilon^2 t$. Eq. (3.61) has the solution (3.23). Inserting Eqs. (3.13) and (3.24) into Eq. (3.62) and using the solvability condition, we obtain the CGLE (3.29). Its uniform solution with $r=0$ is Eq. (3.30).

Appendix 3.B Coefficients of the amplitude equations

3.B.1 Hopf and period doubling bifurcations

The coefficients δ and μ of the amplitude equation, Eq. (3.26), as obtained from the solvability condition for Eq. (3.21), are:

$$\delta = \frac{1}{8}\left(\frac{1+e^{-i\Omega}C^{(1)}[e^{i2\theta}]_2}{1-e^{-i\Omega}C^{(1)}[e^{i2\theta}]_2} - i\cot\frac{\Omega}{2}\right), \tag{3.63}$$

$$\mu = -\frac{2C^{(2)}[e^{i\theta},e^{i\theta}]_2 C^{(2)}[e^{i2\theta},e^{-i\theta}]_1}{e^{i\Omega}\{e^{i2\Omega}-C^{(1)}[e^{i2\theta}]_2\}} - \frac{C^{(3)}[e^{i\theta},e^{i\theta},e^{-i\theta}]_1}{e^{i\Omega}} \sim \frac{\pi^4 R_0^4/M}{1-e^{-i2\Omega}C^{(1)}[e^{i2\theta}]_2}$$
$$= \frac{\pi^3 R_0^2/\rho_0}{1-e^{-i2\Omega}C^{(1)}[e^{i2\theta}]_2}. \tag{3.64}$$

Here $C^{(1)}[e^{i2\theta}]_2 = Q_2$ is given by Eq. (3.9), and

$$C^{(2)}[e^{i\theta},e^{i\theta}]_2 = 8\pi^2 R_0^2 e^{-M}\sum_{n=2}^{N}\frac{nM^{n-2}}{(n-2)!}K_{cc}^{11}(n)\int_{-\pi}^{\pi}e^{-i2\xi}g(\xi)d\xi$$
$$\sim \pi^2 R_0^2\int_{-\pi}^{\pi}e^{-i2\xi}g(\xi)d\xi, \tag{3.65}$$

$$C^{(2)}[e^{i2\theta},e^{-i\theta}]_1 = 4\pi^2 R_0^2 e^{-M}\sum_{n=2}^{N}\frac{nM^{n-2}}{(n-2)!}K_{c2c}^{1}(n)\int_{-\pi}^{\pi}e^{-i\xi}g(\xi)d\xi$$
$$\sim -\frac{\pi^{5/2}R_0^2}{8\sqrt{M}}\int_{-\pi}^{\pi}e^{-i\xi}g(\xi)d\xi, \tag{3.66}$$

$$C^{(3)}[e^{i\theta},e^{i\theta},e^{-i\theta}]_1=\frac{16}{3}\pi^4R_0^4e^{-M}\sum_{n=3}^{N}\frac{nM^{n-3}}{(n-3)!}K^1_{ccc}(n)\int_{-\pi}^{\pi}e^{-i\xi}g(\xi)d\xi$$
$$\sim-\frac{\pi^{9/2}R_0^4}{2\sqrt{M}}\int_{-\pi}^{\pi}e^{-i\xi}g(\xi)d\xi, \tag{3.67}$$

in the limit as $M\gg 1$ (cf. Section 5 of Ref. [26]). In this limit, $|C^{(1)}[e^{i2\theta}]_2| = |\int_{-\pi}^{\pi}e^{-i2\xi}g(\xi)d\xi|/2\le 1/2$, and therefore

$$\text{Re}\,\delta=\frac{1-|C^{(1)}[e^{i2\theta}]_2|^2}{8|1-e^{-i\Omega}C^{(1)}[e^{i2\theta}]_2|^2}\Longrightarrow \text{Re}\,\delta\ge\frac{3}{32}\frac{1}{(1+|C^{(1)}[e^{i2\theta}]_2|)^2}>\frac{1}{24}>0 \tag{3.68}$$
$$\text{Re}\,\mu=\frac{\pi^4R_0^4}{M}\frac{1-\text{Re}(e^{-i2\Omega}C^{(1)}[e^{i2\theta}]_2)}{|1-e^{-2i\Omega}C^{(1)}[e^{i2\theta}]_2|^2}\Longrightarrow$$
$$\text{Re}\,\mu\ge\frac{\pi^4R_0^4}{M}\frac{1-|C^{(1)}[e^{i2\theta}]_2|}{(1+|C^{(1)}[e^{i2\theta}]_2|)^2}>\frac{2\pi^4R_0^4}{9M}>0. \tag{3.69}$$

3.B.2 Pitchfork bifurcation

In this case, all the coefficients in the amplitude equations are real. The coefficients δ and μ are given by the real part of Eq. (3.63) with $\Omega=0$ and by Eq. (3.64) with $\Omega=0$, respectively. Thus, both δ and μ are positive in the limit as $M\gg 1$. The other coefficients appearing in Eqs. (3.43)-(3.44) are

$$Q_\eta=\frac{\partial}{\partial\eta}\Big(\ln C^{(1)}[e^{i\theta}]_1\Big)\Big|_{\eta_c},\ \delta=\frac{1+C^{(1)}[e^{i2\theta}]_2}{8(1-C^{(1)}[e^{i2\theta}]_2)},$$

$$\gamma_1=\frac{1}{4}C^{(2)}[1,e^{i\theta}]_1-\frac{C^{(2)}[e^{i\theta},e^{i\theta}]_2}{2(1-C^{(1)}[e^{i2\theta}]_2)},$$

$$\gamma_2=-\frac{C^{(2)}[e^{-i\theta},e^{i2\theta}]_1}{1-C^{(1)}[e^{i2\theta}]_2},\ \gamma_3=\frac{C^{(2)}[1,e^{i\theta}]_1}{2}, \tag{3.70}$$

In the limit as $M\gg 1$, we can use Eqs. (3.65)-(3.67) together with

$$C^{(2)}[1,e^{i\theta}]_1=\frac{\pi M}{\rho_0}\frac{\partial C^{(1)}[e^{i\theta}]_1}{\partial M}\sim\frac{\pi C^{(1)}[e^{i\theta}]_1}{2\rho_0}=\frac{\pi}{2\rho_0},$$
$$C^{(3)}[1,1,e^{i\theta}]_1=\frac{2\pi^4R_0^4}{3}\frac{\partial^2C^{(1)}[e^{i\theta}]_1}{\partial M^2}\sim-\frac{\pi^2C^{(1)}[e^{i\theta}]_1}{6\rho_0^2}=-\frac{\pi^2}{6\rho_0^2}, \tag{3.71}$$

to calculate these coefficients. Recall that $C^{(1)}[e^{i\theta}]_1 = 1$ at η_c. We obtain:

$$\begin{aligned}
&\frac{C^{(3)}[1,1,e^{i\theta}]_1 - 2(C^{(2)}[1,e^{i\theta}]_1)^2}{4\pi^2} \sim -\frac{1}{6\rho_0^2},\\
&Q_\eta \sim -\frac{\sqrt{\pi M}}{2\eta_c}\left(\frac{2}{\sqrt{\pi M}} - \cos\frac{\eta_c}{2}\right),\\
&\gamma_1 \sim \pi^2 R_0^2\left(1 + \frac{1}{8M} - \gamma_0\right), \quad \gamma_2 \sim \frac{\gamma_0 \pi^2 R_0^2}{4M},\\
&\gamma_0 = \frac{1}{1 - \frac{1}{\sqrt{\pi M}}\cos\frac{\eta_c}{2} - \frac{p}{2}\left(1 + \cos\frac{\eta_c}{2}\right)},\\
&\gamma_3 \sim \frac{\pi^2 R_0^2}{2M}, \quad \delta \sim \frac{2\gamma_0 - 1}{8}, \quad \mu \sim \frac{\pi^4 R_0^4 \gamma_0}{M}.
\end{aligned} \tag{3.72}$$

To find γ_0 in the limit as $M \gg 1$, we have used its definition, $\gamma_0^{-1} = 1 - C^{(1)}[e^{i2\theta}]_2$, and the linear stability condition, $Q_1 = 1$ (for $\xi_0 = \pi$):

$$-p + 2\frac{1-p}{\eta_c}\sin\frac{\eta_c}{2} = \frac{2}{\sqrt{\pi M}} \Longrightarrow \frac{1-p}{\eta_c}\sin\frac{\eta_c}{2} = \frac{p}{2} + \frac{1}{\sqrt{\pi M}}, \tag{3.73}$$

which, inserted in Eq. (3.9) for $C^{(1)}[e^{i2\theta}]_2$, yields

$$\begin{aligned}
C^{(1)}[e^{i2\theta}]_2 \;=\; & \frac{1}{2}\left(p + \frac{1-p}{\eta_c}\sin\eta_c\right) = \frac{p}{2} + \frac{1-p}{\eta_c}\sin\frac{\eta_c}{2}\cos\frac{\eta_c}{2} = \frac{p}{2}\left(1 + \cos\frac{\eta_c}{2}\right)\\
& + \frac{1}{\sqrt{\pi M}}\cos\frac{\eta_c}{2}.
\end{aligned} \tag{3.74}$$

References

[1] S. Ramaswamy, The Mechanics and Statistics of Active Matter, Annu. Rev. Condens. Matter Phys. **1**, 323 (2010).

[2] T. Vicsek and A. Zafeiris, Collective motion, Phys. Rep. **517**, 71 (2012).

[3] M.C. Marchetti, J.F. Joanny, S. Ramaswamy, T.B. Liverpool, J. Prost, M. Rao and R.A. Simha, Hydrodynamics of soft active matter, Rev. Mod. Phys. **85**, 1143 (2013).

[4] D. Bi, X. Yang, M.C. Marchetti, and M.L. Manning, Motility-Driven Glass and Jamming Transitions in Biological Tissues, Phys. Rev. X **6**, 021011 (2016).

[5] V. Hakim and P. Silberzan, Collective cell migration: a physics perspective, Rep. Prog. Phys. **80**, 076601 (2017).

[6] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, Novel type of phase transition in a system of self-driven particles, Phys. Rev. Lett. **75**, 1226 (1995).

[7] Y.-L. Chou, R. Wolfe, and T. Ihle, Kinetic theory for systems of self-propelled particles with metric-free interactions, Phys. Rev. E **86**, 021120 (2012).

[8] T. Ihle, Large density expansion of a hydrodynamic theory for self-propelled particles, Eur. Phys. J. Spec. Top. **224**, 1303 (2015).

[9] T. Ihle, Chapman-Enskog expansion for the Vicsek model of self-propelled particles, J. Stat. Mech.: Theor. Expts. **(2016)**, 083205.

[10] M. C. Miguel, J. T. Parley, and R. Pastor-Satorras, Effects of Heterogeneous Social Interactions on Flocking Dynamics, Phys. Rev. Lett. **120**, 068303 (2018).

[11] T. Ihle, Kinetic theory of flocking: Derivation of hydrodynamic equations, Phys. Rev. E **83**, 030901(R) (2011).

[12] W. R. DiLuzio, L. Turner, M. Mayer, P. Garstecki, D. B. Weibel, H. C. Berg, and G. M. Whitesides, *Escherichia coli* swim on the right-hand side, Nature (London) **435**, 1271 (2005).

[13] E. Lauga, W. R. DiLuzio, G. M. Whitesides, and H. A. Stone, Swimming in Circles: Motion of Bacteria near Solid Boundaries, Biophys. J. **90**, 400 (2006).

[14] D. Takagi, J. Palacci, A. B. Braunschweig, M. J. Shelley, and J. Zhang, Hydrodynamic capture of microswimmers into sphere-bound orbits, Soft Matter **10**, 1784 (2014).

[15] S. J. Ebbens, Active colloids: Progress and challenges towards realising autonomous applications, Current Opinion in Colloid & Interface Science **21**, 14 (2016).

[16] B. Liebchen and D. Levis, Collective Behavior of Chiral Active Matter: Pattern Formation and Enhanced Flocking, Phys. Rev. Lett. **119**, 058002 (2017).

[17] S. Bazazi, J. Buhl, J.J. Hale, M.L. Anstey, G.A. Sword, S.J. Simpson and I.D. Couzin, Collective motion and cannibalism in locust migratory bands, Current biology, **18**, 735 (2008).

[18] C.A. Yates, R. Erban, C. Escudero, I.D. Couzin, J. Buhl, I.G. Kevrekidis, P.K. Maini and D.J. Sumpter, Inherent noise can facilitate coherence in collective swarm motion, Proceedings of the Proc. Natl. Acad. Sci. USA **106**, 5464 (2009).

[19] Y. Kuramoto, Self-entrainment of a population of coupled non- linear oscillators, in *International Symposium on Mathematical Problems in Theoretical Physics*, edited by H. Araki, Lecture Notes in Physics, Vol. 39 (Springer, Berlin, 1975), pp. 420–422.

[20] J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, and R. Spigler, The Kuramoto model: a simple paradigm for synchronization phenomena, Rev. Mod. Phys. **77**, 137 (2005).

[21] H. Hong and S. H. Strogatz, Kuramoto Model of Coupled Oscillators with Positive and Negative Coupling Parameters: An Example of Conformist and Contrarian Oscillators, Phys. Rev. Lett. **106**, 054102 (2011).

[22] R. Hegselmann and U. Krause, Opinion dynamics and bounded confidence, models, analysis and simulation, J. Artif. Soc. Social Simul. **5**, 2 (2002).

[23] T. Kurahashi-Nakamura, M. Mäs, and J. Lorenz, Robust Clustering in Generalized Bounded Confidence Models, J. Artif. Soc. Social Simul. **19**, 7 (2016).

[24] M. Haghani and M. Sarvi, Following the crowd or avoiding it? Empirical investigation of imitative behaviour in emergency escape of human crowds, Animal Behaviour **124**, 47 (2017).

[25] L. L. Bonilla and S. W. Teitsworth, *Nonlinear wave methods for charge transport* (Wiley-VCH, 2010).

[26] L. L. Bonilla and C. Trenado, Crossover between parabolic and hyperbolic scaling, oscillatory modes and resonances near flocking, Phys. Rev. E **98**, 062603 (2018).

[27] Y. Kuramoto and T. Tsuzuki, Persistent Propagation of Concentration Waves in Dissipative Media Far from Thermal Equilibrium, Prog. Theor. Phys. **56**, 356 (1976).

[28] G.B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974).

[29] H. Chaté and P. Manneville, Phase Diagram of the Two-Dimensional Complex Ginzburg-Landau Equation, Physica A **224**, 348 (1996).

[30] I. S. Aranson and L. Kramers, The world of the complex Ginzburg-Landau equation, Rev. Mod. Phys. **74**, 99 (2002).

[31] P. Manneville, *Instabilities, chaos and turbulence, An Introduction to Nonlinear Dynamics and Complex Systems* (Imperial College P., 2004).

[32] J. C. Neu, Vortices in complex scalar fields, Phys. D (Amsterdam) **43**, 385 (1990).

[33] M. C. Cross and P. C. Hohenberg, Pattern formation outside of equilibrium, Rev. Mod. Phys. **65**, 851 (1993).

[34] K. Huang, *Statistical Mechanics, 2nd ed* (Wiley, New York 1987).

[35] G. Grégoire and H. Chaté, Onset of Collective and Cohesive Motion, Phys. Rev. Lett. **92**, 025702 (2004).

[36] H. Chaté, F. Ginelli, G. Grégoire, and F. Raynaud, Collective motion of self-propelled particles interacting without cohesion, Phys. Rev. E **77**, 046113 (2008).

[37] A.P. Solon, H. Chaté, and J. Tailleur, From Phase to Microphase Separation in Flocking Models: The Essential Role of Nonequilibrium Fluctuations, Phys. Rev. Lett. **114**, 068101 (2015).

[38] C. Huepe and M. Aldana, New tools for characterizing swarming systems: A comparison of minimal models, Physica A **387**, 2809 (2008).

[39] J.-B. Caussin, A. Solon, A. Peshkov, H. Chaté, T. Dauxois, J. Tailleur, V. Vitelli, and D. Bartolo, Emergent Spatial Structures in Flocking Models: A Dynamical System Insight, Phys. Rev. Lett. **112**, 148102 (2014).

[40] A. Attanasi, A. Cavagna, L. Del Castello, I. Giardina, S. Melillo, L. Parisi, O. Pohl, B. Rossaro, E. Shen, E. Silvestri, and M. Viale, Finite-Size Scaling as a Way to Probe Near-Criticality in Natural Swarms, Phys. Rev. Lett. **113**, 238102 (2014).

[41] G. Baglietto and E.V. Albano, Nature of the order-disorder transition in the Vicsek model for the collective motion of self-propelled particles, Phys. Rev. E **80**, 050103(R) (2009).

[42] H. Chaté, F. Ginelli, and R. Montagne, Simple Model for Active Nematics: Quasi-Long-Range Order and Giant Fluctuations, Phys. Rev. Lett. **96**, 180602 (2006).

[43] A. M. Menzel, Collective motion of binary self-propelled particle mixtures, Phys. Rev. E **85**, 021912 (2012).

[44] O. Chepizhko, E. G. Altmann, and F. Peruani, Optimal Noise Maximizes Collective Motion in Heterogeneous Media, Phys. Rev. Lett. **110**, 238101 (2013).

[45] N. M. Temme, *Special Functions: An introduction to the classical functions of mathematical physics* (Wiley, 1996).

[46] J. C. Neu, *Singular Perturbations in the Physical Sciences*, Graduate Studies in Mathematics, Vol. 147 (American Math- ematical Society, Providence, RI, 2015).

Chapter 4

Tracking collective cell motion

Contents

4.1 Introduction

Confluent motion of epithelial cell monolayers [1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24] is crucial in many biological processes, such as morphogenesis [3], tissue repair [6, 17, 18], development [4], and tumor invasion and metastasis [1, 2, 3, 24]. It serves as a relatively simple paradigm for collective motion of cells that retain their cell-cell junctions as they move on a two dimensional (2D) substrate. Collective cell migration also poses challenging questions in soft and active matter physics, as it may exhibit fluid, solid or glass behavior with interesting flocking and jamming/unjamming transitions [12, 19, 25, 26, 27, 28, 29, 30, 31]. Interesting dynamics occurs as an epithelial cell aggregate advances through an empty space, as in wound healing, or it collides and encroaches a different tissue, as in cancer invasion. Moreover, advancing cellular fronts may display wave phenomena [15, 32], grow fingers [16, 33, 34], or breakdown and interpenetration against an oppositely moving front [20, 23].

In this Chapter, we explain the combination of particle dynamics [16] with the active vertex model (AVM) [35] to provide a cellular dynamics perspective (i) on monolayers spreading over an empty space [11, 16, 33, 34] and (ii) monolayers colliding in antagonistic migration assays (AMA) [20, 23]. The resulting model describes the collective migration dynamics of a large number of cells and implements exchanges of neighboring cells automatically (T1 transitions) [35]. It incorporates internal dissipation of cells through a friction parameter, a Vicsek-like velocity alignment of neighboring cells [26, 37, 38], noise and and active forces that may include cell polarity. Particles are not self-propelled and have the ability to start the migration or stop in absence of noise.

In AMA with Human Embryonic Kidney (HEC) cell assemblies, precancerous Ras modified cells displace normal cells [23]. These latter experiments have been interpreted using continuum mechanics in a simple biophysical model through phenomenological couplings [34], without recourse to biochemical signaling mechanisms but without clear relations to cellular processes. In this Chapter, we consider wild type (wt) HEC cells to be solid-like whereas invading Ras cells are fluid-like and push the former backward. As time elapses, there are cell exchanges and islands of one cell type form inside the tissue of the other cells, which characterizes a flocking liquid state [28, 30, 36]. However, in AMA with Madin-Darby Canine Kidney (MDCK) cells, the roles are inverted: Ras cells are solid-like and wt cells are fluid-like. The precise form of the separating interface among monolayers of different cell type depends on cell parameters governing segregation vs aggregation of these cells. We characterize it by topological data analysis (TDA). A measure of cellular diversity in the junction tensions produces islands of one type of cells inside the monolayer of the other cells, which is reflected in TDA of simulations and experiments. Cell cohesion given by the AVM, the cell alignment rule and the active noise force produce fingers in interfaces during assays of cell invasions of empty spaces rendering unnecessary to assume a different phenotype for lead cells [16].

Recent experiments have connected metastasis in colorectal cancer to wound healing and tumor invasion of tissue using appropriate molecular markers [24]. Thus, our description of spreading of cellular tissue and antagonistic migration assays using our modified active vertex model might be relevant for metastatic cancer. In particular, we shall show the role of cellular junction tensions in cell invasion, agglomeration and segregation. Understanding precise biochemical mechanisms influencing cell-cell contact and confluent cellular tissue may help develop therapies for metastatic cancers.

Chapter 4 is organized as follows. In Section 4.2 we are going to describe the models we simulate and discuss the numerical values we use to reproduce in (i) an aggregate spreading to an empty space and (ii) the collision of two different cellular monolayers in antagonistic migration assays. The results of the numerical simulations are detailed in Section 4.3 and in Section 4.4 we study the formation of islands in the AMA using TDA.

4.2 Model

Particles based-models have been useful to understand and capture specific aspects of migrating cells [16]. However, to model and reproduce properly the migration of an epithelial monolayer, it is more appropriate to use the Vertex Model. It started with the study of physics of foams in the 1970's and it has been used in a wide variety of dynamic systems.

In our numerical simulations, we use an active vertex model AVM [35] and simulate it by adapting the SAMoS software [39]. The AVM combines the Vertex Model for confluent epithelial tissues [25, 40] with active matter dynamics [35]. Let us describe first the Vertex model, then the AVM and our modification of its dynamics.

4.2.1 The Vertex Model

The Vertex model assumes that all cells in the epithelium are roughly the same height and thus the entire system can be well approximated as a two-dimensional convex polygons with no holes and overlaps that cover the surface. The conformation of the tissue in the Vertex model is computed as a configuration that simultaneously optimizes area and perimeter of all cells. In addition, two neighboring cells share a single edge which is a simplification of the reality. Three junction lines typically meet at a vertex, although vertices with a higher number of contacts are also possible. The model tissue is therefore a mesh consisting of polygons (i.e., cells), edges (i.e., cell junctions), and vertices (i.e., meeting points of three or more cells), see Fig. 4.1. Each configuration of the mesh has the following associated energy

Figure 4.1: Each cell in the Vertex model is represented as a polygon with n sides and it is going to be defined by three parameters, the area A_i, the perimeter P_i and the junction connection between the vertices μ and ν.

$$E_{\mathrm{VM}} = \sum_{i=1}^{N} \left[\frac{K_i}{2} (A_i - A_i^0)^2 + \frac{\Gamma_i}{2} P_i^2 \right] + \sum_{\langle \mu,\nu \rangle} \Lambda_{\mu\nu}\, l_{\mu\nu}. \tag{4.1}$$

Here N is the total number of cells, A_i is the area of the cell i, A_i^0 is its reference area, and K_i is the area modulus, i.e., a constant with units of energy per area squared measuring how hard it is to change the area of the cell. P_i is the cell perimeter and Γ_i (with units of energy per length squared) is the perimeter modulus that determines how hard it is to change perimeter P_i. $l_{\mu\nu}$ is the length of the junction between vertices μ and ν, and $\Lambda_{\mu\nu}$ is the tension of that junction (with units of energy per length). The sum in the last term is over all pairs of vertices that share a junction. Note that the model allows for different cells to have different area and perimeter moduli as well as reference areas, allowing for modeling of tissues containing different cell types. The cell area and perimeter can be written in terms of vertex coordinates. Thus, vertex positions together with their connectivities uniquely determine the energy of the epithelial sheet. The main assumption of the Vertex model is that the tissue will always be in a configuration which minimizes the total energy in Eq. (4.1). To implement the Vertex model, we determine the positions of vertices that minimize E_{VM} for a given set of parameters K_i, Γ_i, and $\Lambda_{\mu\nu}$. Cell rearrangements are modeled by introducing moves that change appropriately the connectivity among cells.

While the moduli K_i and Γ_i are positive, $\Lambda_{\mu\nu} < 0$. When the cell i shares junctions only with others of the same type, $\sum_{\langle \mu,\nu \rangle} \Lambda_{\mu\nu} l_{\mu\nu} = \Lambda_{\mu\nu} \sum_{\langle \mu,\nu \rangle} l_{\mu\nu} = \Lambda_{\mu\nu} P_i$, and this term can be put together with the perimeter term, thereby yielding $\frac{\Gamma_i}{2}(P_i - P_i^0)^2$ plus an unimportant constant, provided $P_i^0 = -\Lambda_{\mu\nu}/\Gamma_i > 0$. Thus the junction tension $\Lambda_{\mu\nu}$ determines the target perimeter of a type of cell. Let us assume that there are two cell types, 1 and 2, with moduli K_j, Γ_j, $j = 1, 2$, Λ_{11}, Λ_{22}, Λ_{12}, and target areas and perimeters A_j^0, P_j^0, $j = 1, 2$, respectively.

We can complete squares and drop additive constants, thereby obtaining

$$E_{\mathrm{VM}} = \sum_{j=1}^{2} \sum_{i_j=1}^{N_j} \left[\frac{K_j}{2} (A_{i_j} - A_j^0)^2 + \frac{\Gamma_j}{2} (P_{i_j} - P_j^0)^2 \right] + (2\Lambda_{12} - \Lambda_{11} - \Lambda_{22}) \sum_{\langle \mu,\nu \rangle} l_{\mu\nu}, \quad (4.2)$$

in which $N_1 + N_2 = N$.

Clearly, $\Lambda_{12} < (\Lambda_{11} + \Lambda_{22})/2$ implies that energy is minimized when the number of junctions between both types of cells increases. Cells of different types therefore tend to mix. Conversely, when $\Lambda_{12} > (\Lambda_{11} + \Lambda_{22})/2$ cells of different type segregate, as suppressing junctions between cells of different type minimizes energy. There is also a competition between the two first terms in Eq. (4.2) to minimize energy. Assume $\Lambda_{12} = (\Lambda_{11} + \Lambda_{22})/2$ and therefore the last term in Eq. (4.2) vanishes. The shape index $p_j^0 = P_j^0/\sqrt{A_j^0} = |\Lambda_{jj}|/(\Gamma_j \sqrt{A_j^0})$ controls the ratio of the type j cell perimeter to its area. For the Vertex model, the value $p_*^0 = 3.812$ (which corresponds to pentagons) separates solidlike and fluidlike behavior of the tissue [27]. For $p^0 < p_*^0$, cortical tension is prevalent over cell-cell adhesion, cells do not exchange neighbors and the monolayer is solidlike. For $p^0 > p_*^0$, cell-cell adhesion dominates, neighbor exchanges occur, and the cellular tissue behaves like a fluid [27].

4.2.2 Active Vertex Model

To introduce dynamics in the Vertex model, we have to go from polygon vertices $\mathbf{r}_\mu$ to polygon centers that represent cells, $\mathbf{r}_i$, considering these centers as particles and introducing dynamics for them [35]. The core assumption of the AVM is that the tissue configurations that optimize the energy in Eq. (4.1) correspond to the Voronoi tessellations of the plane with polygons as cells and cell centers acting as Voronoi seeds. Given a Voronoi tessellation, we consider its dual Delaunay triangulation, comprising Voronoi seeds and the edges joining them (triangles), which have the property that no seed is inside the circumcircle of any triangle; see Fig. 4.2. From a Voronoi tesselation it is straightforward to obtain the dual Delaunay triangulation and vice versa. However, working with Delaunay triangulations has an advantage: they retain their nature when triangle vertices move by flipping edges conveniently [35], whereas Voronoi tessellations do not. The latter have to be reset after motion of polygon vertices.

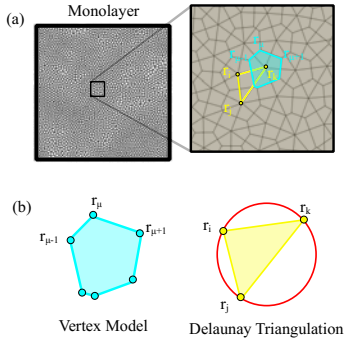

Figure 4.2: Voronoi tessellation and Delaunay triangulation. (a) Zoom of a monolayer showing the vertices of polygons ($\mathbf{r}_\mu$) in the Voronoi tessellation and the centers ($\mathbf{r}_i$) of polygons that are vertices of Delaunay triangles. (b) Voronoi tessellation and Delaunay triangulation.

In the AVM, the area A_i in Eq. (4.1) of the cell i is the area of the associated Voronoi polygon, Ω_i, given by the following discrete version of Green's

formula:

$$A_i = \frac{1}{2} \sum_{\mu \in \Omega_i} (\mathbf{r}_\mu \times \mathbf{r}_{\mu+1}) \cdot \mathbf{N}_i, \tag{4.3}$$

where $\mathbf{r}_\mu$ is the position of vertex μ, and $\mathbf{N}_i$ is a unit vector perpendicular to the surface of the polygon. For the 2D tissue $\mathbf{N}_i$ is directed along the z axis and therefore does not depend on the position of the vertices. The sum in Eq. (4.3) is over all vertices of the Voronoi cell and we close the loop with $\mu + 1 = 1$ when μ equals the total number of vertices in the cell, N_{Ω_i}. The cell perimeter is

$$P_i = \frac{1}{2} \sum_{\mu \in \Omega_i} |\mathbf{r}_\mu - \mathbf{r}_{\mu+1}|. \tag{4.4}$$

The relation between the vertices $\mathbf{r}_\mu$ of the Voronoi polygons (i.e., cells) and the vertices $\mathbf{r}_i$ of the Delaunay triangles (seeds of the Voronoi polygons, i.e., cell centers) is

$$\mathbf{r}_\mu = \frac{\lambda_1 \mathbf{r}_i + \lambda_2 \mathbf{r}_j + \lambda_3 \mathbf{r}_k}{\lambda_1 + \lambda_2 + \lambda_3}. \tag{4.5}$$

Here $\mathbf{r}_i$, $\mathbf{r}_j$ and $\mathbf{r}_k$ are position vectors of the corners of the triangle and λ_i, $i = 1, 2, 3$, are the barycentric coordinates; cf. Fig. 4.2, and Ref. [35] for details. The forces $\mathbf{F}_i = -\nabla_{\mathbf{r}_i} [E_{\mathrm{VM}} + V_{\mathrm{soft}}(|\mathbf{r}_i - \mathbf{r}_j|)]$ are [35]

$$\begin{aligned}
\mathbf{F}_i = &- \sum_{k=1}^{N} \frac{K_k}{2} \left(A_k - A_k^0 \right) \sum_{\nu \in \Omega_k} \left[\mathbf{r}_{\nu+1,\nu-1} \times \mathbf{N}_k \right]^T \left[\frac{\partial \mathbf{r}_\nu}{\partial \mathbf{r}_i} \right] \\
&- \sum_{k=1}^{N} \Gamma_k P_k \sum_{\nu \in \Omega_k} \left(\hat{\mathbf{r}}_{\nu,\nu-1} - \hat{\mathbf{r}}_{\nu+1,\nu} \right)^T \left[\frac{\partial \mathbf{r}_\nu}{\partial \mathbf{r}_i} \right] \\
&- \sum_{k=1}^{N} \sum_{\nu \in \Omega_k} \left[\Lambda_{\nu-1,\nu} \hat{\mathbf{r}}_{\nu,\nu-1} - \Lambda_{\nu,\nu+1} \hat{\mathbf{r}}_{\nu+1,\nu} \right]^T \left[\frac{\partial \mathbf{r}_\nu}{\partial \mathbf{r}_i} \right] \\
&+ k \sum_{\langle j,i \rangle} (2a - |\mathbf{r}_i - \mathbf{r}_j|) \frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|} \Theta(2a - |\mathbf{r}_i - \mathbf{r}_j|).
\end{aligned} \tag{4.6}$$

Here $\left[\frac{\partial \mathbf{r}_\nu}{\partial \mathbf{r}_i} \right]$ is the 3×3 Jacobian matrix connecting coordinates of cell centres with coordinates of the dual Voronoi tessellation, and the non-commutative row-matrix product $[\cdot]^T [\cdot]$ is a 3×1 column vector. $\Theta(x) = 1$ if $x > 0$, else $\Theta(x) = 0$, is the Heaviside unit step function. We have included a range repulsive force of short range a that avoids self intersections of the triangulation for very obtuse triangles [35]. In the AVM, the usual dynamics for the cell centers is a gradient flow of the energy in Eq. (4.1), that is overdamped dynamics with forces

$\mathbf{F}_i$ given by Eq. (4.6), plus active forces $f_a\mathbf{n}_i$, and stochastic forces $\boldsymbol{\nu}_i$ [35]

$$\gamma\,\dot{\mathbf{r}}_i = f_a\mathbf{n}_i + \mathbf{F}_i + \boldsymbol{\nu}_i, \quad \gamma_r\dot{\theta}_i = \boldsymbol{\tau}_i\cdot\mathbf{N}_i + \nu_i^r(t), \tag{4.7}$$

where $\dot{\mathbf{r}}_i = d\mathbf{r}_i/dt$, $\boldsymbol{\tau}_i$ is the torque acting on the polarity $\mathbf{n}_i = (\cos\theta_i, \sin\theta_i)$, $\mathbf{N}_i$ is the local normal to the cell surface (the unit length vector in the z-direction), γ_r is the orientational friction, and $\nu_i^r(t)$ is a zero mean Gaussian white noise responsible for orientational randomness, such that $\langle\nu_i^r(t)\nu_j^r(t')\rangle = 2D_r\delta_{ij}\delta(t-t')$. Terms aligning cell velocity or shape to polarity or terms aligning the polarity of different cells can be included in the energy of Eq. (4.1) [35]. A particularly simple dynamics follows from $f_a = v_0$ (constant active force), $\boldsymbol{\nu}_i = \boldsymbol{\tau}_i = \mathbf{0}$ [27]. The AVM describes naturally cell motion and accounts for patterns of the confluent tissue observed on multiple scales, from cell sizes to much larger distances. Furthermore, cell contacts are generated dynamically from the positions of cell centers.

4.2.3 Dynamics including velocity alignment and inertia.

In this work, we shall modify the AVM dynamics. Instead of Eq. (4.7), we shall use the particle dynamics of Ref. [16] but with different forces between particles. As discussed in Ref.[41] , trajectories of motile cells can be explained by assuming that their acceleration is a certain functional of velocity. Despite the mass of the cell being so small that inertia is negligible compared with typical forces exerted on the cell, the formula for acceleration resembles Newton's second law [41]. In this formula, a linear damping term represents dissipative processes coming from friction with substrate, with other cells, or rupture of adhesion bonds. Active memory terms, which are linear in the velocity, may propel single cells and account for the observed non-monotonic velocity autocorrelation [41]. When considering cellular tissue, Sepúlveda et al model cells as actively motile particles and replace the memory terms by Vicsek-like alignment "forces" [37], and inter-particle and random "forces" [16]. Thus, the acceleration in these models is a consequence of the collective motion of cells and the interaction with the environment and it does not follow from Newton's second law with a mass given by that of a single cell. However, we will continue denoting by forces (per unit mass) the terms comprising the acceleration [16]. In contrast to Eq. (4.7) cells in Ref. [16] are not self-propelled, so that they can stop their motion and start moving again if

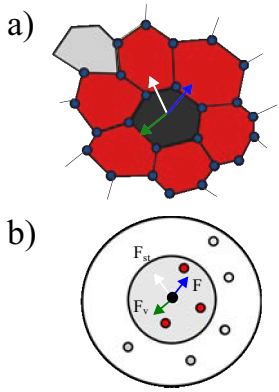

Figure 4.3: Cells represented in (a) and particles in (b) colored in red are considered neighbors of cell/particle in black. Three forces act on the cell/particle, the stochastic, the alignment force and the force generated by a potential defined in [16] for particles or the vertex forces for cells.

there are missing cells in their neighborhood and the active force is zero,

$$\dot{\mathbf{r}}_i = \mathbf{v}_i, \quad \dot{\mathbf{v}}_i = -\alpha \mathbf{v}_i + \sum_{\langle j,i \rangle} \left[\frac{\beta}{n_i} (\mathbf{v}_j - \mathbf{v}_i) + \mathbf{f}_{ij} \right] + \boldsymbol{\varphi}_i + \sigma_0 \boldsymbol{\eta}_i(t), \quad \tau \dot{\boldsymbol{\eta}}_i = -\boldsymbol{\eta}_i + \boldsymbol{\xi}_i(t). \tag{4.8}$$

Here, the sum is over the nearest neighbors of the vertex i of the Delaunay triangulation, n_i is the number of these neighbors, the friction coefficient α comes from internal cell friction or adhesion to the substrate or other cells. The term containing the coefficient β tries to synchronize the velocity of the nearest neighbor cells that of the ith cell and it is similar to the Vicsek model [26, 37, 38, 42]. $\mathbf{f}_{ij}$ is the force per unit mass exerted by cell j on cell i (that has mass m_i). In our simulations we use $\sum_{\langle j,i \rangle} \mathbf{f}_{ij} = \mathbf{F}_i / m_i$, where $\mathbf{F}_i$ is given by Eq. (4.6), and not by an interparticle potential as in Ref. [16] m_i is a reference mass, for example $m_i = \gamma \gamma_r^2 / D_r$. The active forces are $\boldsymbol{\varphi}_i + \sigma_0 \boldsymbol{\eta}_i(t)$. In Ref. [16], $\boldsymbol{\varphi}_i = 0$ and $\boldsymbol{\eta}_i(t)$ is a zero mean Ornstein-Uhlenbeck noise, representing a stochastic force with nonzero correlation time τ. $\boldsymbol{\xi}_i(t)$ is a zero-mean delta-correlated Gaussian white noise. For spreading tests, we have used the numerical values of the parameters indicated in Table 4.1. In our combination of the AVM with particle dynamics, our particles are at the vertices of the Delaunay triangulation (cell centers) and they are no longer point particles as in Ref. [16]. For point particles, the cohesive structure provided by the Delaunay triangulation is absent and has to be achieved by convenient forces $\mathbf{f}_{ij}$, which are repulsive at short distances, attractive at long distances and derive from a convenient potential [16].

α	β	τ	σ_0	K	Γ	Λ
h^{-1}	h^{-1}	h	$\frac{\mu\mathrm{m}}{\mathrm{h}^{3/2}}$	-	-	-
0.534	41.36	0.56	95	1	0.1	-1

Table 4.1: Parameters for our model corresponding to the experiments with MDCK cells in Ref. [16].

4.2.4 Boundaries and Initial conditions.

To handle the limits of the sheet layer, cells at the boundary between a cellular monolayer and the empty space, or between tissues, are special, see Fig. 4.4 to illustrate. They may form actin cables, thereby having a line tension and a bending stiffness [35].

$$E_{lt} = \frac{1}{2} \sum_{\langle i,j \rangle} \lambda_{ij} (l_{ij} - l_0)^2 \tag{4.9}$$

$$E_{bend} = \frac{1}{2} \sum_i \zeta_i (\theta_i - \pi)^2, \quad \cos(\theta_i) = \frac{\vec{r}_{ji} \cdot \vec{r}_{ki}}{|\vec{r}_{ji}||\vec{r}_{ki}|} \tag{4.10}$$

Here the modulus λ_{ij} is the line tension of the edge connecting vertices i and j, $l_{ij} = |r_{ij}| (r_{ij} = r_i - r_j)$ is the edge length (of preferred magnitude l0), ζ_i is the bending stiffness of angle θ_i at the boundary particle i, and r_j and r_k are the positions of boundary particles to the left and

to the right of particle i. The line tension energy of Eq. (4.9) tries to keep boundary edges at a length l0 whereas the bending energy of Eq. (4.10) tries to keep the boundary line flat. The sums in these formulas are over boundary particles only and we assume that each boundary cell has exactly two boundary neighbors [35].

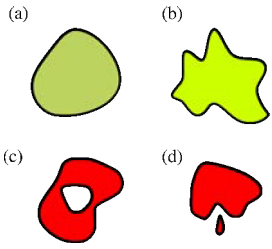

Figure 4.4: (a) and (b) are allowed boundaries in te AVM, however it is not possible to have a hole inside a bulk (c) or split the bulk into to different domains (c).

A random configuration of the particles comprising a confluent cell monolayer is usually different from those configurations observed in experiments. Thus, we have to carry out an initialization stage until the particle configuration is compatible with their observed velocity distributions. For spread tests, we proceed as follows. We set a square box of size 1 mm^2 area, see Fig. 4.6, $N \approx 4000$ particles (comparable to the number of cells in the experiments), the packing ratio and the particle mean velocity. Then, we numerically solve Eqs. (4.1) and (4.8) with forces $\boldsymbol{\varphi}_i = \mathbf{0}$ and $\sum_{\langle j,i\rangle} \mathbf{f}_{ij} = \mathbf{F}_i/m_i$, $\mathbf{F}_i$ given by Eq. (4.6), until the velocity probability density functions (PDFs) of the experiments are fitted. The parameters adjusted to the experimental data at early time (30 min after stencil removal) are listed in Table 4.1. We stop the initialization stage when the distribution of mean distances between particles is close to the initial distribution as observed in experiments and displayed in Fig. 4.5. From this simulation, we obtain the particle positions $\mathbf{r}_i$ and initial random directions for the particle velocities. As we can see in Fig. 4.6, the velocity field obtained from the simulations, Fig. 4.6(b), is very similar to that measured by PIV analysis [16], Fig. 4.5(a).

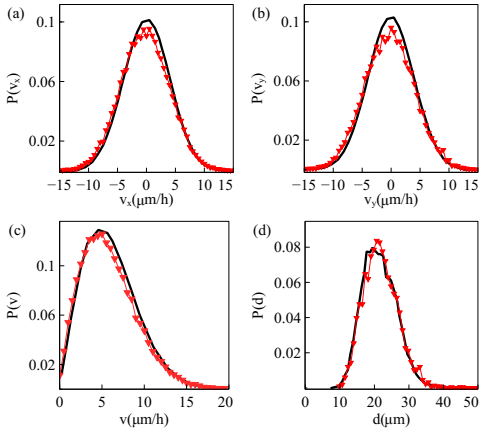

Figure 4.5: Probability distribution function (PDF) for particle velocities: (a) v_x, (b) v_y, (c) $v = |\mathbf{v}|$; and (d) mean distance d between neighboring particles; after the initialization procedure (red triangles) as compared to the experimentally observed PDF (black line) [16]. Parameter values are those in Table 4.1.

For AMA, we choose a random configuration having the same number of wt and Ras cells separated by a vertical straight line and we set a known velocity distribution from experiments [23]. This represents the situation of the two monolayers when they first make contact. See details in the next Section.

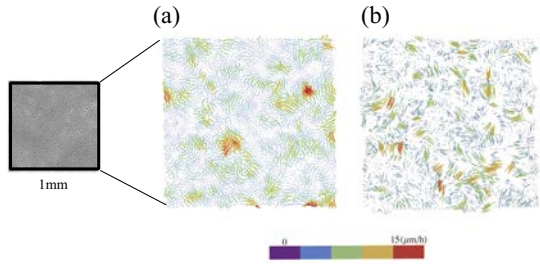

Figure 4.6: Velocity field obtained from (a) experiments [16], (b) simulations after the initialization procedure. Parameter values are those in Table 4.1.

4.3 Results of numerical simulations

In this Chapter, we present the results obtained from the simulations of two different tissue configurations, (a) a cellular monolayer spreads over an empty space, and (b) two monolayers comprising wild type and modified cells collide. In each case, the simulations are compared to relevant experimental observations.

4.3.1 Collective movements in a migrating epithelium

Inspired by wound healing phenomena and experiments on tissue scratching (explained briefly in Subsection 1.2.1) and in order to clarify the mechanisms involved in a free edge of an epithelium, we study the movement of an epithelium which encroaches on a virgin substrate with the experimental protocol consisting of micro-fabricated stencils whose removal increases the motility of the epithelium. This technique has several advantages, dimensions are well controlled, the geometry is almost perfect and cells do not suffer an injury. During the motion of an epithelial strip after removing the stencil, we can distinct three different phases,

First phase of motion. Minutes after the stencil removal, the first rows of the cells located at the edge of the monolayer start their motion. Their direction of motion is perpendicular to the edge of the band and their speed is around $25\mu m h^{-1}$.

Second phase of motion. After two hours, the PIV recorded from the experiments reveals the complex movements that can appear inside the bulk of the tissue, cf. Fig. 4.7. Cells do not

move independently, their velocities are correlated and the presence of cellular flows shows the existence of motion inside the monolayer [11]. The average speed over a whole band of cells is equal to $8\mu mh^{-1} \pm 1\mu mh^{-1}$. However, there are some areas where their speed can reach bigger values. These areas are not only found at the edges of the monolayer but also in the middle of the bulk.

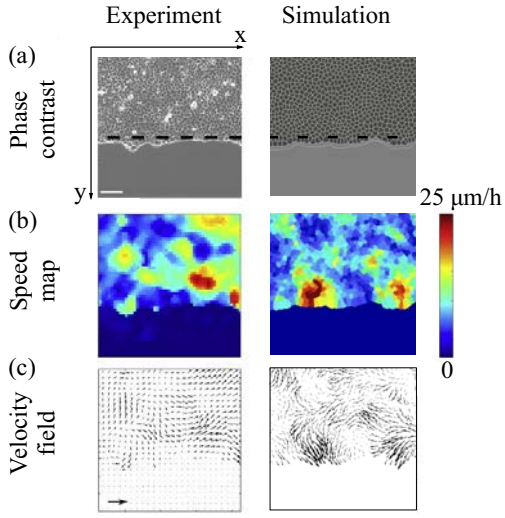

Figure 4.7: Cell velocity field after 2 h of stencil removal in an invasion configuration. Left panels are obtained from experimental data adapted from Ref. [10], and right panels from simulations of the model with the parameters of Fig. 4.8 and Table 4.1. (a) Phase contrast visualizing cells, (b) profile of cell speed (modulus of velocity), (c) velocity field.

Third phase of motion. The fronts of advancing cells in Figs. 4.7 and 4.8 clearly show the formation of fingers. The AVM keeps cells together while the term proportional to β in Eq. 4.8 induces a common average direction in their motion. This effect becomes stronger the larger β is, which promotes and enforces finger formation. Thus, unlike the particle model of Ref. [16], we do not need a longer range attractive potential interaction between cells. A comparison of our simulation results in Figs. 4.7 and 4.12 to the experiments reported in Refs. [10, 43] shows that the appearance and size of the cell velocity field are reproduced qualitatively. We find these areas without having to postulate the existence of special leader cells.

Our numerical simulations of spreading configurations show that cells inside a finger move faster than those at other portions of the interface. We have observed that the average velocity of finger cells may oscillate irregularly about some average value with a short period of about one hour. Fig. 4.9 shows the average velocity of 9 finger cells during a 7 hour time interval. The velocity of a single cell in the finger oscillates somewhat more irregularly in a similar fashion. For much longer time intervals, the average velocity may experience an

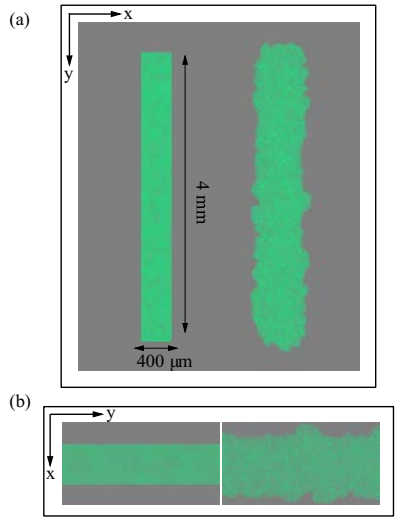

Figure 4.8: Initial configuration and configuration after 20 h of stencil removal showing the formation of fingers according to the numerical simulation of the model. (a) Full view, (b) zoom. Initial box size is 1.6 mm^2, $P^0 = 10$, $A^0 = \pi$, and shape index $p^0 = 5.65$. Parameter values are those in Table 4.1.

overall upward trend. The average velocity of boundary cells in flat regions also oscillates with time but it does not show a definite behavior overlong time intervals: it may even display a downward trend. In experiments, the velocity of cells leading inter-facial fingers has also been observed to oscillate rapidly and irregularly with periods of about one hour or less, which is similar to the findings based on numerical simulations of our model; see Fig. 101A of Ref. [43]. Some models based on continuum mechanics predict longer periods of tens of hours [43].

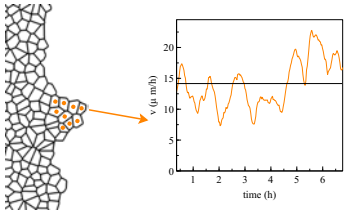

Figure 4.9: Average velocity of the marked cells during finger expansion. The velocity of each cell oscillates in a similar but somewhat more irregular manner (not shown).

In our simulations, we consider a narrow strip configuration as that in Fig. 4.8(a), which is similar to those in Ref. [43]. We adapt the SAMoS code [39] to simulate the AVM with dynamics given by Eqs. (4.8) and (4.6), in which $\boldsymbol{\varphi}_i = \mathbf{0}$. Parameter values are those in Table 4.1. Cells migrate on the surface maintaining their junctions with their neighbors, which is

enforced by the term proportional to β in Eq. (4.8). During healing, noisy forcing in Eq. (4.8) makes some cells to move faster that the others while keeping their contacts. This is the origin of the fingers or instabilities of the interface with the cell free space, which are illustrated by Fig. 4.8(b). In addition, cells on the interface, or close to it, may grow beyond the target area A^0 in Eq. 4.1. As they do so, each cell has a probability to divide into two daughter cells, which equals $r_d(A - A^0)dt$. Here dt is the time step and r_d is the division rate. We have normalized the target area to $A^0 = \pi$, and we check whether every 10 time steps if $A > A_0$. With these parameters, there is some cell division near the interface of the confluent layer and the empty space. Fig. 4.10 and Fig. 4.11 show the cell division in the simulations and in an experiment respectively. Depending on the cell type, the general shape of the velocity

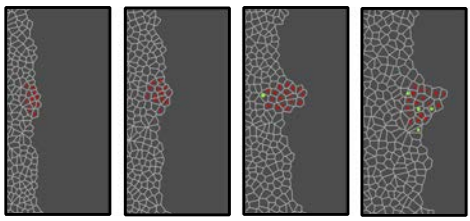

Figure 4.10: Evolution of a finger and adjacent cells for times $t = 0, t = 2.8h, t = 5.6$ and $t = 8h$. Red dots indicate the initial cells and green dots indicate those cells resulting from division.

Figure 4.11: Evolution of adjacent cells for times reported in [43]. Red dots indicate the initial cells and green dots indicate those cells resulting from division. For MDCK cells, the same neighbors and structure remains unbroken along the time.

fields, proliferation or coordinated movements change. In particular, we have focused on the

evolution and behavior of MDCK cells presented in Fig. 4.11. For the MDCK epithelium, the movements are more coordinated and extend over characteristic lengths than for the (Normal Rat Kidney) NRK. Indeed, MDCK share their neighbors along the time unlike the other type of cells. This characteristic has been also reproduced in our simulations, see Fig. 4.10.

Having calculated numerically the velocity field, we can quantify the orientational motion inside the epithelium. Take for example, the configuration after 35 h of stencil removal is shown in Fig. 4.12. In addition to the velocity field and the speed (modulus of the velocity vector) map, we have depicted a density map of the polar order parameter S_{pol},

$$S_{\mathrm{pol}} = \frac{1}{N}\sum_{i=1}^{N} \cos\vartheta_i, \quad \cos\vartheta_i = \frac{v_x(i)}{\sqrt{v_x(i)^2 + v_y(i)^2}}. \tag{4.11}$$

Here ϑ_i is the angle that the velocity vector of the ith cell forms with the outer normal to the strip (the x axis in Fig. 4.8). Fig. 4.12(b) depicts the density plot of the cellular polar order parameter, $\cos\vartheta_i$, after 35 h of stencil removal (similar to experimental data reported in Fig. 92 of Ref. [43]). Fig. 4.13 shows that an ensemble average of the polar order parameter (over 5 realizations, smooth line) increases with time and follows the same trend as the measurements reported in Ref. [10] (jagged line). At early times, S_{pol} in Fig. 4.13 does not exhibit a particular trend. The angles are distributed homogeneously and are not located in specific areas. After a while, the cells start orient themselves perpendicular to the strip, specially at the edges of the tissue, as shown in Fig. 4.12. This effect occurs in strips of width larger than 300 μm. On shorter strips, their two sides are no longer independent and the appearance of a finger changes the motion of the whole strip.

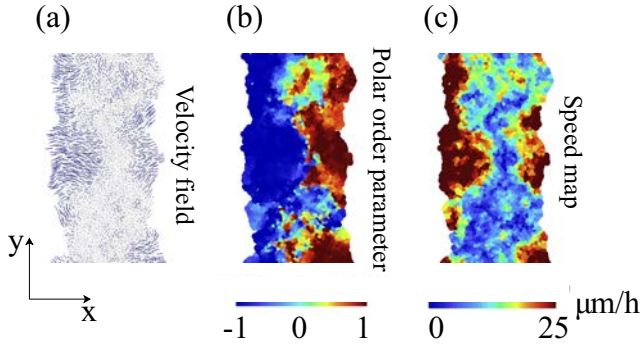

Figure 4.12: (a) Numerically simulated cell velocity field, (b) local polar order parameter $\cos\vartheta_i$, and (c) speed ($|\mathbf{v}|$) map after 35 h of stencil removal in an invasion configuration for a 400 μm wide strip. Parameter values as in Fig. 4.7.

The velocity field in Fig. 4.12(a) (from simulations) and Fig. 4.14(c) (from an experiment) exhibit swirl patterns [11]. To characterize them, we have depicted in Fig. 4.14(a) the

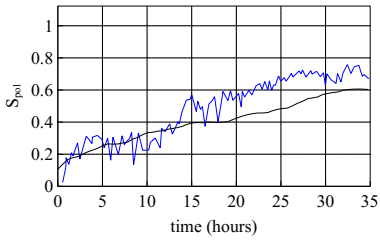

Figure 4.13: Evolution of the polar order parameter $S_{\mathrm{pol}}(t)$ corresponding to Fig. 4.7. Here $t = 0$ corresponds to 1.5 h after stencil removal [10]. An average over 5 simulations exhibits the same trend as measurements reported in Ref. [10].

correlation function for the x-component of the velocity field:

$$I(|\mathbf{r}|, t) = \frac{\langle v_x^*(\mathbf{r}', t)\, v_x^*(\mathbf{r}' + \mathbf{r}, t)\rangle_{\mathbf{r}'}}{\sqrt{\langle v_x^*(\mathbf{r}', t)^2\rangle_{\mathbf{r}'} \langle v_x^*(\mathbf{r}' + \mathbf{r}, t)^2\rangle_{\mathbf{r}'}}}, \quad v_x^*(\mathbf{r}, t) = v_x(\mathbf{r}, t) - \langle v_x(\mathbf{r}, t)\rangle_{\mathbf{r}}. \tag{4.12}$$

Here the averages are spatial averages over $\mathbf{r}'$ and also ensemble averages over simulations with different initial conditions. Fig. 4.14(b) depicts the correlation length defined by the first zero of the correlation function and the swirl size defined by its first local minimum. Empty and blue squares correspond to values given by different simulations. The best fits to straight

Figure 4.14: (a) Spatial correlation function $I(r, t)$ corresponding to Fig. 4.12(a) for different times. (b) Correlation length given by the first zero of $I(r)$ (empty squares) and swirl size given by the first local minimum of $I(r)$ (blue squares). Dashed line from swirl sizes in Ref. [11]. (c) Image from [6].

lines are also shown and compared to a similar line for Angelini *et al*'s experimental data [11]. Clearly correlation length and swirl size increase with time, indicating that cells feel each other on increasingly larger regions as time elapses. This has been observed in other experiments and simulations [10, 23]. The correlation lengths given by our simulations agree quite well with values reported in the literature for similar observation times [10, 11, 23].

A comparison of our simulation results in Figs. 4.7 and 4.8 to the experiments reported in Refs. [10, 43] shows that the appearance and size of the cell velocity field are reproduced qualitatively. Our simulations show that the area and velocity of cells both increase as their distance to the boundary of the cellular tissue decreases. Fig. 4.15 shows that cells near the interface in a spreading configuration have larger areas than cells far from the interface. This is particularly noticeable in the fingers, the cells in them are faster and have a larger area than the cells elsewhere. The cells far from the tissue border are compressed and have smaller area than boundary ones. This prediction of the under-damped AVM with dynamics as in Eq. (4.8) has been observed in experiments; see Fig. 4 from Ref. [44]. In experiments, they investigate the correlation between the area of the cells and their position in the epithelium. They conclude that the area of finger cells reaches larger values than in the simulations, which is related to the fact that we use a fixed target. As we have discussed, the Vertex model tries to maintain the whole sheet with cells with an area not far from the reference area, for that reason, in our simulations, there is not so much difference between the area of the cells as we can observe in the experiments.

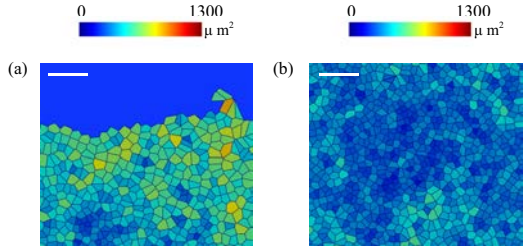

Figure 4.15: Areas of cells during a simulation of a spreading configuration: (a) Area of cells near the interface, (b) area of cells far from the interface. Our simulations exhibits the same trend as measurements reported in Ref. [44]. White bar=100μm.

4.3.2 Results for the collision configuration

Recently, Moitrier *et al.* have reported confrontation assays between antagonistically migrating cell sheets [23] whose procedure has been explained in Subsection. 1.2.2. In their experiment, the two confluent cellular monolayers (wild type and modified Ras HEK cells) advance toward an intermediate empty space, collide and the Ras monolayer displaces the wt one. The experiment shows that the velocities of the cells decay exponentially fast the farther they are from the advancing fronts [23]. If $x = L(t)$ is the position of the monolayer front, the velocity of the cells at position $x < L$ is $V^{\mathrm{wt}} \exp[(x - L)/\lambda^{\mathrm{wt}}]$ for the wt and $-V^{\mathrm{Ras}} \exp[-(x - L)/\lambda^{\mathrm{Ras}}]$ for the Ras cells at $x > L$. After the collision, these velocity functions remain the same but now V^{wt} and V^{Ras} acquire a common and lower value $-V^{\mathrm{interface}}$. Moitrier *et al* interpret their experiments by comparing with simple solutions of a 1D continuum model [23]. In our simulations, we use the SAMoS code to simulate the AVM cellular model with dynamics given by Eq. (4.8). The invading Ras cells (magenta)

move to the left whereas the wt cells (green) are pushed backward because they experience aversion to mixing with Ras cells. We model this situation by adding a negative active force $\boldsymbol{\varphi}_i^{\mathrm{Ras}} = a^{\mathrm{Ras}} \exp[-[x - L(0)]/\lambda^{\mathrm{Ras}}]$ to Ras cells in Eq. (4.8) for $x > L(0)$ (not included in Ref. [16]), whereas wt and Ras cells do not experience an active force if $x < L(0)$. We use $\lambda^{\mathrm{Ras}} = 410\,\mu$m, $a^{\mathrm{Ras}} = 9\,\mu\mathrm{m/h}^2$, $L(0) = 0$. The active force $\boldsymbol{\varphi}^{\mathrm{Ras}}$ keeps Ras cells moving to the left and pushing wt ones. Therefore we no longer need the synchronization force proportional to β to keep cells moving in the same direction. Fig. 4.17 shows finger formation for the active force $\boldsymbol{\varphi}^{\mathrm{Ras}}$ and for $\beta = 13.85\ \mathrm{h}^{-1}$, which is smaller than the value in Table 4.1. Other parameters are as indicated in Table 4.2.

α	β	τ	σ_0	K_j	Γ_j	Figure #
h^{-1}	h^{-1}	h	$\frac{\mu\mathrm{m}}{\mathrm{h}^{3/2}}$	-	-	-
0.0602	13.85	1.66	55.88	1	1	8
0.42	0.602	1.66	13.97	1	1	9, 10

Table 4.2: Two sets of parameters corresponding to the experiments with HEK cells in Ref. [23].

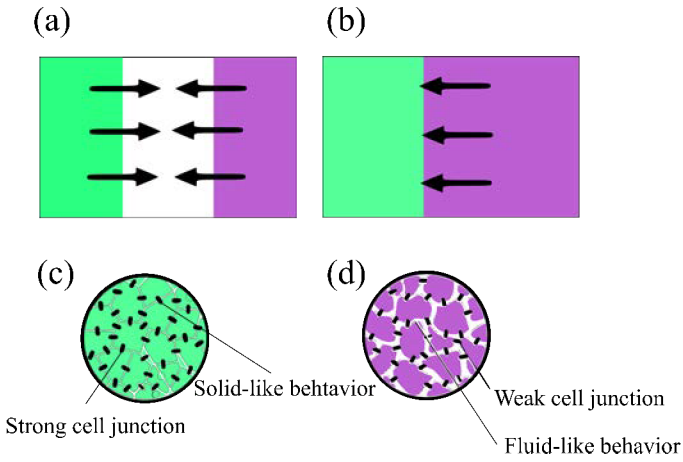

Figure 4.16: Sketch of the antagonistic migration assay. (a) and (b) initial configuration and invasion from purple population respectively. (c) green cells maintain a strong cell junction between their neighbors. In (d) purple cells have a fluid-like behavior and the cell junctions are weak. Modified figure from [9].

Our underdamped AVM uses more features of wt and Ras cells obtained from the experiments than kept by continuum models. The latter lose features at distances close to the cell size. Continuum models fit friction, viscosity and strength of active forces for the two cell populations to explain how Ras cells invade the wt monolayer [23].

The AVM allows us to study tissues that behave differently. In our simulations, 5000 cells are split into two populations with different properties specified by the junction tensions Λ_{ij}, $j = 1, 2$, which affect each pair of cell-cell contacts. The simulations producing Figs. 4.17,

4.18 and 4.19 have open boundaries because we have focused on the interface between populations. We have fixed $K = \Gamma = 1$ and $-6.8 = \Lambda_{22} < \Lambda_{11} = -6.2$, which produce shape parameters p^0 of 3.50 (green cells) and 3.84 (magenta cells), below and above the transition value $p_0^0 = 3.812$, respectively. Thus, Ras magenta cells are fluid-like (supercritical shape index) and their density is smaller than that of the solid-like wt cells. This is consistent with the observation that wt cells have larger mean traction force amplitudes than Ras cells [23], see sketch in Fig. 4.16. Our aim is to analyze the effect of Λ_{12} on the AMA. Both monolayers occupy the right and left portions of a 4.4 mm wide, 3.1 mm tall box. In Figs. 4.17-4.19, we show a 1 mm $\times$2.5 mm region. In our simulations, we start from a configuration with

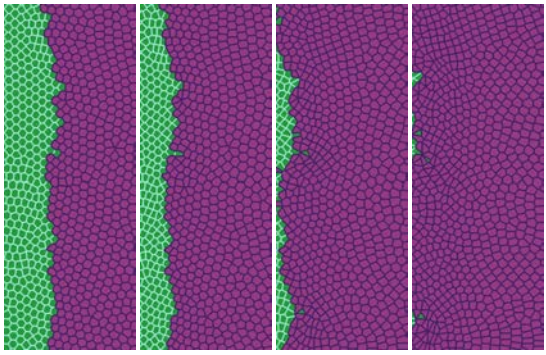

Figure 4.17: Simulation of the antagonistic migration assay. Junction tensions are $\Lambda_{11} = -6.2$, $\Lambda_{22} = -6.8$, which yield shape indices 3.50 (green cells) and 3.84 (magenta cells), respectively. Other parameters are listed in the first row of Table 4.2, and $\Lambda_{12} = -7.0 < \frac{1}{2}(\Lambda_{11} + \Lambda_{22})$ correspond to weak population mixing. Snapshots are taken at times 2 h, 6.5 h, 13 h, 20 h.

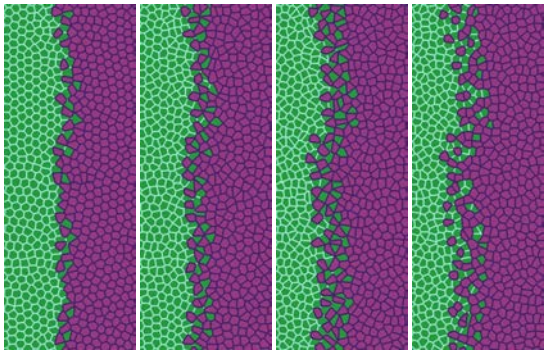

Figure 4.18: Simulation of the antagonistic migration assay. First and second snapshots: $\Lambda_{12} = -7.5 < \frac{1}{2}(\Lambda_{11} + \Lambda_{22})$ (population mixing); third and fourth snapshots: $\Lambda_{12} = -6.0 > \frac{1}{2}(\Lambda_{11} + \Lambda_{22})$ (population segregation). Other parameters are listed in the second row of Table 4.2 whereas times are as in Fig. 4.17.

the cell populations separated by a straight vertical interface at $L(0) = 0$. The active force $\boldsymbol{\varphi}$

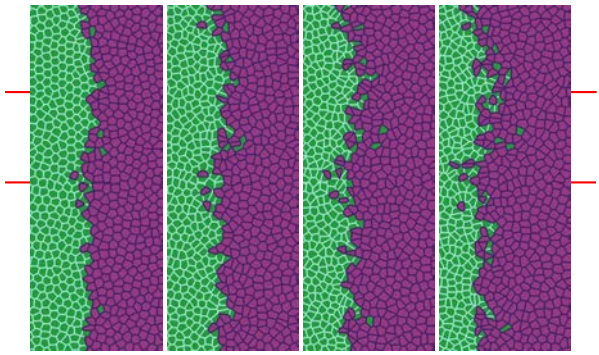

Figure 4.19: Simulation of the antagonistic migration assay. Parameters and times are as in Fig. 4.18, except that one fifth of the overall population (randomly placed green and magenta cells) have $\Lambda_{12} = -7.5$ (population mixing) and the other four fifths have $\Lambda_{12} = -6.0$ (population segregation). The marked region has similar size to that reported in experiments [23].

pushes Ras cells with $x > L(0)$ to the left, whereas $\boldsymbol{\varphi} = \mathbf{0}$ for any cell to the left of $x = L(0)$. The junction tension Λ_{12} in Fig. 4.17 ($\Lambda_{12} = -7.0$) and in the two left panels of Fig. 4.18 ($\Lambda_{12} = -7.5$) favors population mixing. Ras (magenta) cells push wt (green) cells backwards at a velocity close to the observed $V^{\text{interface}}$, meanwhile creating a rugged interface between cell populations. As time elapses, fingers and some isolated islands (lagging wt in the Ras assembly and advancing Ras islands in the receding wt assembly) appear. These effects are more pronounced the smaller Λ_{12} is, as shown by comparison of Figs. 4.17 and 4.18. It is possible to create some realistic mixing of the populations by changing the junction tension Λ_{12} with time. The first two snapshots in Fig. 4.18 have $\Lambda_{12} = -7.5 < \frac{1}{2}(\Lambda_{11} + \Lambda_{22})$, which favors population mixing. Then the interface between cell populations becomes very rugged and there appear islands of one cell type inside a layer of the other type. The third and fourth snapshots in Fig. 4.18 have been obtained with $\Lambda_{12} = -6.0 > \frac{1}{2}(\Lambda_{11} + \Lambda_{22})$ that favors population segregation. The interface becomes smoother and the islands shrink and tend to disappear.

We have also focused on the effects of cellular alignment. There are two terms in Eq. (4.8) that try to synchronize cell velocities: the term proportional to β and the active force $\boldsymbol{\varphi}$, which pushes the Ras cells to the left. Although the values of β used to draw Figs. 4.17-4.19 are smaller than that in Table 4.1, different β still make a difference in the behavior during tissue collision, specially in the Ras population. Fig. 4.17 exhibits global polar migration because its β value is larger than that in Figs. 4.18 and 4.19, but types of cells are not mixed despite having a favorable value $\Gamma_{12} = -7.0$. The smaller value of β in Figs. 4.18 and 4.19 creates a weaker polar alignment than that in Fig. 4.17. The different patterns observed in these figures illustrate that cell alignment affects importantly the shape and configuration of the interface.

While the rightmost panel of Fig. 4.18 is similar to some of the experimental data [23], we can obtain a similar formation of islands and fingers by assuming that Λ_{12} is randomly distributed among cells. In particular, we assume that one fifth of magenta and green cells have

$\Lambda_{12} = -7.5$, which favors mixing of populations, while the remaining ones have $\Lambda_{12} = -6.0$ and favor population segregation. The result is depicted in Fig. 4.19, which exhibits behavior similar to experimental observations [23]. The topological data analyses of the next Section characterize the geometry of the interface between cell types in antagonistic migration assays.

4.4 Formation of islands and Topological Data Analysis

Experiments and numerical simulations of cell monolayers produce time series of images that make it possible to identify the structure of interfaces and to compare their time evolution. It is quite cumbersome to process manually these time series. Here we use Topological Data Analysis (TDA) as a computational tool to process automatically time series of images. We next illustrate how to use TDA for this purpose and how to interpret the obtained results. We focus on specific parts of selected snapshots of images from experiments and then on time series of images from numerical simulations. While we have few images of interfaces from experiments, we can generate arbitrarily many from numerical simulations. Having many images, the automatic TDA tool enables us to describe in detail the topological changes of the interfaces and to implement hierarchical clustering strategies, thereby classifying the evolving interface structures. Fig. 4.20 shows the interfaces between two colliding confluent cellular monolayers in an AMA [23]. In this experiment, magenta Ras cells make green wild type cells move back, cf. third and fourth snapshots in the cover of Soft Matter, vol. 15 [23]. The interface between the two cell populations is rather rough, it exhibits fingers, and there are islands or pockets of green cells left behind by the advance of the magenta front. To quantify these phenomena in an automatic way, we proceed as follows. Using Matlab, we transform the images in matrices of ones (green) and zeros (magenta). Then we extract the positions of green/magenta interfaces, represented by the point clouds shown in Fig. 4.20, and process them using TDA. We pursue a similar strategy for images extracted from numerical simulations of our underdamped AVM, which yields a more complete picture of the evolution of interfaces

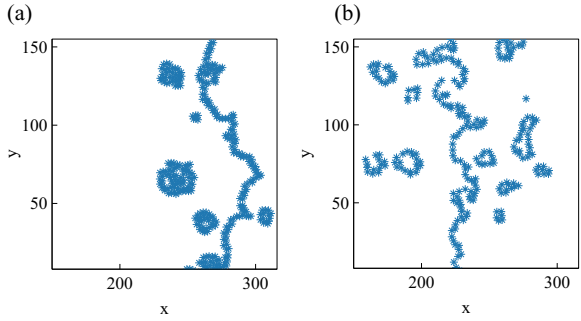

Figure 4.20: Structure of the interface between colliding layers corresponding to two snapshot sof the collision of two confluent cellular monolayers in Moitrier et al's experiment [23]. These profiles correspond to (a) the third and (b) the fourth panels (counting from the left) in the cover of Soft Matter corresponding to Ref. [23]

4.4.1 Persistent homology

A finite set of data points may be considered a sampling from the underlying topological space. Homology distinguishes topological spaces (e.g., annulus, sphere, torus, or more complicated surface or manifold) by quantifying their connected components, topological circles, trapped volumes, and so forth. Persistent homology characterizes the topological features of clouds of point data or particles at different spatial resolutions [48]. Highly persistent features span a wide range of spatial scales. Persistent features are more likely to represent true features of the data/pattern under study than to constitute artifacts of sampling, noise, or parameter choice [47]. To find the persistent homology of a cloud of point data/set of particles, we must first view them as a simplicial complex C. Roughly speaking, a simplicial complex is defined by a set of vertices (points or particles) and collections of k-simplices. The latter are the convex hulls of subsets with $k+1$ vertices, comprising also faces; see the Appendix for precise definitions. Defining a distance function on the underlying space (the euclidean distance, for instance), we can generate a filtration of the simplicial complex, which is a nested sequence of increasingly bigger subsets. More precisely, a filtration of a simplicial complex C is a family of sub-complexes $\{C(r)|r \in R\}$ of C such that $C(r) \subset C(r')$ whenever $r \leq r'$. The filtration value of a simplex $S \in C$ is the smallest r such that $S \in C(r)$. The motivation for studying the homology of simplicial complexes is the observation that two shapes can be distinguished by comparing their holes. For $k \in N$, the Betti number b_k counts the number of k-dimensional holes. A k-dimensional Betti interval $[r_b, r_d)$ represents a k-dimensional hole that is created at the filtration value r_b, exists for $r_b \leq r < r_d$ and disappears at value r_d. We are interested in Betti intervals that persist for a large filtration range, they describe how the homology of $C(r)$ changes with r. How do we construct a filtration? The Vietoris-Rips filtration $VR(X, r)$ [47, 48], which we will use here, is constructed as follows,

1. The set of vertices X is the cloud of points under study.
2. Given vertices x_1 and x_2, the edge $[x_1, x_2]$ is included in $VR(X, r)$ if the distance $d(x_1, x_2) \leq r$.
3. If all the edges of a higher dimensional simplex are included in $VR(X, r)$, the simplex belongs to $VR(X, r)$.

A default choice for the distance d to study homology of 2D particle configurations is the Euclidean metric. Fig. 4.21 displays two simplexes of a Vietoris-Rips filtration for the point cloud in Fig. 4.20. Notice the appearance and disappearance of holes and isolated components as the threshold distance r to connect points increases. This filtration is governed by three parameters,

1. The maximum dimension d_{max}. This is the maximum dimension of the simplices to be constructed. The persistent homology (characterized by its Betti numbers) can be computed up to dimension $d_{max} - 1$. In this case $d_{max} = 2$, we consider points (0-simplices), edges (1-simplices), and triangles (2-simplices).

2. The maximum filtration value r_{max} and the number of divisions N. These values define the filtered simplicial complexes to be constructed, for $r \in \{0, \frac{r_{max}}{N-1}, \frac{2r_{max}}{N-1}, \frac{(N-2)r_{max}}{N-1}, r_{max}\}$.

Notice that for a set of P points, the full simplicial complex will have about $2^P - 1$ simplices in it. Therefore, d_{max} and r_{max} are usually slowly increased to get information without reaching computational limits. The computation is not too sensitive to the specific value of N. When r_{max} is greater than the diameter of the point cloud, all possible edges form and join all the points in one simplex. In the next two Sections, we apply TDA to experimental and numerical

Figure 4.21: Visualization of the complexes $VR(X, r)$ for the point cloud depicted in Fig. 4.20 when (a) $r = 6$ and (b) $r = 10$. For large enough r all the components merge in a single one. Holes appear and disappear as new connections are created, reflecting the overall point cloud arrangement.

images.

4.4.2 TDA for experiments

Let us consider the snapshots depicted in Fig. 4.20 and Fig. 4.21 processes the earlier snapshot depicted in Fig. 4.20(a), in which the green and magenta monolayers have made contact and started interpenetrating each other. Ras cells (magenta) are pushing back wt cells (green) towards the left. As they do so, there are islands of wt cells inside the Ras monolayer. How does TDA capture these features? After constructing the Vietoris-Rips filtration, there are two commonly employed graphical representations that visualize the persistent homology of a point cloud: barcodes and persistence diagrams [45]. Barcodes of a homology H_k depict Betti intervals $[r_b, r_d)$ for k-holes ($k > 0$) or connected components ($k = 0$) as the filtration parameter r varies. The homology class H_0 comprises the points forming the green/magenta interfaces. As the size filtration parameter r increases from zero, there appear edges joining these points, thereby forming clusters as illustrated by Fig. 4.21 for specific values of r and indicated by the barcodes in Fig. 4.22(a) for the selected range of r. The class H_1 further distinguishes compact components of the interface that are detached from the main part of the interface and form topological cycles, cf. the corresponding barcode in Fig. 4.22(a). These components are islands of one cell type (phase) inside the bulk of the other phase. Persistence diagrams represent the Betti intervals by points in a birth-death plane. The x axis represents the filtration value r at which components/holes are created. The

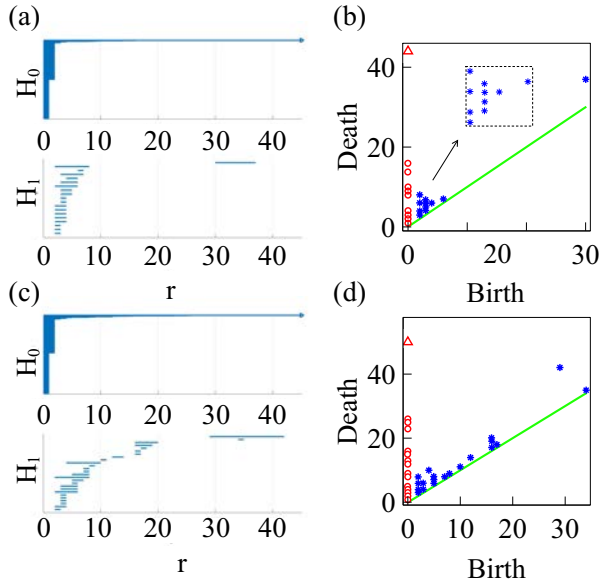

Figure 4.22: Barcodes (left) and persistence diagrams (right) for the homologies H_0 (circles) and H_1 (asterisks) of the interfaces separating cell types in images from experiments and numerical simulations. We use Vietoris-Rips filtrations with parameters N and r_{max}. (a)-(b) TDA from Fig. 4.20(a) (experiments) with $N = 45$, $r_{max} = 45$; (c)-(d) TDA from Fig. 4.20(b) (experiments) with $N = 45$, $r_{max} = 45$;Points in the persistence diagrams mark the beginning (birth) and end (death) of a bar (homology class) in the barcode. Triangles represent a component with infinite persistence. The green line is the diagonal.

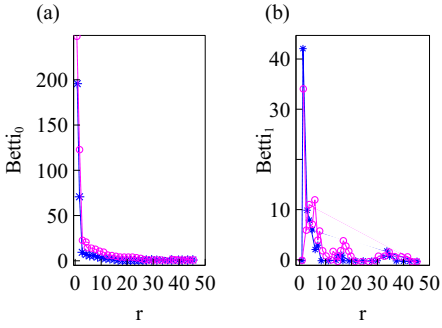

Figure 4.23: a)-(b) Betti numbers versus filtration parameter diagrams for Fig. 4.20(a) (blue asterisks) and 14(b) (magenta circles, later time in the AMA experiment) show that the number of clusters and holes in the interface between aggregates increases with time. As a result of island formation and motion, which increases with time, Panel (a) shows that the number of components decreases more slowly with r for the later time. The additional peaks in Panel (b) near $r = 40$ correspond to islands that have already penetrated further inside the other cell population in the experiment.

y axis represents the filtration value r at which they disappear. Those points less close to the diagonal (green) tend to mark robust underlying geometrical features. Fig. 4.22(b) depicts the persistence diagram corresponding to Fig. 4.20(a). Red circles mark connected components of the interface between cell monolayers and the magnitude of the filtration parameter r at which they disappear. As the filtration parameter increases, points comprising the main front merge rapidly in one component that absorbs neighboring clusters. They correspond to blocks of bars in the H_0 panel of Fig. 4.22(a) that start at the lowest value of r. Blue asterisks represent the appearance (horizontal axis) and disappearance (vertical axis) of holes inside such clusters. The first column of asterisks represents the ten bars in the H_1 panel of Fig. 4.22(a) that start at the same value of r and form four groups of bars, which end at about the same value of r. The remaining bars and asterisks are similarly related. They represent the new holes that form as the clusters merge, which gives an idea of the relative arrangement thereof. Relatively narrow barcodes produce points in the persistence diagram that are close packed. Fig. 4.22(c)-(d) display the barcodes and persistence diagram corresponding to Fig.14(b). Compared to the earlier snapshot of Fig. 4.20(a) and its TDA in Fig. 4.22(a)-(b), there are more islands of each phase in the bulk of the other: the invasion of Ras cells leaves pockets of wt cells inside their midst. The main interface has become more meandering and exhibits more fingers than in the earlier snapshot. As a consequence, the number of clusters or interface components is larger than at the earlier time. Similarly, there are more topological cycles, which reflects the larger number of islands of one cell type in the midst of the other cell type. Barcodes and persistence diagram are more spread out. This is further quantified by the Betti numbers b_j that count the number of elements in H_j , for $j = 0$ (clusters) and for $j = 1$ (holes), as depicted in Fig. 4.23(a)-(b) for the snapshots shown in Fig. 4.20.

4.4.3 TDA of numerical simulations

As indicated in the previous Section, to observe island formation, we have to tune the (negative) junction tensions when simulating antagonistic migration assays. In particular, $\Lambda_{12} < \frac{1}{2}(\Lambda_{11} + \Lambda_{22})$ facilitates mixing of wt and modified cell populations whereas $\Lambda_{12} > \frac{1}{2}(\Lambda_{11} + \Lambda_{22})$ produces population segregation. In Fig. 4.18, Λ_{12} switches from population 2 mixing to segregation after the two first snapshots. Then the pockets of green cells left behind by the advance of the interface shrink and start disappearing, as shown in the third and fourth snapshots of Fig. 4.18. If mixing is weaker, as in Fig. 4.17, the interface forms pronounced fingers, there are less islands and we do not need to change the junction tensions with time. In Fig. 4.19, Λ_{12} randomly takes on a mixing value for one fifth of Ras and wt cells and on a segregation value for the others. The results of changing interface and island formation are qualitatively similar to those observed in experiments.

Let us now interpret the evolution shown in the panels of Fig. 4.19 using TDA. Fig. 4.24 and Fig. 4.25 show the barcodes, persistence diagrams and Betti numbers for the marked Sections of the leftmost and rightmost panels in Fig. 4.19. As before, we represent the interfaces by point clouds. At $r = 0$, each point of the interface is a component. For the more regular interface of the leftmost panel in Fig. 4.19, increasing r produces point

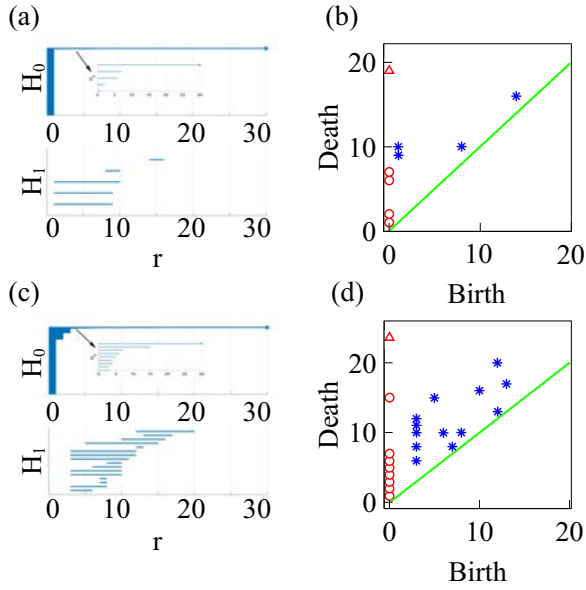

Figure 4.24: Barcodes (left) and persistence diagrams (right) for the homologies H_0 (circles) and H_1 (asterisks) of the interfaces separating cell types in images from experiments and numerical simulations. (a)-(b) TDA from the leftmost panel in Fig. 4.19 (numerical simulations) with $N = 60$, $r_{max} = 30$; (c)-(d) TDA from the rightmost panel in Fig. 4.19 (numerical simulations) with $N = 30$, $r_{max} = 30$.

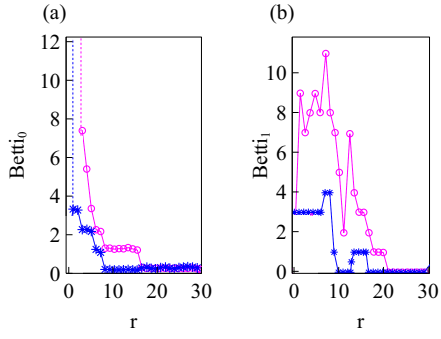

Figure 4.25: (a)-(b) Same for the numerical simulations considered in Fig. 4.24 corresponding to the leftmost and rightmost panels in Fig. 4.19. As a result of island formation and motion, which increases with time, Panel (c) shows that the number of components decreases more slowly with r for the later time. The peaks in Panels (a) and (b) are similar for r below 20.

components appearing as the short H_0 bars in Fig. 4.24(a). These bars end at similar filtration values and appear as a single red circle in the persistence diagram of Fig. 4.24(b). The main three islands correspond to the three intermediate bars in the inset of Fig. 4.24(a), which disappear at larger filtration values. The lowest circle in Fig. 4.24(b) represents the point components, the three intermediate ones represent the islands in the barcode and their sizes. All clusters finally merge in the main front represented by the arrow on top of the vertical axis in Fig. 4.24(b). Analysis of H_1 confirms that the intermediate circles/bars are round islands and not strings. Each component corresponds to a cycle represented by the three largest H_1 bars in Fig. 4.24(a) and the two first asterisks in Fig. 4.24(b), one of which represents the two bars of similar length. The two shortest bars represent holes formed as components merge during the filtration process and correspond to the two asterisks closer to the diagonal in Fig. 4.24(b).

Figs. 4.24(c)-(d) correspond to the more meandering interface of the rightmost panel in Fig. 4.19. There are more points in the cloud representing the interface, whose irregularity results in different extinction values of r for the associate H_0 bars. The main seven islands correspond to the intermediate bars in the inset of Figs. 4.24(c), and their extinction values in the persistence diagram give an idea of the distance to the main front or to another island. The fact that they are islands (enclosed by a boundary) is inferred from the H_1 bars in Figs. 4.24(c). They correspond to the seven bars that appeared first, which are also represented by the first column five asterisks in Figs. 4.24(d) having smaller r. Two of the asterisks correspond to two islands of similar size length each, which have bars of similar size. The length of the bars in the barcode or the distance of the asterisks from the diagonal in the persistence diagram give an idea of the island size. Additional H_1 bars represent holes created during the filtration process as components merge and give an idea of the relative arrangement of the islands or of the fingers in the main front. They are represented ary the additional asterisks in Figs. 4.24(d). The Betti numbers in Figs. 4.25(a)-(b) show a larger number of island and holes as time increases from the leftmost snapshot in Figs. 4.19 to the rightmost one. Compared to the TDA of experiments in Figs. 4.22(a)-(d), there are no gaps between bars and asterisks appearing for large r in Figs. 4.24(a)-(d). The reason is that the distance of islands to the main front is smaller for the simulation than for the experiment.

4.5 Conclusions

We have modeled how epithelial cell aggregates (i) advance through empty spaces (wound healing, tissue spreading) and (ii) collisions between aggregates (tumoral invasion) using an AVM with dynamics for cell centers that includes collective tissue forces [35], velocity alignment and inertia [16]. The AVM implements exchanges of neighboring cells automatically (T1 transitions). Compared with particle models with underdamped dynamics studied in [16], our model accounts for fingering instabilities in spreading tissue without having to distinct between type of cells.

Compared to continuum models [34], stochasticity enables our model to reproduce

the observed fast irregular oscillation of cell velocities in fingers [43] and the spatial auto-correlation of the velocity [11]. Indeed, our underdamped AVM predicts that cells at the interface and the fingers have larger area than those well inside the tissue, which has been corroborated by recent experiments [44]. We also observe in numerical simulations of tissue spreading that the velocity of the fastest cell in a finger may oscillate with a short period in a range between 30 minutes to about one hour. A similar short period oscillation has been observed in experiments; cf Figure 101A in L. Petitjean's PhD thesis [43]. Thus, for spreading tissue, detailed comparison to experimental data provides a quantitatively accurate description of cell motion (speed, velocity correlation function and polar order parameter).

For antagonistic migration assays, we have reproduced collisions in which one cell population pushes back another whereas both populations mix forming different types of interfaces. The key element to model mixing is to keep different junction parameters for the two colliding tissues, the invading cells are liquid-like whereas the receding tissue comprises solid-like cells. In addition, a fraction of cells favor mixing, the others segregation, and that these cells are randomly distributed in space. Thus characterized, numerical simulations produce outcomes similar to those observed in experiments [23].

To characterize automatically the dynamics of islands and the rugged interface between aggregates, we have introduced topological data analyses of experiments and time series from numerical simulations. In collisions between aggregates, the interface between wt and Ras cell populations roughens and islands appear. The persistence diagrams of Homology classes 0 (clusters) and 1 (cycles) spread out and the number of these classes given by the corresponding Betti numbers increases. Using time series of data generated by numerical simulations, we have explained how to cluster interfaces using distance matrices based on the bottleneck distance between their persistence diagrams, which are stable to perturbations in the process. Despite the amount of data from experiments being limited, disruptive events such as island and cluster formation can be automatically captured by topological data analyses of numerical simulations and contrasted with experiments. Similarly, the Wasserstein distance between images enables us to track and classify automatically the evolving shapes of interfaces between cell populations by using time series from experimental or numerical studies. These techniques of topological data analysis are scalable and could be used in studies involving large amounts of data whenever available.

Our results have allowed to extract parameter values and to determine biologically relevant physical mechanisms for characterizing confluent motion of cellular aggregates, as described above. In particular, (i) cells at the interface are larger, inform the aggregate motion and are influenced by it, without needing leader cells to form fingers at the interface; and (ii) in colliding cellular aggregates, the solid or liquid like character of the cells (as determined by their junction parameters) decides the way the invasion goes.

References

[1] P. Friedl, P. B. Noble, P. A. Walton, D. W. Laird, P. J. Chauvin, R. J. Tabah, M. Black, and K. S. Zänker, Migration of coordinated cell clusters in mesenchymal and epithelial cancer explants in vitro, Cancer Res. **55**, 4557 (1995).

[2] P. Friedl and K. Wolf, Tumour-cell invasion and migration: diversity and escape mechanisms, Nature Cancer Res. **3**, 362(2003).

[3] P. Friedl and D. Gilmour, Collective cell migration in morphogenesis, regeneration and cancer, Nature Rev. Cell Biol. **10**, 445 (2009).

[4] C. J. Weijer, Collective cell migration in development, J. Cell Sci. **122**(18), 3215 (2009).

[5] O. du Roure, A. Saez, A. Buguin, R. H. Austin, P. Chavrier, P. Silberzan, and B. Ladoux, Force mapping in epithelial cell migration, Proc. Natl. Acad. Sci. U. S. A. **102**, 2390 (2005).

[6] M. Poujade, E. Grasland-Mongrain, A. Hertzog, J. Jouanneau, P. Chavrier, B. Ladoux, A. Buguin and P. Silberzan, Collective migration of an epithelial monolayer in response to a model wound, Proc. Natl. Acad. Sci. U. S. A. **104**, 15988 (2007).

[7] X. Trepat, M.R. Wasserman, T.E. Angelini, E. Millet, D.A. Weitz, J.P. Butler, and J.J. Fredberg, Physical forces during collective cell migration, Nature Phys. **5**, 426 (2009).

[8] P. Rørth, Collective Cell Migration, Ann. Rev. Cell Dev. Biol. **25**, 407 (2009).

[9] A. Shellard, and R. Mayor, (2020), Rules of collective migration: from the wildebeest to the neural crest. Philosophical Transactions of the Royal Society B, 375(1807), 20190387.

[10] L. Petitjean, M. Reffay, E. Grasland-Mongrain, M. Poujade, B. Ladoux, A. Buguin, and P. Silberzan, Velocity fields in a collectively migrating epithelium, Biophys. J. **98**, 1790 (2010).

[11] T. Angelini, E. Hannezo, X. Trepat, J. J. Fredberg, and D. A. Weitz, Cell Migration Driven by Cooperative Substrate Deformation Patterns, Phys. Rev. Lett. **104**, 168104 (2010).

[12] T. Angelini, E. Hannezo, X. Trepat, M. Marquez, J.J. Fredberg, and D.A. Weitz, Glass-like dynamics of collective cell migration, PNAS **108**, 4714 (2011).

[13] X. Trepat and J. J. Fredberg, Plithotaxis and emergent dynamics in collective cellular migration, Trends in Cell Biology **21**(11), 638 (2011).

[14] P. Rørth, Fellow travellers: emergent properties of collective cell migration, EMBO Rep. **13**, 984 (2012).

[15] X. Serra-Picamal, V. Conte, R. Vicent, E. Anon, D.T. Tambe, E. Bazellieres, J.P. Butler, J.J. Fredberg, and X. Trepat, Mechanical waves during tissue expansion, Nature Phys. **8**, 628 (2012).

[16] N. Sepúlveda, L. Petitjean, O. Cochet, E. Grasland-Mongrain, P. Silberzan, and V. Hakim, Collective Cell Motion in an Epithelial Sheet Can Be Quantitatively Described by a Stochastic Interacting Particle Model, PLOS Comput. Biol. **9**, e1002944 (2013).

[17] A. Brugués, E. Anon, V. Conte, J.H. Veldhuis, M. Gupta, J. Colombelli, J.J. Muñoz, G.W. Brodland, B. Ladoux, and X. Trepat, Forces driving epithelial wound healing, Nature Phys. **10**, 683 (2014).

[18] A. Ravasio, I. Cheddadi, T. Chen, T. Pereira, H. T. Ong, C. Bertocchi, A. Brugues, A. Jacinto, A. J. Kabla, Y. Toyama, X. Trepat, N. Gov, L.N. de Almeida, and B.Ladoux, Gap geometry dictates epithelial closure efficiency, Nature Comm. **6**, 7683 (2015).

[19] J.-A. Park, J.-H. Kim, D. Bi, J.A. Mitchel, N. T. Qazvini, K. Tantisira, C. Y. Park, M. McGill, S.-H. Kim, B. Gweon, J. Notbohm, R. Steward Jr, S. Burger, S.H. Randell, A. T. Kho, D. T. Tambe, C. Hardin, S. A. Shore, E. Israel, D. A. Weitz, D. J. Tschumperlin, E. P. Henske, S.T. Weiss, M. L. Manning, J. P. Butler, J.M. Drazen, and J.J. Fredberg, Unjamming and cell shape in the asthmatic airway epithelium, Nature Mat. **14**, 1040 (2015).

[20] S. Porazinski, J. de Navascués, Y. Yako, W. Hill, M. R. Jones, R. Maddison, Y. Fujita, and C. Hogan, EphA2 Drives the Segregation of Ras-Transformed Epithelial Cells from Normal Neighbors, Current Biol. **26**, 3220 (2016).

[21] H. B. Taylor, A. Khuong, Z. Wu, Q. Xu, R. Morley, L. Gregory, A. Poliakov, W. R. Taylor, and D. G. Wilkinson, Cell segregation and border sharpening by Eph receptor-ephrin-mediated heterotypic repulsion, J. R. Soc. Interface **14**, 20170338 (2017).

[22] V. Hakim and P. Silberzan, Collective cell migration: a physics perspective, Rep. Prog. Phys. **80**, 076601 (2017).

[23] S. Moitrier, C. Blanch-Mercader, S. Garcia, K. Sliogeryte, T. Martin, J. Camonis, P. Marcq, P. Silberzan, and I. Bonnet, Collective stresses drive competition between monolayers of normal and Ras-transformed cells, Soft Matter **15**, 537 (2019).

[24] K. Ganesh, H. Basnet, Y. Kaygusuz, A. M. Laughney, L. He, R. Sharma, K. P. O'Rourke, V. P. Reuter, Y.-H. Huang, M. Turkekul, E. E. Er, I. Masilionis, K. Manova-Todorova, M. R. Weiser, L. B. Saltz, J. Garcia-Aguilar, R. Koche, S. W. Lowe, D. Pe'er, J. Shia, and J. Massagué, L1CAM defines the regenerative origin of metastasis-initiating cells in colorectal cancer, Nature Cancer **1**, 28 (2020).

[25] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R.A. Simha, Hydrodynamics of soft active matter, Rev. Mod. Phys. **85**, 1143 (2013).

[26] E. Méhes and T. Vicsek, Collective motion of cells: from experiments to models, Integrative Biol, **6**, 831 (2014).

[27] D. Bi, X. Yang, M.C. Marchetti, and L. Manning, Motility-Driven Glass and Jamming Transitions in Biological Tissues, Phys. Rev. X **6**, 021011 (2016).

[28] C. Malinverno, S. Corallino, F. Giavazzi, M. Bergert, Q. Li, M. Leoni, A. Disanza, E. Frittoli, A. Oldani, E. Martini, T. Lendenmann, G. Deflorian, G.V. Beznoussenko, D. Poulikakos, K. H. Ong, M. Uroz, X. Trepat, D. Parazzoli, P. Maiuri, W. Yu, A. Ferrari, R. Cerbino, and G. Scita, Endocytic reawakening of motility in jammed epithelia, Nature Mat. **16**, 587 (2017).

[29] F. Giavazzi, C. Malinverno, S. Corallino, F. Ginelli, G. Scita, and R. Cerbino, Giant fluctuations and structural effects in a flocking epithelium, J. Phys. D: Appl. Phys. **50**, 384003 (2017).

[30] F. Giavazzi, M. Paoluzzi, M. Macchi, D. Bi, G. Scita, L. Manning, R. Cerbino, and C. Marchetti, Flocking Transition in Confluent Tissues, Soft Matter **14**, 3471 (2018).

[31] A. Palamidessi, C. Malinverno, E. Frittoli, S. Corallino, E. Barbieri, S. Sigismund, G. V. Beznoussenko, E. Martini, M. Garre, I. Ferrara, C. Tripodo, F. Ascione, E. A. Cavalcanti-Adam, Q. Li, P. P. Di Fiore, D. Parazzoli, F. Giavazzi, R. Cerbino, and G. Scita, Unjamming overcomes kinetic and proliferation arrest in terminally differentiated cells and promotes collective motility of carcinoma, Nat. Mater. **18**, 1252 (2019).

[32] P. Rodríguez-Franco, A. Brugués, A. Marín-Llauradó, V. Conte, G. Solanas, E. Batlle, J. J. Fredberg, P. Roca-Cusachs, R. Sunyer, and X. Trepat, Long-lived force patterns and deformation waves at repulsive epithelial boundaries, Nature Mater. **16**, 1029 (2017).

[33] G. Y. Ouaknin and P. Z. Bar-Yoseph, Stochastic collective movement of cells and fingering morphology: no maverick cells, Biophys. J. **97**, 1811(2009).

[34] R. Alert, C. Blanc-Mercader, and J. Casademunt, Active Fingering Instability in Tissue Spreading, Phys. Rev. Lett. **122**, 088104 (2019).

[35] D. L. Barton, S. Henkes, C. J. Weijer, and R. Sknepnek, Active Vertex Model for Cell-Resolution Description of Epithelial Tissue Mechanics, PLoS Comput. Biol. **13**, e1005569 (2017).

[36] R. Alert and X. Trepat, Physical Models of Collective Cell Migration, Ann. Rev. Cond. Matter Phys. **11**, (2020) and arXiv:1905.07675.

[37] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, Novel type of phase transition in a system of self-driven particles, Phys. Rev. Lett. **75**, 1226 (1995).

[38] T. Vicsek, and A. Zafeiris, Collective motion, Phys. Rep. **517**, 71 (2012).

[39] Soft Active Matter on Surfaces (SAMoS);. Available from: https://github.com/sknepneklab/SAMoS.

[40] H. Honda and G. Eguchi, How much does the cell boundary contract in a monolayered cell sheet?, J. Theor. Biol. **84**, 575 (1980).

[41] D. Selmeczi, S. Mosler, P.H., Hagedorn, N. B., Larsen and H., Flyvbjerg. (2005). Cell motility as persistent random motion: theories from experiments, Biophysical journal, **89**, 912.

[42] L. L. Bonilla and C. Trenado, Contrarian compulsions produce exotic time-dependent flocking of active particles, Phys. Rev. E **99**, 012612 (2019).

[43] L. Petitjean, Réponse active d'un épithélium à une stimulation mécanique, PhD Thesis, University Paris VI – Pierre et Marie Curie. 26 September 2011.

[44] J.Q. Lv, P.C. Chen, W.T. Góźdź, and B. Li, Mechanical adaptions of collective cells nearby free tissue boundaries, Journal of Biomechanics, 109763 (2020).

[45] C. Topaz, L. Ziegelmeier, and T. Halverson, Topological Data Analysis of Biological Aggregation Models, PLOS One **10**, e0126383 (2015).

[46] R. Ghrist, Elementary Applied Topology. CreateSpace Independent Publishing Platform, 2014.

[47] G. Carlsson, Topology and data. AMS Bulletin **46**, 255 (2009).

[48] A. Zomorodian and G. Carlsson, Computing persistent homology, Discrete Comput. Geom. **33**, 249 (2005).

Conclusions

The three main goals presented at the beginning of this thesis are: i) the analysis of the flocking in the two dimensional Vicsek model, carrying out a linear stability analysis of the disordered state and studying the possible bifurcations from the kinetic equation discrete in time and space, ii) the study of a modified Vicsek model including a different mechanism of synchronization aiming to understand and analyze the ordered state solutions of the amplitude equations corresponding to period-doubling, Hopf, or pitchfork bifurcations of the disordered state and iii) the qualitative and quantitative comparison between numerical simulations and topological data analysis, and the experiments available in the literature of epithelial monolayers. The main conclusions of the analyses carried out in this thesis are summarized below.

Flocking in the Vicsek Model

The standard Vicsek Model has been studied by analyzing its kinetic equation for one-particle distribution function in the limit of infinitely many particles by assuming molecular chaos. We have carried out a linear stability analysis of this state and studied the possible bifurcations issuing from it. By using systematically bifurcation theory for its Enskog kinetic equation, we have observed that the picture that flocking emerges from the bifurcation analysis is intricate. The amplitude equations for the bifurcation nodes near the critical value of noise are two coupled equations for a disturbance of the number density and a current density. We have shown that their solutions exhibit an interplay between parabolic and hyperbolic behavior in two different time scales when the distance to the critical value of the bifurcation control parameter goes to zero. In this limit we have found that there appear oscillation frequencies that give rise to resonance phenomena if the alignment rule contains a periodic function of time. We have confirmed the existence of these resonances with direct simulations of the VM.

Flocking in the modified Vicsek model

We have presented a *modified* Vicsek model in which particles can align their velocity with the local average direction or with the opposite direction. To understand the results, we have uses the same methodology to study the effect of modifying the probability density of the noise in the alignment rule by which VM particles change their velocities. The theoretical results obtained show new exotic phases arising in the behavior of the system and the equation for the complex amplitude of the bifurcating solution is a modified complex Ginzburg-Landau equation.

We have also observed how the novel exotic phases depend on the presence of contrarian compulsions. For small contrarian compulsions, there is a nonzero optimal noise value for which polarization is maximal and reflects a trend to homogeneous flocking as in the standard Vicsek model. However, when contrarian compulsions are prevalent, increasing the alignment noise may transform incoherent particle motion to a phase displaying period two polarization. If we relax the contrarian rule so that particles may select deflection with a large angle, we find that the flocking order parameter may oscillate periodically in time.

Tracking collective cell motion

The results obtained from the numerical studies of how epithelial cells aggregates, have shown an agreement with experiments from P. Silberzan's group. We have used a convenient model of active vertex and a dynamics for the cell centers that includes, forces that try to align their velocities to neighboring cells, friction with the substrate, inertia, and stochastic forces. We have simulated numerically this model in two different cases related to wound healing and to invasion of one cell collective by another one: (i) a cellular monolayer spreading on empty space, and (ii) the collision of two different cell populations in an antagonistic migration assay.

For (i), we have shown that inertia yields larger cells in the boundary than in the interior of the expanding tissue. We have also shown that fingers and cells located at the interface of the monolayer have larger area than cells in the middle of the monolayer and that finger cells oscillate. Both effects have been observed in experiments.

Concerning the antagonistic migration assays (ii), we have reproduced the collision of two monolayers from different populations. We have shown that cells can form different interfaces during the collision due to the fluid-like and solid-like behavior of the populations of cancerous and normal cells, respectively. To characterize the dynamics of the islands arising at the interface, we have introduced topological data analyses of experiments and numerical simulations. The overall analysis have allowed to us to extract parameter values that determine biologically relevant physical mechanisms for characterizing confluent motion of cellular aggregates.

Future work

For the discrete space-time Enskog equation studied in Chapter 2, the bifurcation solutions issuing from the disordered state depend on a small amplitude parameter that scales as the square root of the distance of the bifurcation parameter (e.g., alignment noise) to its critical value. The bifurcating states satisfy partial differential equations (PDEs) for slowly varying time and space scales. They still contain the small-amplitude parameter and their character depends on it: the bifurcation equations are hyperbolic for $\epsilon = 0$ and parabolic for $\epsilon \neq 0$. On short time scales, the terms proportional to can be ignored and the resulting hyperbolic equations may exhibit solutions that have a nontrivial space dependence and oscillate in time.

For sufficiently long times, the extra diffusive and source terms dampen the oscillations and make the solutions approach traveling waves. These phenomena will be explored in our future work.