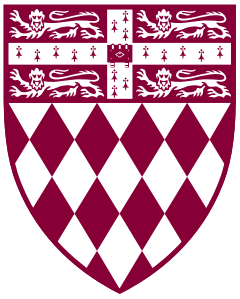
On the irreversibility of collective motion in polar active matter



Øyvind Lossius Borthne

Supervisor: Prof. Michael E. Cates

Department of Applied Mathematics and Theoretical Physics
University of Cambridge

This thesis is submitted for the degree of
Doctor of Philosophy

Fitzwilliam College

August 2021

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Summary and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Summary and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee

Øyvind Lossius Borthne
August 2021

Summary

Title: *On the irreversibility of collective motion in polar active matter*

- Chapter 1 – Introduction to the contents and main results of the thesis.
- Chapter 2 – Review of Itô and Stratonovich calculus, stochastic thermodynamics, (stochastic) entropy production, time-reversal symmetry, the principle of detailed balance, Sekimoto's stochastic energetics and spherical diffusions, with example applications.
- Chapter 3 – Review of the Vicsek model (VM), the 'continuous-time Vicsek model' (CTVM) and mean-field theory of the CTVM. This chapter also includes some unpublished results on the stability of the polar liquid phase of the CTVM in mean-field, in addition to the coarse-grained hydrodynamic theory of the CTVM. The work done in this chapter has been done under the supervision of Michael E. Cates.
- Chapter 4 – Unpublished analytical and numerical results on the 'Active XY model' (AXYM), including a linear stability analysis of the polar liquid phase of the AXYM and 'non-extensive AXYM' (NeAXYM). This work has been done under the supervision of Michael E. Cates. Sections 4.4 and 4.6 contain numerical investigations and a linear stability analysis of the coarse-grained hydrodynamic theory of the AXYM in $d = 2$, which has been published in paper 1. in the publications section, and has been done under the supervision of Michael E. Cates and Étienne Fodor.
- Chapter 5 – Unpublished results on the construction of the entropy production rate (EPR) in the CTVM, AXYM and NeAXYM, in addition to both numerical and analytical investigations of the EPR at high densities. The work in this chapter has been done under the supervision of Michael E. Cates and Étienne Fodor.
- Chapter 6 – Numerical and analytical investigation of the EPR in two hydrodynamic theories of polar active matter; the hydrodynamic Vicsek model (HVM) and diffusive flocking model (DFM). The relationship between the scaling of the EPR with noise

intensity and violation of time-reversal symmetry is investigated. The work in this chapter has been published in paper 1. in the publications section, and has been done under the supervision of Michael E. Cates and Étienne Fodor.

- Chapter 7 – Concluding remarks and summary the main findings in the thesis, including reflections and an outlook on potential future directions.

Throughout the thesis, all figures and tables have either been prepared for the dissertation or are taken from paper 1. in the publications section where not explicitly stated otherwise.

Øyvind Lossius Borthne

Acknowledgements

First and foremost, I would like to express my sincere gratitude to Prof. Michael E. Cates and Dr. Étienne Fodor, whose insight, encouragement and guidance has been indispensable, both to the scientific vision that has shaped my PhD, its successful realisation, as well as to the moral support and motivation which has assisted me along the way. Their expertise and pioneering ideas have been a continuous source of inspiration for aspiring to my own scientific ambitions, and their supervision has provided me with the necessary platform for achieving them.

I am also grateful to my friends and fellow coworkers in the Soft Matter group at DAMTP in Cambridge. Particularly Dr. Robert Jack and Dr. Ronojoy Adhikari, whom have shed light on and provided invaluable insight to different aspects of my work and the challenges I have faced along the way. I have also had the pleasure of working with fellow PhD students Fernando Caballero, Austen Bolitho, Irene Li, Jakub Dolezal, Timothy Ekeh, Günther Turk, Paul Rohrbach and Lukas Kikuchi, as well as postdocs Patrick Pietzonka, Julian Kappler, Rajesh Singh, Jules Guioth, Camille Scalliet and Joseph Peterson, all of whom have invited to, and initiated, intriguing scientific discussions, and perhaps most importantly, have contributed to a stimulating, enjoyable and vibrant social environment within the department.

Others outside of my academic circle have also been essential throughout my studies, both morally and financially. Without the loving support of my closest family, including Siri, who have patiently supported and motivated me throughout my studies, reaching this far would have seemed all the more unachievable. Also, I would like to express my deepest gratitude to Aker Scholarship for providing me with the financial resources necessary to pursue a PhD at Cambridge, and their support both in preparation for my studies as well as along the way.

Table of contents

List of figures

List of tables

Publications

Parts of this thesis have been published in:

1. Borthne Ø L, Fodor É, Cates M E,
   "Time-reversal symmetry violations and entropy production in field theories of polar active matter",
   New J. Phys. **22** 123012

Chapter 1

Introduction

Active matter comprises a vast range of systems out of thermodynamic equilibrium, for which the constituent particles undergo a self-propelled motion on the basis of a sustained transfer of energy with the environment [1–3]. They are present at all scales in the living world, and important examples are found in the cytoskeleton [4] and swimming microorganisms such as E. coli and myxobacteria [1], and range all the way through fish [5], birds [6] and human crowds [7]. Many experimental realisations, including self-catalytic colloids [8], self-propelled robots [9], and vibrated discs [10] and polar rods [11], have further facilitated systematic and intriguing investigations of synthetic active matter in the lab. From a theoretical point of view, one of the central problems in the study of active systems is to understand how one may unambiguously quantify their departure from equilibrium, and the extent to which their non-equilibrium nature on the microscopic scale imprints on the global dynamics [12–15].

By defying relaxation to equilibrium, many active systems display a novel phenomenology without any counterpart in inanimate systems. For example, in systems where the active constituents interact via alignment, a transition to a steady collectively moving (or flocking) phase may occur [1, 2, 16, 17]. In some well-known examples, including schools of fish and flocks of wildebeest or birds, the alignment arises as the result of some intelligent decision-making. This is not a requirement, however, and similar behaviour has also been reproduced in the lab in systems where the interaction is mediated by the solvent [11], as well as in suspensions of elongated active objects interacting via steric repulsion [1], illustrated in fig. 1.1. To the statistical physicist, this therefore provides hope that minimal models can capture the essential physics involved and uncover novel universal behaviour [18–20].

Perhaps more important to our agenda, however, is the presence of steady currents such as those seen in systems that undergo collective motion, since they represent a genuine signature of the departure from equilibrium. In particular, *time-reversal symmetry* (TRS) dictates that

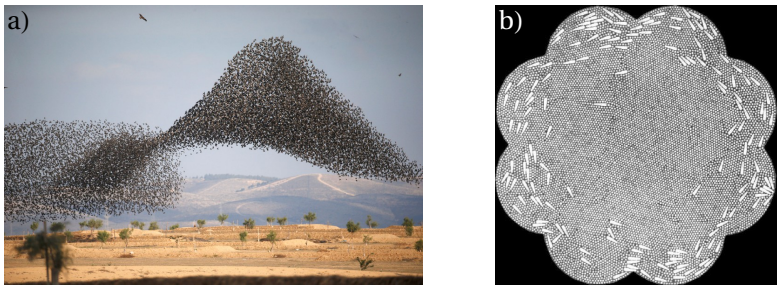


Fig. 1.1 Collective motion in biological and synthetic systems: a) starling murmuration and b) vibrated polar rods (figure taken from [11]).

the statistics of equilibrium dynamics played forward and backward in time must be the same, thus precluding the existence of steady-state currents [21–25]. One should note, however, that activity alone is not sufficient to guarantee the existence of global currents. For example, in suspensions of isotropic active particles interacting via volume exclusion, the steady-state dynamics can lead to so-called motility-induced phase separation (MIPS) [3, 26, 27], which on global scales may appear deceivingly similar to equilibrium phase separation [12]. Of course, the former may occur even in the absence of attractive interactions, which is not true for its equilibrium counterpart. Because of this, it is pertinent to ask also to what extent the activity really matters in such systems, and more specifically, when the micro-irreversibility can be detected from the global collective dynamics [28].

Building on the idea that non-equilibrium is characterised by the breaking of TRS, stochastic thermodynamics defines the *entropy production rate* (EPR) as a measure of the difference between the statistics of forward and reversed time fluctuations [25]. Importantly, even though stochastic thermodynamics was initially developed for 'thermodynamically consistent' systems for which heat can be meaningfully defined, the generality the EPR construction allows it to be considered also for systems where such concepts are not immediately appropriate. The EPR then remains a measure of irreversibility, albeit in an informatic sense, and can be utilised to define both equilibrium and the departure from it [12, 29–33]. To distinguish this concept from the thermodynamic EPR, however, on some occasions we will follow the convention of referring to it as an *informatic* EPR (IEPR) [28].

Conceptually, the EPR compares realisations of the dynamics played forward and reversed in time. However, in order to make such a comparison, it is necessary to define an appropriate protocol for time-reversal (see fig. 1.2). While some quantities, such as position and velocity,

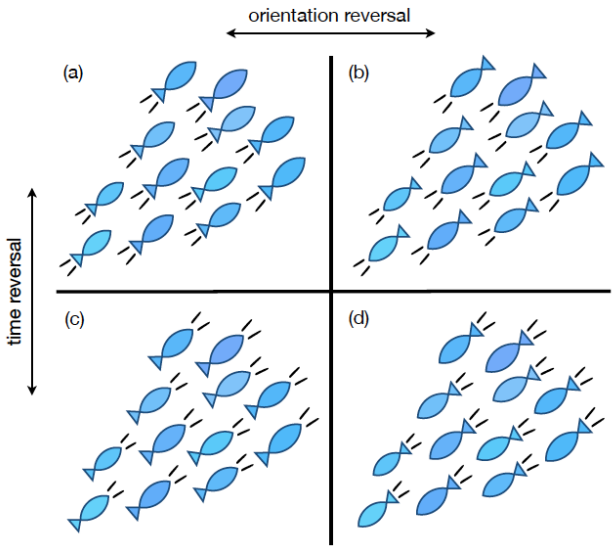


Fig. 1.2 Reprinted, with permission, from arXiv submission [28]. The figure illustrates four combinations of time and orientation reversals that can be considered, e.g. in a simplified planar model of a shoal of fish. The fish orientations are specified by their head-to-tail axes, and their swimming directions by the wake lines formed due to their motion. Panel (a) illustrates how the fish normally swim in forward time. On reversing time or head-to-tail orientations, but not both, fish swim tail first as in (b) or (c). Combining both time and and orientation reversal, all the fish swim in the opposite direction as illustrated in (d). For a single fish, (a) and (d) occur with equal probability, assuming rotational invariance so that the fish does not preferentially swim in any particular direction. For both a single fish *and* a shoal, (b) and (c) both occur with lower probability than (a). Consequently, the EPR of the single fish model will be 'small' when (d) is chosen for time-reversal, or 'large' if on instead chooses (b) or (c). For a shoal, assuming for sake of argument that it is denser at the front than the rear, (d) will also have a lower probability than (a) of occurring. Correspondingly, the EPR of the shoal model constructed using (d) for time-reversal should *also* have a 'large' EPR.

have fixed signatures under time-reversal, others, such as active propulsive forces that act along e.g. a head-to-tail axis, do not in general [34–36]. In thermodynamically consistent systems where the EPR can be connected to a heat flow, the thermodynamics restricts this choice since there can then only be one thermodynamic EPR [28]. On the other hand, for the systems we will see in this thesis, oftentimes several protocols can be utilised, providing different definitions of the EPR.

Motivated by the conceptual understanding of what the EPR should be – i.e. a measure of the difference between 'movies' of the dynamics played forward and backward in time – some of our work in this thesis, particularly pertaining to chapters 5 and 6, seeks to explore how this actually gets reflected quantitatively. Or, asked differently, what does it mean for

the EPR to be 'small' or 'large', and can this be interpreted in terms of time-forward and backwards realisations? We will see, in particular, that the collective behaviour and formation of global currents has important consequences for how the EPR scales with the strength of random fluctuations present on the hydrodynamic scale. In this sense, we can say that TRS is broken strongly or weakly, or that the dynamics are effectively equilibrium when the EPR respectively diverges, remains finite or approaches zero with vanishing noise.

To realise the limit of 'small hydrodynamic fluctuations' in microscopic models, we will consider dynamics where it is assumed that the particles can be treated as point-like, and take the limit of high density. In chapters 3 and 4 we therefore review the construction of continuous-time analogs of the much celebrated *Vicsek model* (VM) [17] of flocking active particles, and show that the mean-field theory (MFT) of the models are exact at infinite density. This allows us to perform exact calculations of the EPR in this limit, which can be compared with numerical results.

In summary, we will structure the thesis as follows. In chapter 2 we review stochastic thermodynamics more thoroughly, including the construction of the EPR, and introduce a toolbox that will be useful for the remainder of the thesis. We first apply this to analyse in chapter 3 a continuous-time formulation of the VM that we refer to as the *continuous-time Vicsek model* (CTVM). Then, in chapter 4, we will see that by only slightly modifying the alignment interaction of the CTVM, leading to a model we refer to as the *Active XY model* (AXYM), the resulting phase diagram is drastically changed, containing in particular a novel *polar cluster* (PC) phase. The EPRs of the two models are then compared in chapter 5, where it is found that the PC phase in the AXYM produces behaviour which is quite different from that of the homogeneous flocking seen in the CTVM. Finally, in chapter 6, we draw more general conclusions by considering from the outset two hydrodynamic theories of flocking, namely the *hydrodynamic Vicsek model* (HVM) and *diffusive flocking model* (DFM), before presenting our concluding remarks in chapter 7.

Chapter 2

Stochastic thermodynamics of non-equilibrium systems

Beginning with the pioneering work of Evans *et al.* [37], who discovered a peculiar symmetry in the distribution function of entropy production in sheared fluids, the subsequent development of **stochastic thermodynamics** represents a paradigm shift in modern non-equilibrium statistical physics. Perhaps most well known are the so-called fluctuation theorems, which express universal properties of the distribution functions of observable quantities such as entropy, heat and work in stochastic dynamics [25, 38]. These include the steady-state fluctuation theorem initially discovered by Evans, which was subsequently proven in much greater generality in the highly influential papers by Kurchan [39] and Lebowitz and Spohn [24], in addition to the celebrated Hummer-Szabo [40] and Jarzynski [41, 42] relations and Crooks fluctuation theorem [43, 44]. Still, these remain some of the only existing examples of truly general results of non-equilibrium statistical physics valid arbitrarily far from the equilibrium and linear response regimes. Other significant, and more recent, developments include the thermodynamic uncertainty relations, that quantify the trade off between precision and dissipation in non-equilibrium systems [45–55].

Central to stochastic thermodynamics is the concept of a **stochastic entropy production** [56], first introduced in the papers by Maes [57], Kurchan [39], and Lebowitz and Spohn [24], which can be consistently defined on the level of individual realisations of dynamics that are described by stochastic differential equations (SDEs, also referred to as *Langevin equations*). Importantly, this construction is valid even when the source of fluctuations cannot *a priori* be identified as thermal, or from a bath with a prescribed temperature. For such dynamics, the entropy production remains a measure of the non-equilibrium nature of the dynamics, and should be interpreted as quantifying the extent to which the dynamics violate time-reversal symmetry (TRS) [12, 31, 36, 58]. In this chapter, our primary agenda will be to review this

construction, which then sets the stage for discussing the irreversibility of collective motion in the following chapters.

First, in section 2.1, we highlight some results from stochastic analysis that will be used throughout the remainder of this thesis. Importantly, however, we do assume that the reader possesses some proficiency in stochastic calculus, and so for sake of brevity our focus will be limited. Then, in section 2.2, we introduce the entropy balance equation for stochastic dynamics that are described by SDEs. Here we identify the relevant generalised thermodynamic forces that drive probability currents, and are responsible for the entropy production in the system and reservoir. Furthermore, the entropy balance equation is closely related to time-reversal on the level of individual trajectories, and in section 2.3 we explore this relationship and show how entropy production can be related to a violation of the principle of detailed balance, which also leads us to define the stochastic entropy production. In section 2.4 we then further study the statistics of this quantity, and briefly sketch a much celebrated proof by Lebowitz and Spohn [24] of the steady-state fluctuation theorem. Finally, in sections 2.5 and 2.6 we discuss applications of both the stochastic analysis and the stochastic thermodynamics that are relevant to our work in the following chapters, before concluding in section 2.7.

2.1 Stochastic calculus of processes driven by white noise

Modern non-equilibrium statistical mechanics of stochastic dynamics relies to a great extent the mathematical analysis of differential equations driven by white noise. In light of this, we begin this chapter by recalling some preliminaries and central results from stochastic calculus that we employ frequently throughout the remainder of this thesis, which also serves the dual purpose of fixing our notation and conventions. However, in order to avoid digression our discussion is necessarily focused, implying that some background knowledge of the topic is assumed on the part of the reader. Specifically, our attention in this section will be centered on the various definitions of the stochastic integral, in addition to transformations of stochastic processes and their resulting laws. Most of the results we present here can be found in any standard textbook on stochastic calculus, e.g. [59–61], and we will only reference explicitly results that are taken from a specific resource.

The dynamics we study in this thesis are the solutions to stochastic differential equations (SDEs), or *Langevin equations*, that are driven by white noise. Conventionally, *standard* (or *unit*) Gaussian white noise is defined as the d-dimensional stochastic process $\boldsymbol{\eta}$ which

satisfies $\langle \eta_\alpha \rangle = 0$ and has autocorrelation

$$\left\langle \eta_\alpha(t)\eta_\beta(t')\right\rangle = \delta_{\alpha\beta}\delta(t-t'), \tag{2.1}$$

where $\langle\cdot\rangle$ denotes an average over realisations of $\boldsymbol{\eta}$. From the Gaussian property this completely fixes the statistics of $\boldsymbol{\eta}$, since all higher order moments can be determined from the expectation and autocorrelation.

Throughout, we formally treat the white noise process as the time-derivative of a standard Brownian motion (or Wiener process) $\boldsymbol{W}$, so that integrating it gives

$$\boldsymbol{W}(\tau) = \int_0^\tau \mathrm{d}t\, \boldsymbol{\eta}(t). \tag{2.2}$$

We will need to take care, however, when integrating $\boldsymbol{\eta}$ against other processes. More precisely, for a $d \times d'$-matrix valued process $b(t) \equiv \left(b_{\alpha\beta}(t)\right)$, the *Itô integral* of b against the d'-dimensional unit white noise $\boldsymbol{\eta}$ is defined as

$$\int_0^\tau \mathrm{d}t\, b\boldsymbol{\eta} \equiv \lim_{N\to\infty} \sum_{n=0}^{N-1} b(t_n)\left(\boldsymbol{W}(t_{n+1}) - \boldsymbol{W}(t_n)\right), \tag{2.3}$$

where $t_n = n\Delta t$, and $\Delta t = t/N$.

The fact that we discretise the integral in eq. (2.3) using a 'left Riemann-Stieltjes sum', with b evaluated at t_n, turns out to be crucial when the limit is taken. For example, the *Stratonovich integral*

$$\int_0^\tau \mathrm{d}t\, b\circ\boldsymbol{\eta} \equiv \lim_{N\to\infty} \sum_{n=0}^{N-1} \frac{b(t_n)+b(t_{n+1})}{2}\left(\boldsymbol{W}(t_{n+1}) - \boldsymbol{W}(t_n)\right), \tag{2.4}$$

which uses mid-point discretisation (trapezoidal rule) will in general yield a different result. More generally, there clearly are an infinite number of ways to discretise the stochastic integral, and all yield distinct definitions. Despite this, for us it will be sufficient to consider only the Itô and Stratonovich choices. To distinguish these we indicate that we are employing Stratonovich discretisation by $\circ$ as in eq. (2.4), also when there is no explicit integration performed.

Equipped with the definitions in eqs. (2.3) and (2.4), we say that a d-dimensional process $\boldsymbol{X}$ is a solution to the SDE

$$\dot{\boldsymbol{X}} = \boldsymbol{a}(\boldsymbol{X},t) + b(\boldsymbol{X},t)\boldsymbol{\eta}, \tag{2.5}$$

driven by d'-dimensional white noise $\boldsymbol{\eta}$, if

$$\boldsymbol{X}(\tau) = \boldsymbol{X}(0) + \int_0^\tau \mathrm{d}t\, \boldsymbol{a}(\boldsymbol{X},t) + \int_0^\tau \mathrm{d}t\, b(\boldsymbol{X},t)\boldsymbol{\eta}, \tag{2.6}$$

where $\boldsymbol{a}$ is a d-dimensional *drift term* and $b = \left(b_{\alpha\beta}\right)$ is a $d \times d'$-matrix valued *diffusion coefficient*. Of course, a similar definition can be used for Stratonovich SDEs, although as we will see momentarily, transforming between these is straightforward so a second definition is superfluous. Thus, from eq. (2.6), solutions to the SDE in eq. (2.5) may be obtained at least formally via e.g. standard Picard iterations.

Mathematically, the reason the Itô and Stratonovich choices give different results is that the sample paths of Brownian motion are very 'rough'. The standard way to quantify this roughness is via the *quadratic covariation* process, which measures the total squared variation of the process over a mesh that tends to zero. As we shall see throughout this thesis, the quadratic covariation will prove a very useful tool when we seek to transform between Itô and Stratonovich integrals. Since we are interested in this merely from a computational perspective, we define the quadratic variation $[\cdot,\cdot]$ in this thesis (in the non-standard way) for processes $\boldsymbol{X}$ that solve the SDE in eq. (2.5) as the symmetric bilinear form which satisfies[1]

$$\begin{cases} \left[X_\alpha, X_\beta\right] = 0, \\ \left[X_\alpha, \dot{X}_\beta\right] = 0, \\ \left[\dot{X}_\alpha, \dot{X}_\beta\right] = B_{\alpha\beta}, \end{cases} \tag{2.7}$$

where $B = bb^T$. In particular, the quadratic covariation only cares about the diffusion term of the SDE (the 'rough part') and vanishes for terms that are of so-called finite total variation – and are more well-behaved.

For two processes X and Y of finite quadratic covariation, one may show that the following very useful transformation property between the Itô and Stratonovich integrals holds:

$$X \circ \dot{Y} = X\dot{Y} + \frac{1}{2}\left[\dot{X}, \dot{Y}\right]. \tag{2.8}$$

Because of this, we will refer to the term $\left[\dot{X}, \dot{Y}\right]/2$ as a *spurious drift*. Typically in applications we encounter in this thesis, Y will be some complicated function of X or vice versa when we wish to transform between the two integrals. However, as we shall see, by using only a few properties of the quadratic covariation the spurious drift can fairly straightforwardly be

[1] We are really considering the *time-derivative* of the quadratic covariation the way it is normally defined in the mathematical literature. However, since we only use it as a computational tool, this difference will not be relevant for us.

calculated. Note also that from the bilinearity of the quadratic covariation, eq. (2.8) extends trivially to vector valued processes.

First, the *Kunita-Watanabe identity* (see e.g. theorem 28.1 in [62]) states that the quadratic covariation does not care about multiplicative factors that are sufficiently regular. Explicitly, for a function f and stochastic process $\boldsymbol{X}$ of finite quadratic covariation,

$$\left[f(\boldsymbol{X})\dot{X}_\alpha, \dot{X}_\beta\right] = f(\boldsymbol{X})\left[\dot{X}_\alpha, \dot{X}_\beta\right]. \tag{2.9}$$

For processes $\boldsymbol{X}$ that solve the SDE in eq. (2.5), this can be shown straightforwardly from eq. (2.7) by setting $\dot{\boldsymbol{Y}} = \left(\dot{\boldsymbol{X}}, f(\boldsymbol{X})\dot{\boldsymbol{X}}\right)^T$.

To motivate the second property we will first need to introduce *Itô's lemma*, which extends the chain rule from ordinary calculus to transformations of stochastic processes driven by white noise. This will also prove useful for other computations we perform, which further motivates introducing it here. For a function f and d-dimensional stochastic process $\boldsymbol{X}$ of finite quadratic covariation, this states that

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\boldsymbol{X}) = \dot{\boldsymbol{X}} \cdot \nabla f(\boldsymbol{X}) + \frac{1}{2}\left[\dot{X}_\alpha, \dot{X}_\beta\right] \nabla_\alpha \nabla_\beta f(\boldsymbol{X}). \tag{2.10}$$

Thus, if $\boldsymbol{X}$ satisfies an SDE of the form of eq. (2.5), and we set $\boldsymbol{Y} = \boldsymbol{f}(\boldsymbol{X})$ for some d'-dimensional function $\boldsymbol{f}$, then it follows that

$$\begin{cases} \left[Y_\alpha, Y_\beta\right] = 0, \\ \left[Y_\alpha, \dot{Y}_\beta\right] = 0, \\ \left[\dot{Y}_\alpha, \dot{Y}_\beta\right] = \left(\nabla_\gamma f_\alpha\right)\left(\nabla_\delta f_\beta\right)\left[\dot{X}_\gamma, \dot{X}_\delta\right], \end{cases} \tag{2.11}$$

To derive this, one simply has to apply Itô's lemma to $\boldsymbol{Z} = (\boldsymbol{X}, \boldsymbol{f}(\boldsymbol{X}))^T$ and the definition of the quadratic covariation in eq. (2.7).

From Itô's lemma in eq. (2.10) and the transformation identity in eq. (2.8), a particularly appealing property of Stratonovich calculus can be derived. Namely, that it satisfies the ordinary chain rule. By this we mean that for a function f and stochastic process $\boldsymbol{X}$,

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\boldsymbol{X}) = \dot{\boldsymbol{X}} \circ \nabla f(\boldsymbol{X}). \tag{2.12}$$

Note that in eq. (2.12) we use the convention that when $\circ$ appears between two vector quantities (here $\dot{\boldsymbol{X}}$ and ∇) an inner (scalar) product is implied, which we also assume throughout this thesis.

A third and final property related to the quadratic covariation that we mention here is the *Lévy characterisation of Brownian motion*. As the name suggests, this states that $\boldsymbol{X}$ is a standard Brownian motion if and only if $\langle X_\alpha \rangle = 0$ and

$$\left[\dot{X}_\alpha, \dot{X}_\beta\right] = \delta_{\alpha\beta}. \tag{2.13}$$

As we shall see, this allows us to determine when certain transformations of white noise remain white.

Having introduced the quadratic covariation, we remark that one may of course perform the transformation between the Itô and Stratonovich integrals by appealing to an explicit transformation identity. However, we find it is much easier to work with the quadratic covariation, especially when the transformations become layered and involved. For example, for functions f and g and a stochastic process X, eqs. (2.8) and (2.11) immediately imply that for $Y = g(X)$

$$f(X) \circ \dot{Y} = f(X)\dot{Y} + \frac{1}{2} f'(X) g'(X) \left[\dot{X}\right], \tag{2.14}$$

where we denote by $\left[\dot{X}\right] \equiv \left[\dot{X}, \dot{X}\right]$. Similarly, for a 1-dimensional SDE of Stratonovich form

$$\dot{X} = a(X) + b(X) \circ \eta, \tag{2.15}$$

since $\left[\frac{\mathrm{d}}{\mathrm{d}t} b(X), \eta\right] = b' \left[\dot{X}, \eta\right] = b'b$, it can be written in Itô form as

$$\dot{X} = a(X) + \frac{1}{2} b'(X) b(X) + b(X)\eta. \tag{2.16}$$

A dynamics that solves the SDE in eq. (2.5) can also be described by its corresponding *law* (or *distribution*) $p(\boldsymbol{x}, t)$, defined as the probability of observing the process $\boldsymbol{X}$ at position $\boldsymbol{x} \in \mathcal{V}$ at time t, where we denote by $\mathcal{V}$ the volume within which the process $\boldsymbol{X}$ lives. The distribution p can be obtained as the solution to the *Fokker-Planck equation* (FPE)

$$\partial_t p = \mathcal{L}^\dagger p. \tag{2.17}$$

The linear operator $\mathcal{L}^\dagger$ in eq. (2.17) is the L^2-adjoint of the *infinitesimal generator* $\mathcal{L}$. For a dynamics that solves the SDE in eq. (2.5), the infinitesimal generator is given explicitly by

$$\mathcal{L} = \boldsymbol{a} \cdot \nabla + \frac{1}{2} B_{\alpha\beta} \nabla_\alpha \nabla_\beta, \tag{2.18}$$

where $B = bb^T$ as before. Note that in eq. (2.18) we apply the Einstein summation convention to repeated Greek indices $\alpha, \beta, \ldots$, which we will also do throughout this thesis. Importantly,

the infinitesimal generator fully characterises the dynamics in eq. (2.5), meaning that there is a one-to-one correspondence between $\mathcal{L}$ and the SDE.

To make the FPE well-posed, we have to specify both an initial distribution $p_0(\boldsymbol{x})$, which may be a point mass centered at some $\boldsymbol{x}_0 \in \mathcal{V}$, in addition to appropriate boundary conditions. Throughout this chapter, we will assume for simplicity that the boundary $\partial\mathcal{V}$ is either periodic or that $\mathcal{V}$ is infinite. Regardless, one always assumes that the *probability current* $\boldsymbol{j}$ defined by

$$\boldsymbol{j} = \boldsymbol{a}p - \frac{1}{2}\nabla \cdot (Bp) \tag{2.19}$$

integrates to zero over $\partial\mathcal{V}$ so that there is no net flux of probability mass across $\partial\mathcal{V}$, i.e.

$$\int_{\partial\mathcal{V}} \boldsymbol{j} \cdot \mathrm{d}\boldsymbol{S} = 0, \tag{2.20}$$

where $\mathrm{d}\boldsymbol{S}$ is the surface measure on $\partial\mathcal{V}$. From the continuity equation

$$\partial_t p = -\nabla \cdot \boldsymbol{j}, \tag{2.21}$$

which is equivalent to the FPE, this ensures that the probability density p is conserved as it should.

To simplify some calculations in the following, we also make the assumption that B is divergence free, meaning that $\nabla_\alpha B_{\alpha\beta} = 0$, which allows us to write

$$\boldsymbol{j} = \boldsymbol{a}p - \frac{1}{2}B\nabla p, \tag{2.22}$$

although most results we present here generalise straightforwardly also when $\nabla_\alpha B_{\alpha\beta} \neq 0$. Finally, we remark that the steady-state solution to the FPE will be denoted by p_{ss}, and satisfies $\partial_t p_{\mathrm{ss}} = 0$, or equivalently $\nabla \cdot \boldsymbol{j}_{\mathrm{ss}} = 0$, where $\boldsymbol{j}_{\mathrm{ss}} = \boldsymbol{a}p_{\mathrm{ss}} - B\nabla p_{\mathrm{ss}}/2$ is the steady-state probability current.

With this we conclude our introduction to the stochastic analysis we will utilise in the remainder of this thesis. In the following sections of this chapter, we will aim to build a consistent *dynamical* statistical mechanics of solutions to the SDE in eq. (2.5). This will introduce the entropy production as a natural measure of the irreversibility of stochastic dynamics, which we will investigate for polar active systems in later parts of this thesis.

2.2 Statistical mechanics of stochastic dynamics

Throughout the following sections, we will aim to provide a brief introduction to the dynamical statistical mechanics referred to as **stochastic thermodynamics**, which has seen a cascade of breakthrough developments occur within the past two or three decades [25, 38]. In essence, from the viewpoints of Sekimoto [21] and Seifert [56], the theory takes seriously both the first and second laws of thermodynamics on the level of individual realisations of stochastic dynamics. These may, for example, solve the SDE in eq. (2.5) (or a stochastic *partial* differential equation [38]), or obey some Markovian dynamics with discrete jumps driven by Poisson noise [63]. Both of these descriptions allow for the separation of the slow degrees of freedom of the system and the fast, noisy, action of one or more reservoirs. Importantly, they also facilitate the description of systems that are driven arbitrarily far from equilibrium by external forces or fluxes of e.g. thermal or chemical energy. Because of this, novel results that relate the classical thermodynamic quantities of entropy, heat and work have been obtained also in the regime where linear response to equilibrium breaks down [40–44]. Applications of stochastic thermodynamics are vast, and it has been successfully employed in many studies of biological processes and biomolecules [64], including molecular machines [65–67], enzymes [68, 69] and colloidal systems [29, 34, 70].

Among the most important developments are the fluctuation theorems, that "express universal properties of the probability distribution of trajectory level observables such as heat and work" [25]. The first of these was observed as a particular symmetry in the probability distribution of entropy production in computer simulations of sheared fluids by Evans *et al.* [37], and subsequently proven rigorously for chaotic dynamics by Gallavotti and Cohen [71]. However, it was not until the seminal contributions by Maes [57], Kurchan [39], and Lebowitz and Spohn [24] that the significant generality of this result became more clear, when it was proven also for the stochastic dynamics of SDEs and Markov chains, and it was realised that the entropy production could be obtained by considering the difference between path probabilities of forward and reversed time dynamics. Others, such as the Jarzynski relation [41, 42] and Crooks fluctuation theorem [43, 44], relate free-energy differences between the initial and final states of a system to the probability distribution of the work required to transform between them.

Another very interesting and more recent avenue of research concerns the thermodynamic uncertainty relations, that quantify the trade off between precision and dissipation, or entropy production, in systems that are arbitrarily far from equilibrium [45–47]. These have, amongst other things, been used to derive bounds on the efficiency of molecular motors [65, 66], which have also been compared with experimental data of kinesin motors that convert ATP to mechanical power [67]. For our purposes in the remainder of this thesis, however, the

thermodynamic uncertainty relations will not be directly relevant, and so we will not discuss them in any further detail.

Our presentation in the following will again necessarily be restricted. Moreover, since the models that we deal with in this thesis are influenced by fluctuations that are not obviously thermal, some of the language from thermodynamics will not be appropriate to our setting. Because of this, we grant the *entropy production rate* (EPR) (also referred to as *informatic* EPR (IEPR) [28]) centre stage, which, as we shall see, can meaningfully be defined via an entropy balance equation for all systems described by the SDE in eq. (2.5). The EPR should then be interpreted as an informatic measure of irreversibility since it quantifies, in an appropriate sense, the difference between the statistics of the forward and reversed time dynamics, although it cannot in general be connected to a heat flow. In section 2.6, we will see that this formalism reduces to the familiar one of Sekimoto's stochastic energetics [21] for an equilibrium colloid suspended in a viscous fluid, and also how the theory works out for an example non-equilibrium system where energy and heat currents cannot be usefully defined.

We begin in this section by introducing the entropy balance equation for dynamics that obey eq. (2.5). Then, in section 2.3, we introduce the trajectory level formulation of the stochastic dynamics via the path ensemble and action, and show that equilibrium can be defined by requiring that the principle of detailed balance holds. We progress from this by introducing the stochastic EPR, which provides us with the trajectory level definition of entropy production that we study for the remainder of this thesis. Finally, in sections 2.5 and 2.6 we look at applications of both the stochastic analysis from section 2.1 and the stochastic thermodynamics that will be useful to us in the following chapters.

In the following, we label the degrees of freedom of our system by $\boldsymbol{X}$, which usually contains both positions and momenta of the particles in the system (not the reservoir), and assume that they evolve according to an SDE of the form of eq. (2.5). We further assume that particles are restricted to live in some phase space volume $\mathcal{V}$, which is either infinite or has periodic boundaries. In line with our introduction to stochastic calculus in section 2.1, we denote the probability (density) of observing $\boldsymbol{X}$ in the configuration $\boldsymbol{x}$ at time t by $p(\boldsymbol{x},t)$.

From these definitions, we introduce the *system* (*Shannon*) *entropy* $\mathcal{S}^{\mathrm{sys}}$ in the standard way as

$$\mathcal{S}^{\mathrm{sys}} = -\int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, p(\boldsymbol{x},t) \log\left(h p(\boldsymbol{x},t)\right), \tag{2.23}$$

where we take the Boltzmann constant $k_B = 1$ and h has dimensions of a phase space volume element. To build a dynamical theory, we begin by investigating the rate of change of $\mathcal{S}^{\mathrm{sys}}$. To determine this, we differentiate eq. (2.23) with respect to time and use the fact that p is conserved (cf. eqs. (2.20) and (2.21)). Subsequent to an integration by parts, where the

boundary term vanishes, we immediately obtain

$$\dot{\mathcal{S}}^{\mathrm{sys}} = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \boldsymbol{j} \cdot \boldsymbol{F}^{\mathcal{S}}. \tag{2.24}$$

In eq. (2.24), we have defined the *entropic thermodynamic force*

$$\boldsymbol{F}^{\mathcal{S}} = -\nabla \log(hp), \tag{2.25}$$

responsible for driving p towards increasing values of the system entropy $\mathcal{S}^{\mathrm{sys}}$. One should note here that $\boldsymbol{F}^{\mathcal{S}}$ is of course not a mechanical force, even though we employ this terminology – instead, it acts on the space of probability distributions. Notice also that $\boldsymbol{F}^{\mathcal{S}}$ is 'conservative', and as we shall see it performs no 'work' on $\boldsymbol{j}$. More specifically, we will see that in steady-state, $\boldsymbol{j}_{\mathrm{ss}} \cdot \boldsymbol{F}^{\mathcal{S}}_{\mathrm{ss}} = 0$, where $\boldsymbol{F}^{\mathcal{S}}_{\mathrm{ss}} = -\nabla \log(hp_{\mathrm{ss}})$, and so the system entropy production $\dot{\mathcal{S}}^{\mathrm{sys}}$ vanishes in steady-state as also follows from eq. (2.23).

When the dynamics solve the SDE in eq. (2.5), and $B(\boldsymbol{x},t) = bb^T$ is invertible for all $\boldsymbol{x} \in \mathcal{V}$ and t, $\boldsymbol{F}^{\mathcal{S}}$ enters in the fundamental decomposition of thermodynamic forces

$$\boldsymbol{F} = \boldsymbol{F}^{\mathcal{S}} + \boldsymbol{F}^{\mathcal{Q}}, \tag{2.26}$$

which follows immediately from eq. (2.22) for $\boldsymbol{j}$. Explicitly, in eq. (2.26), $\boldsymbol{F}$ is the total force responsible for driving the current $\boldsymbol{j}$, which is defined by the linear Onsager relation

$$\boldsymbol{j} = L\boldsymbol{F}, \tag{2.27}$$

where the Onsager matrix $L \equiv Bp/2$. Furthermore, $\boldsymbol{F}^{\mathcal{Q}}$ is defined by

$$\boldsymbol{F}^{\mathcal{Q}} \equiv 2B^{-1}\boldsymbol{a}, \tag{2.28}$$

and is the thermodynamic force associated with the flow of energy between the system and the reservoir. In particular, when fluctuations are thermal, we will see in section 2.6 that one may identify the entropy produced by $\boldsymbol{F}^{\mathcal{Q}}$, i.e.

$$\dot{\mathcal{S}}^{\mathrm{res}} = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \boldsymbol{j} \cdot \boldsymbol{F}^{\mathcal{Q}} \tag{2.29}$$

with the entropy production in the reservoir due to heat dissipation. Notice also that $\boldsymbol{F}^{\mathcal{Q}}$ is independent of the distribution p.

Combining eqs. (2.24) and (2.26), it thus follows that we have the *entropy balance equation* [24, 25]

$$\dot{\mathcal{S}} = \dot{\mathcal{S}}^{\text{sys}} + \dot{\mathcal{S}}^{\text{res}}, \tag{2.30}$$

where the total EPR $\dot{\mathcal{S}}$ is given by $\dot{\mathcal{S}} = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \boldsymbol{j} \cdot \boldsymbol{F}$, or equivalently

$$\dot{\mathcal{S}} = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \boldsymbol{j}^T L^{-1} \boldsymbol{j}. \tag{2.31}$$

Our objective in the following sections will be to demonstrate that this expression for $\dot{\mathcal{S}}$ indeed makes sense as the physical total entropy production rate. For now, however, we observe that since $B = bb^T$ is positive definite it follows that

$$\dot{\mathcal{S}} \geq 0, \tag{2.32}$$

with equality holding if and only if $j = 0$. Because of this, we define the equilibrium distribution, if it exists, as the steady-state distribution p_{ss} which solves $j_{\text{ss}} = 0$. Thus, in equilibrium $\boldsymbol{F}$ vanishes and the entropic force and the force due to energy flow between the system and reservoir exactly balance, i.e.

$$\boldsymbol{F}^{\mathcal{S}}_{\text{ss}} = -\boldsymbol{F}^{\mathcal{Q}} \quad \text{in equilibrium.} \tag{2.33}$$

In the next section, we will will proceed by introducing the trajectory level description of the stochastic dynamics via a path ensemble and action. This allows us to discuss the probabilities of individual realisations of the dynamics, and our agenda will be to see how time-reversal can be used at this level to connect the entropy balance equation in eq. (2.30) with a violation of the principle of detailed balance.

2.3 Time-reversal and the principle of detailed balance

A major achievement of stochastic thermodynamics has been to successfully connect the entropy production of the stochastic dynamics at the level of individual realisations with the corresponding deterministic description in terms of the probability distribution over configurations $p(\boldsymbol{x},t)$ introduced in the previous section. Here, we explore this relationship, and show that there is an intimate link between the entropy balance equation in eq. (2.30) and time-reversal at the trajectory level. To do this, we first introduce the path integral formalism of the dynamics described by the SDE in eq. (2.5), which allows us to construct the

distribution over trajectories for the time-reversed dynamics. Then, we show that equilibrium is characterised by a path level principle of detailed balance.

In the following, we let $p(\boldsymbol{x},\tau|\boldsymbol{x}_0,-\tau)$ be the *transition probability density* which solves the FPE in eq. (2.17) with initial condition $p_0(\boldsymbol{x})=\delta(\boldsymbol{x}-\boldsymbol{x}_0)$ started at time $t=-\tau$. The path integral formulation of the stochastic dynamics expresses the transition probability density p as a functional integral over a *path* (*transition*) *probability density* $\mathcal{P}$, so that

$$p(\boldsymbol{x},\tau|\boldsymbol{x}_0,-\tau)=\int_{\boldsymbol{x}_0}^{\boldsymbol{x}}\mathcal{D}[\boldsymbol{y}]\,\mathcal{P}([\boldsymbol{y}]|\boldsymbol{x}_0)\,. \tag{2.34}$$

In eq. (2.34), the path integral is performed over trajectories $\boldsymbol{y}(t)$ connecting $\boldsymbol{y}(-\tau)=\boldsymbol{x}_0$ and $\boldsymbol{y}(\tau)=\boldsymbol{x}$ in a time 2τ. Formally, we consider $\mathcal{P}$ to be defined as the continuum limit

$$\mathcal{P}\left([\boldsymbol{y}]|\boldsymbol{x}_0\right)=\lim_{N\to\infty}\prod_{n=1}^{N}p(\boldsymbol{y}_n,\tau_n|\boldsymbol{y}_{n-1},\tau_{n-1}), \tag{2.35}$$

where $\boldsymbol{y}_n=\boldsymbol{y}(\tau_n)$, $\tau_n=n\Delta\tau$ and $\Delta\tau=2\tau/N$. An explicit derivation of eq. (2.34) can then be performed by expanding $p(\boldsymbol{y}_n,\tau_n|\boldsymbol{y}_{n-1},\tau_{n-1})$ in small $\Delta\tau$, and we refer the reader to e.g. [72] for further details.

Importantly, all that we require for our agenda in the following is the result which states that unit Gaussian white noise $\boldsymbol{\eta}$ has the path probability density [73]

$$\mathcal{P}[\boldsymbol{\eta}]\propto\exp\left(-\frac{1}{2}\int_{-\tau}^{\tau}\mathrm{d}t\,\eta^2\right). \tag{2.36}$$

From this, one may argue that by inverting the SDE in eq. (2.5) to write the noise $\boldsymbol{\eta}$ in terms of the dynamics $\boldsymbol{X}$, we can obtain the path probability density for $\boldsymbol{X}$ by substitution back into eq. (2.36) provided the diffusion coefficient b is invertible. Thus, at least heuristically we should have that when $\boldsymbol{x}(t)$ is a typical sample path,

$$\mathcal{P}([\boldsymbol{x}]|\boldsymbol{x}_0)\propto\exp\left(-\frac{1}{2}\int_{-\tau}^{\tau}\mathrm{d}t\;(\dot{\boldsymbol{x}}-\boldsymbol{a})^T\,B^{-1}\,(\dot{\boldsymbol{x}}-\boldsymbol{a})\right). \tag{2.37}$$

Of course, this argument ignores the Jacobian of the transformation which could potentially be non-trivial. In fact, it turns out that the Jacobian depends on the particular time-discretisation employed when deriving the path integral representation (which may be chosen independently from the discretisation of the SDE), and for the 'time-symmetric' discretisation choice it is given by

$$J=\frac{1}{\sqrt{\det B}}\exp\left(-\frac{1}{2}\int_{-\tau}^{\tau}\mathrm{d}t\,\nabla\cdot\boldsymbol{a}\right), \tag{2.38}$$

which we adopt in the following. However, despite the contribution of J to $\mathcal{P}$, we choose to define the action A on the basis of eq. (2.37) for later convenience, explicitly

$$A = \frac{1}{2}\int_{-\tau}^{\tau} \mathrm{d}t\, (\dot{\boldsymbol{x}} - \boldsymbol{a})^T B^{-1} (\dot{\boldsymbol{x}} - \boldsymbol{a}), \tag{2.39}$$

which we will refer to as the *Freidlin-Wentzell action* (FWA) [74].[2]

Central to what we will discuss in the following is the discrete T-symmetry (or *time-reversal*), defined simply as a reversal of time, i.e.

$$\mathrm{T} : \boldsymbol{X}(t) \mapsto \boldsymbol{X}(-t), \tag{2.40}$$

and we will denote by $\overleftarrow{\boldsymbol{X}} = \mathrm{T}\boldsymbol{X}$ the corresponding time-reversed trajectory. For stochastic dynamics $\boldsymbol{X}$, we will be interested in exploring the statistics of the time-reversal $\overleftarrow{\boldsymbol{X}}$. By investigating this, we will be able to derive both an SDE for the time-reversed dynamics, in addition to the principle of detailed balance. As we shall see, the principle of detailed balance states that in equilibrium when $j_{\mathrm{ss}} = 0$, there is no net flux of probability mass between any pair of states. Consequently, as we will see, the dynamics $\boldsymbol{X}$ and $\overleftarrow{\boldsymbol{X}}$ are statistically equivalent in equilibrium.

Our treatment in the following is inspired in particular by the excellent review of macroscopic fluctuation theory by Bertini *et al.* in [75], where a similar calculation is performed for certain hydrodynamic driven diffusive systems. Following their arguments, we first note that since the stochastic process $\boldsymbol{X}$ is described by the SDE in eq. (2.5), it is also a *Markov process*. This has the property that the time-reversed process $\overleftarrow{\boldsymbol{X}}$ is also a Markov process, with the same stationary distribution p_{ss} as the forward-time process $\boldsymbol{X}$. Assuming then that this process also admits a description in terms of an SDE of the form

$$\dot{\overleftarrow{\boldsymbol{X}}} = \boldsymbol{a}^* + b\boldsymbol{\eta}, \tag{2.41}$$

we will seek to determine the drift $\boldsymbol{a}^*$ for $\overleftarrow{\boldsymbol{X}}$. To motivate this, note that the assumption in eq. (2.41) is equivalent to the statement that we may reverse the dynamics of $\boldsymbol{X}$ by applying an appropriate external 'force' so that $\boldsymbol{a}^* = \boldsymbol{a} + \Delta\boldsymbol{a}$ [75].

[2] Although the procedure we have sketched to arrive at the path integral formalism is that of the standard 'Onsager-Machlup' approach, we choose to refer to A as the 'Freidlin-Wentzell action' because we have discarded the contribution from J. The calculations we perform, however, are valid beyond the small noise regime within which Freidlin-Wentzell theory is formulated.

Now we let p_0 be some arbitrary initial probability distribution and define the probability $\mathfrak{p}[\boldsymbol{x}]$ of observing the sample path $\boldsymbol{x}(t)$ of the forward time dynamics in eq. (2.5) as

$$\mathfrak{p}[\boldsymbol{x}] = p_0(\boldsymbol{x}_0)\mathcal{P}([\boldsymbol{x}]|\boldsymbol{x}_0)\,. \tag{2.42}$$

Correspondingly, given our assumption in eq. (2.41), we have a distribution $\mathfrak{p}^*[\boldsymbol{x}]$ for the time-reversed process $\overleftarrow{\boldsymbol{X}}$ given by

$$\mathfrak{p}^*[\boldsymbol{x}] = p_0(\boldsymbol{x}_0)\mathcal{P}^*([\boldsymbol{x}]|\boldsymbol{x}_0)\,. \tag{2.43}$$

As before, $\mathcal{P}^* \propto \mathrm{e}^{-A^*}$, and the Jacobian J^* and FWA A^* follow for the time-reversed dynamics from eq. (2.41), and can be written

$$A^* = \frac{1}{2}\int_{-\tau}^{\tau} \mathrm{d}t\ (\dot{\boldsymbol{x}} - \boldsymbol{a}^*)^T B^{-1}\,(\dot{\boldsymbol{x}} - \boldsymbol{a}^*)\,, \tag{2.44}$$

and

$$J^* = \frac{1}{\sqrt{\det B}} \exp\left(-\frac{1}{2}\int_{-\tau}^{\tau} \mathrm{d}t\, \nabla\cdot\boldsymbol{a}^*\right). \tag{2.45}$$

Furthermore, we will denote by $\mathfrak{p}_{\mathrm{ss}}[\boldsymbol{x}]$ and $\mathfrak{p}^*_{\mathrm{ss}}[\boldsymbol{x}]$ the probabilities of observing $\boldsymbol{x}(t)$ under the forward and time-reversed dynamics started in steady-state, respectively. Observe then, that since $\overleftarrow{\boldsymbol{X}}$ is the time-reversal of $\boldsymbol{X}$, we must have

$$\overleftarrow{\mathfrak{p}}_{\mathrm{ss}}[\boldsymbol{x}] = \mathfrak{p}^*_{\mathrm{ss}}[\boldsymbol{x}], \tag{2.46}$$

where we have defined $\overleftarrow{\mathfrak{p}}_{\mathrm{ss}}[\boldsymbol{x}] \equiv \mathfrak{p}_{\mathrm{ss}}[\overleftarrow{\boldsymbol{x}}]$.

From eq. (2.46) the drift term $\boldsymbol{a}^*$ for the time-reversed process follows readily. To see this, we simply rewrite eq. (2.46) using eqs. (2.42) and (2.43) to obtain

$$\log\frac{p_{\mathrm{ss}}(\boldsymbol{x}(\tau))}{p_{\mathrm{ss}}(\boldsymbol{x}(-\tau))} = \log\frac{\mathcal{P}^*([\boldsymbol{x}]|\boldsymbol{x}(-\tau))}{\mathcal{P}([\overleftarrow{\boldsymbol{x}}]|\boldsymbol{x}(\tau))}. \tag{2.47}$$

Moreover, computing the right-hand side of eq. (2.47) from eqs. (2.38), (2.39), (2.44) and (2.45) gives

$$\begin{aligned}\log\frac{\mathcal{P}^*([\boldsymbol{x}]|\boldsymbol{x}(-\tau))}{\mathcal{P}([\overleftarrow{\boldsymbol{x}}]|\boldsymbol{x}(\tau))} = &\int_{-\tau}^{\tau} \mathrm{d}t\ (\boldsymbol{a}+\boldsymbol{a}^*)^T B^{-1}\circ\dot{\boldsymbol{x}} + \frac{1}{2}\int_{-\tau}^{\tau}\mathrm{d}t\ (\boldsymbol{a}+\boldsymbol{a}^*)^T B^{-1}(\boldsymbol{a}-\boldsymbol{a}^*)\\ &+\frac{1}{2}\int_{-\tau}^{\tau}\mathrm{d}t\,\nabla\cdot(\boldsymbol{a}-\boldsymbol{a}^*)\,,\end{aligned} \tag{2.48}$$

while the left-hand side can be expressed in terms of the conservative entropic force $\boldsymbol{F}_{\mathrm{ss}}^{\mathcal{S}}$ introduced in section 2.2, i.e.

$$\log \frac{p_{\mathrm{ss}}(\boldsymbol{x}(\tau))}{p_{\mathrm{ss}}(\boldsymbol{x}(-\tau))} = -\int_{-\tau}^{\tau} \mathrm{d}t\, \boldsymbol{F}_{\mathrm{ss}}^{\mathcal{S}} \circ \dot{\boldsymbol{x}}. \tag{2.49}$$

Notice that we are employing the Stratonovich convention in the stochastic integrals in eqs. (2.48) and (2.49). Although we opt not to delve into the details here, this follows because we (conveniently) decided on the time-symmetric discretisation of the path integral. Of course, these results are independent of this choice, however more care would have had to be taken were we to make a different one.

Now, since eqs. (2.48) and (2.49) must hold for all paths $\boldsymbol{x}$ and times τ, we may conclude that

$$\begin{cases} \boldsymbol{a} + \boldsymbol{a}^* = -B\boldsymbol{F}_{\mathrm{ss}}^{\mathcal{S}}, \\ (\boldsymbol{a}+\boldsymbol{a}^*)^T B^{-1} (\boldsymbol{a}-\boldsymbol{a}^*) + \nabla \cdot (\boldsymbol{a}-\boldsymbol{a}^*) = 0. \end{cases} \tag{2.50}$$

Clearly, the first line in eq. (2.50) gives us $\boldsymbol{a}^*$ as promised. The second line can be rewritten using the first, in addition to eq. (2.22) for $\boldsymbol{j}$, to obtain the orthogonality between the entropic force $\boldsymbol{F}_{\mathrm{ss}}^{\mathcal{S}}$ and the steady-state probability current $\boldsymbol{j}_{\mathrm{ss}}$, specifically

$$\boldsymbol{j}_{\mathrm{ss}} \cdot \boldsymbol{F}_{\mathrm{ss}}^{\mathcal{S}} = 0. \tag{2.51}$$

Finally, we conclude that the time-reversed process $\overleftarrow{\boldsymbol{X}}$ obeys the SDE

$$\dot{\overleftarrow{\boldsymbol{X}}} = -\boldsymbol{a} - B\boldsymbol{F}_{\mathrm{ss}}^{\mathcal{S}} + b\boldsymbol{\eta}. \tag{2.52}$$

We can gain deeper insight into how the forward and backward probabilities $\mathfrak{p}_{\mathrm{ss}}$ and $\overleftarrow{\mathfrak{p}}_{\mathrm{ss}}$ differ for some given trajectory $\boldsymbol{X}$ by investigating the quantity

$$\Delta\mathfrak{s}_{\mathrm{ss}}[\boldsymbol{X}] \equiv \log \frac{\mathfrak{p}_{\mathrm{ss}}[\boldsymbol{X}]}{\overleftarrow{\mathfrak{p}}_{\mathrm{ss}}[\boldsymbol{X}]}. \tag{2.53}$$

First, we observe that $\Delta\mathfrak{s}_{\mathrm{ss}}$ can be expressed as a sum of two terms

$$\Delta\mathfrak{s}_{\mathrm{ss}}[\boldsymbol{X}] = \Delta\mathfrak{s}_{\mathrm{ss}}^{\mathrm{sys}}[\boldsymbol{X}] + \Delta\mathfrak{s}^{\mathrm{res}}[\boldsymbol{X}], \tag{2.54}$$

where we have defined

$$\Delta\mathfrak{s}_{\mathrm{ss}}^{\mathrm{sys}}[\boldsymbol{X}] \equiv \log \frac{p_{\mathrm{ss}}(\boldsymbol{X}(-\tau))}{p_{\mathrm{ss}}(\boldsymbol{X}(\tau))}, \tag{2.55}$$

$$\Delta\mathfrak{s}^{\mathrm{res}}[\boldsymbol{X}] \equiv \log \frac{\mathcal{P}([\boldsymbol{X}]|\boldsymbol{X}(-\tau))}{\mathcal{P}\left([\overleftarrow{\boldsymbol{X}}]|\boldsymbol{X}(\tau)\right)}. \tag{2.56}$$

From these definitions, one may then show that

$$\Delta\mathfrak{s}_{\mathrm{ss}}^{\mathrm{sys}}[\boldsymbol{X}] = \int_{-\tau}^{\tau} \mathrm{d}t\, \boldsymbol{F}_{\mathrm{ss}}^{\mathcal{S}} \circ \dot{\boldsymbol{X}}, \tag{2.57}$$

$$\Delta\mathfrak{s}^{\mathrm{res}}[\boldsymbol{X}] = \int_{-\tau}^{\tau} \mathrm{d}t\, \boldsymbol{F}^{\mathcal{Q}} \circ \dot{\boldsymbol{X}}. \tag{2.58}$$

In particular, it follows immediately from eq. (2.54) and the force balance in eq. (2.26) that

$$\Delta\mathfrak{s}_{\mathrm{ss}}[\boldsymbol{X}] = \int_{-\tau}^{\tau} \mathrm{d}t\, \boldsymbol{F}_{\mathrm{ss}} \circ \dot{\boldsymbol{X}}, \tag{2.59}$$

where $\boldsymbol{j}_{\mathrm{ss}} = L\boldsymbol{F}_{\mathrm{ss}}$. As can be seen by direct comparison with eq. (2.30), eq. (2.54) expresses the deep connections between time-reversal at the level of individual trajectories and entropy production. Moreover, this relationship holds even when the fluctuations in the SDE in eq. (2.5) are non-thermal, and is therefore a universal property of all such dynamics.

Finally, we observe that in order to have $\Delta\mathfrak{s}_{\mathrm{ss}}[\boldsymbol{X}] = 0$ for all realisations $\boldsymbol{X}$, we must have $F_{\mathrm{ss}} = 0$. Equivalently, we see from eq. (2.59) that

$$\Delta\mathfrak{s}_{\mathrm{ss}} = 0 \qquad \text{if and only if} \qquad j_{\mathrm{ss}} = 0. \tag{2.60}$$

As can be seen from the definition of $\Delta\mathfrak{s}_{\mathrm{ss}}$ in eq. (2.53), this is in fact exactly the principle of detailed balance, which states that in equilibrium the net flux of probability mass between any pair of states vanishes [76], or

$$p_{\mathrm{ss}}(\boldsymbol{X}(-\tau))\mathcal{P}([\boldsymbol{X}]|\boldsymbol{X}(-\tau)) = p_{\mathrm{ss}}(\boldsymbol{X}(\tau))\mathcal{P}\left([\overleftarrow{\boldsymbol{X}}]|\boldsymbol{X}(\tau)\right) \quad \text{in equilibrium.} \tag{2.61}$$

Notice also that when $F_{\mathrm{ss}} = 0$, the entropic thermodynamic force and the force due to energy flow exactly balance, as remarked in section 2.2, so that $\boldsymbol{F}_{\mathrm{ss}}^{\mathcal{S}} = -\boldsymbol{F}^{\mathcal{Q}}$. From eq. (2.52), it therefore follows that

$$\dot{\overleftarrow{\boldsymbol{X}}} = \boldsymbol{a} + b\boldsymbol{\eta}, \tag{2.62}$$

and so in equilibrium the forward and time-reversed dynamics are equivalent.

This concludes our discussion of the relationship between time-reversal at the trajectory level and the principle of detailed balance. In the next section, we will see that we can also evaluate the steady-state expectation $\langle \Delta \mathfrak{s}_{\mathrm{ss}} \rangle$. By reformulating this ensemble average as a temporal average, we may then associate the total EPR $\dot{\mathcal{S}}$ to (long) individual realisations of the dynamics in eq. (2.5).

2.4 Stochastic entropy production and the Gallavotti-Cohen symmetry

In this section, we will first show the expected value of $\Delta \mathfrak{s}_{\mathrm{ss}}$ is given by the total EPR $\dot{\mathcal{S}}$. We will then investigate the higher order moments of $\Delta \mathfrak{s}_{\mathrm{ss}}$, and show that the moment generating function is related to a particular universal symmetry known as the *Gallavotti-Cohen symmetry* [24, 71]. From this we sketch a classic proof of a well-known and much celebrated fluctuation theorem, which can be considered the manifestation of the Gallavotti-Cohen symmetry in the probability distribution for the entropy production in the long time limit. Most of our discussion in this section is directly inspired by the seminal contribution of Lebowitz and Spohn in [24].

Before we begin, we highlight a general result which can be derived fairly straightforwardly from the stochastic analysis introduced in section 2.1. More specifically, this concerns the expected value of certain observables $\mathcal{O}$ that can be written as Stratonovich integrals of some function f against a stochastic process X. That is, we assume that we may write

$$\mathcal{O}[X] = \int_0^\tau \mathrm{d}t\, f(X) \circ \dot{X}. \tag{2.63}$$

From the transformation identity between the Itô and Stratonovich integrals in eq. (2.8), it immediately follows that we may rewrite $\mathcal{O}$ in Itô form as

$$\mathcal{O}[X] = \int_0^\tau \mathrm{d}t \left(f(X)\dot{X} + \frac{1}{2} f'(X) \left[\dot{X}\right] \right). \tag{2.64}$$

Thus, when X solves the SDE in eq. (2.5), we may compute the steady-state expectation $\langle \mathcal{O} \rangle$ of the observable $\mathcal{O}$ as

$$\langle \mathcal{O}[X] \rangle = \tau \int_{\mathcal{V}} \mathrm{d}x \left(f(x)a + \frac{1}{2} B f'(x) \right) p_{\mathrm{ss}} \tag{2.65}$$

$$= \tau \int_{\mathcal{V}} \mathrm{d}x\, j_{\mathrm{ss}}(x) f(x) \tag{2.66}$$

where the final equality follows from an integration by parts. Clearly, this argument extends trivially to arbitrary dimensions d by using the bilinearity of the quadratic covariation.

Now, observe that by using eq. (2.66), it follows from eqs. (2.54), (2.57) and (2.58) that $\langle \Delta\mathfrak{s}_{\mathrm{ss}}^{\mathrm{sys}} \rangle = 0$ and $\langle \Delta\mathfrak{s}_{\mathrm{ss}} \rangle = \langle \Delta\mathfrak{s}^{\mathrm{res}} \rangle$, where

$$\langle \Delta\mathfrak{s}^{\mathrm{res}}[\boldsymbol{X}] \rangle = 2\tau \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \boldsymbol{j}_{\mathrm{ss}} \cdot \boldsymbol{F}^{\mathcal{Q}}. \tag{2.67}$$

Consequently, assuming that we may replace temporal averages by averages over noise realisations, we obtain one of the central results of stochastic thermodynamics; namely that the total steady-state EPR can be obtained from the log-ratio of forward and reversed path probabilities, i.e.

$$\dot{\mathcal{S}} = \lim_{\tau\to\infty} \frac{1}{2\tau} \log \frac{\mathcal{P}[\boldsymbol{X}]}{\overleftarrow{\mathcal{P}}[\boldsymbol{X}]}, \tag{2.68}$$

where $\overleftarrow{\mathcal{P}}[\boldsymbol{X}] \equiv \mathcal{P}[\overleftarrow{\boldsymbol{X}}]$ (for ease of notation we will omit the starting point in the conditional probability functional $\mathcal{P}$ from this point onward). Note also that a finite-time (transient) generalisation of eq. (2.67), for which one starts from a non-stationary distribution p_0 and takes the expected value with respect to the law at time τ, can also be obtained, although the derivation becomes slightly more involved. Because of this, $\boldsymbol{F}^{\mathcal{Q}} \circ \dot{\boldsymbol{X}}$ should really be regarded as the instantaneous EPR due to energy (heat) flow between the system and reservoir.

Equation (2.68) allows us to construct the EPR for a given stochastic dynamics without explicit knowledge of the steady-state distribution p_{ss} or the current $\boldsymbol{j}_{\mathrm{ss}}$. Furthermore, in numerical simulations, we can use eq. (2.58), or equivalently eq. (2.68), to estimate the total EPR by sampling long trajectories and computing the corresponding temporal average [12, 31, 36, 58]. In chapters 5 and 6 we will use analytical tools in addition to this method to compute the EPR in both microscopic and hydrodynamic models of polar active matter.

We now turn our attention to investigating the higher order moments of $\Delta\mathfrak{s}^{\mathrm{res}}$, which will reveal the celebrated Gallavotti-Cohen symmetry. To this end, we introduce the process $\boldsymbol{Y} = (\boldsymbol{X}, \Delta\mathfrak{s}^{\mathrm{res}}[\boldsymbol{X}])^T$ which allows us to consider $\boldsymbol{Y}$ as the solution to the SDE

$$\dot{\boldsymbol{Y}} = \boldsymbol{a}_Y + b_Y \boldsymbol{\eta}, \tag{2.69}$$

where $\boldsymbol{\eta}$ is the same noise that drives $\boldsymbol{X}$, and the drift and diffusion terms $\boldsymbol{a}_Y$ and b_Y are given by

$$\boldsymbol{a}_Y = \left(\boldsymbol{a}, 2\boldsymbol{a}^T B^{-1}\boldsymbol{a} + \nabla\cdot\boldsymbol{a}\right)^T, \tag{2.70}$$

$$b_Y = \left(b^T, 2b^T B^{-1}\boldsymbol{a}\right)^T, \tag{2.71}$$

respectively. This trick allows us to construct the EPR $\Delta\mathfrak{s}^{\mathrm{res}}$ as the solution to an SDE, which we could not do if we considered only $\Delta\mathfrak{s}^{\mathrm{res}}$ by itself [70]. Note also that in eq. (2.70), we take the operator ∇ to act only on the $\boldsymbol{x}$ variables.

The FPE for the probability density $p_Y(\boldsymbol{y},t) \equiv p_Y(\boldsymbol{x},s,t)$ of the process $\boldsymbol{Y}$ can then be written as

$$\partial_t p_Y = -\nabla \cdot \boldsymbol{j}_X - \partial_s j_{\mathfrak{s}}, \tag{2.72}$$

where the probability currents $\boldsymbol{j}_X$ and $j_{\mathfrak{s}}$ are given explicitly by

$$\boldsymbol{j}_X = \boldsymbol{a} p_Y - \frac{1}{2} B \nabla p_Y - \boldsymbol{a} \partial_s p_Y, \tag{2.73}$$

$$j_{\mathfrak{s}} = \left(2\boldsymbol{a}^T B^{-1} \boldsymbol{a} + \nabla \cdot \boldsymbol{a}\right) p_Y - \nabla \cdot (\boldsymbol{a} p_Y) - 2\boldsymbol{a}^T B^{-1} \boldsymbol{a} \partial_s p_Y. \tag{2.74}$$

Note that in addition to assuming that the total probability current $\boldsymbol{j}_Y \equiv (\boldsymbol{j}_X, j_s)$ integrates to zero on the boundary of the composite domain $\mathcal{V} \times \mathbb{R}$, we also assume that both $\int_{\partial \mathcal{V}} \boldsymbol{j}_X(\boldsymbol{x},s,t) \cdot \mathrm{d}\boldsymbol{S} = 0$ for all s and $j_{\mathfrak{s}}(\boldsymbol{x},s,t)|_{s=-\infty}^{s=\infty} = 0$ for all $\boldsymbol{x}$. This ensures that there is no net flow of probability mass between the domains of $\boldsymbol{X}$ and $\Delta\mathfrak{s}^{\mathrm{res}}$.

We may now construct the moment generating function $M_\lambda(t)$ for $\Delta\mathfrak{s}^{\mathrm{res}}$ from the distribution p_Y. That is, we define

$$M_\lambda(t) = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \mu_\lambda(\boldsymbol{x},t) \tag{2.75}$$

where μ_λ is defined by

$$\mu_\lambda(\boldsymbol{x},t) = \int_{-\infty}^{\infty} \mathrm{d}s\, \mathrm{e}^{\lambda s} p_Y(\boldsymbol{x},s,t). \tag{2.76}$$

From the FPE for p_Y in eq. (2.72), it is fairly straightforward to perform the integral in eq. (2.76). Using the appropriate boundary conditions cited above, one finds that $\partial_t \mu_\lambda = \mathcal{L}_\lambda \mu_\lambda$, where the Lebowitz-Spohn operator $\mathcal{L}_\lambda$ is given explicitly by

$$\mathcal{L}_\lambda = 2\lambda(1+\lambda)\boldsymbol{a}^T B^{-1} \boldsymbol{a} - (1+\lambda)\nabla \cdot \boldsymbol{a} - (1+2\lambda)\boldsymbol{a} \cdot \nabla + \frac{1}{2} B_{\alpha\beta} \nabla_\alpha \nabla_\beta. \tag{2.77}$$

In particular, one may verify that that it satisfies the symmetry

$$\mathcal{L}_\lambda = \mathcal{L}^\dagger_{-\lambda-1}, \tag{2.78}$$

which is the celebrated Gallavotti-Cohen symmetry. The brilliant insight of Lebowitz and Spohn was to realise that because of eq. (2.78), the same symmetry is also manifest in the *scaled cumulant generating function* (SCGF) of the EPR. This then implies the celebrated fluctuation theorem of the EPR distribution function.

To see how the fluctuation theorem follows from eq. (2.78), we briefly sketch the argument of Lebowitz and Spohn in [24]. They invoke the so-called *Perron-Frobenius theorem*, which applies to positive linear operators such as $\mathcal{L}$. This implies that $\mathcal{L}_\lambda$ has a maximal positive eigenvalue $\Phi(\lambda)$, which is the same as that of its adjoint $\mathcal{L}_\lambda^\dagger$. In fact, from eq. (2.75), it follows that this is given by the SCGF

$$\Phi(\lambda) = \lim_{t\to\infty} \frac{1}{t} \log M_\lambda(t), \tag{2.79}$$

since for long times $\mu_\lambda(\boldsymbol{x},t) = \mathrm{e}^{\mathcal{L}_\lambda t}\mu_\lambda(\boldsymbol{x},0) \sim \mathrm{e}^{\Phi(\lambda)t}$. From the Gallavotti-Cohen symmetry in eq. (2.78) it therefore follows that

$$\Phi(\lambda) = \Phi(-\lambda - 1), \tag{2.80}$$

By then applying a standard result from large deviation theory known as the *Gärtner-Ellis theorem*, one may show that the probability distribution $p_\mathfrak{s}(s,t) = \int_\mathcal{V} \mathrm{d}\boldsymbol{x}\, p_Y(\boldsymbol{x},s,t)$ can be obtained in the long time limit from the Legendre-Fenchel transform of the SCGF, i.e.

$$\lim_{t\to\infty} -\frac{1}{t} \log p_\mathfrak{s}(st,t) = \sup_\lambda \{\lambda s - \Phi(\lambda)\}. \tag{2.81}$$

From eq. (2.80), one therefore deduces that in the long time limit

$$\frac{p_\mathfrak{s}(st,t)}{p_\mathfrak{s}(-st,t)} \sim \mathrm{e}^{st}, \tag{2.82}$$

which is the fluctuation theorem of the distribution $p_\mathfrak{s}$. The fascinating statement of eq. (2.82) is the rather intuitive fact that producing a positive amount of entropy st in a time t is exponentially more likely than producing the corresponding negative amount $-st$. Moreover, it is significant because it holds also arbitrarily far from equilibrium, and for the EPR in any system described by eq. (2.5).

Next, we approach the final two sections of this chapter, where we discuss applications of the above theory. Specifically, we first introduce in section 2.5 a set of tools for working with Brownian motions on the sphere, and discuss the properties of its dual representations, which will be highly valuable to our work in the coming chapters. Then, in section 2.6 we study a simple model of an active colloid in a thermal bath, which allows us to work out explicitly the above theory.

2.5 Intermezzo: spherical diffusions

Before we proceed in the next section to look at an application of the theory from the previous sections, we take a moment to introduce, en passant, what we refer to as *spherical diffusions*. These will be central to our work in the coming chapters, and some of the results derived here will also be used in section 2.6. In this section, we will show that there is a very appealing dual description of a Brownian particle on a manifold, which is achieved by looking at the SDE in local coordinates as well as in the coordinates of the higher-dimensional embedding Euclidean space. By switching between these different representations, we may choose the one most suited for our purposes.

Here and throughout, we denote by $\mathbb{S}^d = \left\{\boldsymbol{x} \in \mathbb{R}^{d+1} \,|\, x = 1\right\}$ the d-dimensional unit sphere. Perhaps most familiar to the statistical physicist is the spherical diffusion on the 1-sphere $\mathbb{S}^1$, often also referred to as a rotational diffusion. This is represented by a stochastic process $\theta(t)$ which is just a Brownian motion on $\mathcal{V} = [0, 2\pi]$ with periodic boundary conditions. Equivalently, θ solves the SDE

$$\dot{\theta} = \sqrt{2D_r}\xi, \tag{2.83}$$

where $D_r > 0$ is referred to as the rotational diffusion coefficient [77, 78]. Although this definition works well for 1-dimensional spherical diffusions, eq. (2.83) does not extend easily to higher dimensions. Indeed, the spherical diffusion on $\mathbb{S}^2$ is *not* just a 2-dimensional Brownian motion on a square domain with periodic boundaries.

In the following, we would like to construct the dynamical description of a Brownian motion on the sphere. Because of this, we approach this from the angle of differential geometry, by introducing the standard mathematical notion of a Brownian motion on a manifold $\Sigma \subset \mathbb{R}^d$, with the aim of specialising to $\Sigma = \mathbb{S}^{d-1}$. Specifically, this considers the orthogonal projection operator $\mathrm{P}_\Sigma(\boldsymbol{x})$ which maps vectors in $\mathbb{R}^d$ onto the tangent space of Σ at $\boldsymbol{x} \in \Sigma$, and defines the process $\boldsymbol{X}$ which solves the SDE

$$\dot{\boldsymbol{X}} = \mathrm{P}_\Sigma(\boldsymbol{X}) \circ \boldsymbol{\xi} \tag{2.84}$$

as a standard Brownian motion on Σ [62, 79]. Disregarding the Stratonovich discretisation for a moment, eq. (2.84) is actually fairly intuitive. For $\boldsymbol{X}$ on Σ, it tells us to project $\boldsymbol{\xi} \in \mathbb{R}^d$ onto the tangent space of Σ at $\boldsymbol{X}$. Thus, integrating eq. (2.84) we indeed obtain a curve on Σ for all times t. Furthermore, one may show that by choosing the Stratonovich $\circ$, the infinitesimal generator of the process in local coordinates is the Laplace(-Beltrami) operator on Σ, which is the natural generalisation of the generator of Brownian motion to differentiable manifolds.

Specialising now to the $d-1$-dimensional unit sphere $\mathbb{S}^{d-1}$, it is easy to see that the projection operator $\mathrm{P}_{\mathbb{S}^{d-1}}$ takes the form $\mathrm{P}_{\mathbb{S}^{d-1}}(\boldsymbol{e})_{\alpha\beta} = \delta_{\alpha\beta} - e_\alpha e_\beta$ [79, 80]. Thus, the process $\boldsymbol{e}(t) \in \mathbb{R}^d$ which solves

$$\dot{e}_\alpha = \left(\delta_{\alpha\beta} - e_\alpha e_\beta\right) \circ \xi_\beta \tag{2.85}$$

is a standard Brownian motion on $\mathbb{S}^{d-1}$. Throughout this thesis, we will refer to $\boldsymbol{e}$ as a *standard* $d-1$-dimensional spherical diffusion. We also define the spherical diffusion with rotational diffusivity D_r as the solution to the SDE

$$\dot{e}_\alpha = \sqrt{2D_r}\left(\delta_{\alpha\beta} - e_\alpha e_\beta\right) \circ \xi_\beta. \tag{2.86}$$

From both eqs. (2.85) and (2.86), it is straightforward to verify that $|\boldsymbol{e}(t)| = |\boldsymbol{e}(0)| \equiv 1$ for all times t as it should.

To see that this definition indeed reduces to eq. (2.83) on $\mathbb{S}^1$, we now write out the SDE in eq. (2.86) in local coordinates. By making the transformation $\boldsymbol{e} \equiv \boldsymbol{e}(\theta)$, where $\boldsymbol{e}(\theta) = (\cos\theta, \sin\theta)$, we can apply Itô's lemma in Stratonovich form to obtain

$$\dot{\boldsymbol{e}} = \boldsymbol{e}'(\theta) \circ \dot{\theta}, \tag{2.87}$$

where $\boldsymbol{e}' = \mathrm{d}\boldsymbol{e}/\mathrm{d}\theta$. Thus, if we take $\dot{\theta} = \sqrt{2D_r}\boldsymbol{e}'(\theta) \circ \boldsymbol{\xi}$, then we immediately obtain eq. (2.86) for $\boldsymbol{e}$. It therefore remains to show only that $\tilde{\xi} \equiv \boldsymbol{e}'(\theta) \circ \boldsymbol{\xi}$ is indeed a standard white noise. To see this, we simply observe that

$$\left[\frac{\mathrm{d}}{\mathrm{d}t} e'_\alpha, \xi_\alpha\right] = 0. \tag{2.88}$$

Combined with the transformation identity between Stratonovich and Itô integrals in eq. (2.8), we see that we do not pick up a spurious drift in the conversion, and may write $\tilde{\xi} = \boldsymbol{e}' \cdot \boldsymbol{\xi}$. A straightforward application of the Lévy characterisation of Brownian motion then implies that $\tilde{\xi}$ is indeed a unit white noise.

To go beyond the 1-sphere, it is convenient to work instead with the Itô form of eq. (2.86). One may obtain this by first noting that for a $d-1$-dimensional spherical diffusion with $\boldsymbol{e} \in \mathbb{R}^d$,

$$\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(e_\alpha e_\beta\right), \xi_\beta\right] = \sqrt{2D_r}(d-1)e_\alpha. \tag{2.89}$$

Hence, the SDE for $\boldsymbol{e}$ in eq. (2.86) can be written in Itô form as

$$\dot{e}_\alpha = -D_r(d-1)e_\alpha + \sqrt{2D_r}\left(\delta_{\alpha\beta} - e_\alpha e_\beta\right)\xi_\beta, \tag{2.90}$$

and in particular picks up a spurious drift. Interestingly, using the fact that the Itô integral is mean-zero, an immediate consequence of eq. (2.90) is that the autocorrelation function of $\boldsymbol{e}$ is given by

$$\langle \boldsymbol{e}(0) \cdot \boldsymbol{e}(t) \rangle = \mathrm{e}^{-D_r(d-1)t}. \tag{2.91}$$

Notice also that from eq. (2.91), we see directly that D_r has the dimensions of inverse time, and that it can be regarded as the typical time $\boldsymbol{e}$ takes to decorrelate, or 'forget its orientation'.

From eq. (2.90), one may straightforwardly compute that on $\mathbb{S}^2$, in local coordinates $\boldsymbol{e}(\theta,\phi) = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta)^T$, the infinitesimal generator is given by the spherical Laplacian

$$\mathcal{L} = \frac{D_r}{\tan\theta}\partial_\theta + D_r\partial_\theta^2 + \frac{D_r}{\sin^2\theta}\partial_\phi^2 \tag{2.92}$$

Thus, we can read off that the 2-dimensional spherical diffusion solves the SDE

$$\dot{\theta} = \frac{D_r}{\tan\theta} + \sqrt{2D_r}\xi_\theta, \tag{2.93}$$

$$\dot{\phi} = \frac{\sqrt{2D_r}}{\sin\theta}\xi_\phi. \tag{2.94}$$

In general, as remarked above, the infinitesimal generator for the spherical diffusion on $\mathbb{S}^{d-1}$ is just the Laplace-Beltrami operator $\mathcal{L} = D_r\Delta_{\mathbb{S}^{d-1}}$ [80].

Finally, we briefly mention a nice result which is also straightforwardly derivable from the above. It concerns an active Brownian particle, which we will discuss more thoroughly in the following section. In its simplest form, this is a model for a particle with position $\boldsymbol{X}$ which obeys $\dot{\boldsymbol{X}} = v_0\boldsymbol{e}$, where v_0 is some constant bare velocity and $\boldsymbol{e}$ is a $(d-1)$-dimensional spherical diffusion with rotational diffusivity D_r. From eq. (2.91) for the autocorrelation function of $\boldsymbol{e}$, it immediately follows that for long times τ

$$\langle X^2 \rangle \sim 2\frac{v_0^2\tau}{D_r(d-1)}. \tag{2.95}$$

Hence, we may define an effective diffusion constant

$$D_{\text{eff}} \equiv \lim_{\tau\to\infty}\frac{1}{2\tau}\langle X^2\rangle \tag{2.96}$$

$$= \frac{v_0^2}{D_r(d-1)} \tag{2.97}$$

for the active Brownian particle, and a corresponding effective diffusion length scale $\ell_D \sim v_0/D_r$.

2.6 Stochastic thermodynamics of an actively driven colloid

In this section we will work out the consequences of the theory of stochastic thermodynamics introduced in the previous sections explicitly for the example system of an equilibrium colloid suspended in a viscous fluid. This allows us to delineate how stochastic thermodynamics can be used to understand energetics on the level of individual realisations of the dynamics of a system. In this sense, stochastic thermodynamics extends the classical thermodynamics of equilibrium systems by generalising notions of energy, heat and entropy to the true non-equilibrium regime.

We will begin first by briefly recalling the stochastic thermodynamics of an equilibrium colloid suspended in a viscous fluid with position $\boldsymbol{X}$ and velocity $\boldsymbol{V}$ under the influence of some external potential U [22, 23]. This will serve as a benchmark when we introduce active forcing momentarily, with which we may compare how the theory works out differently in the two cases. In the fluid bath, the equilibrium colloid is under the influence of both a viscous drag $\boldsymbol{f}_{\mathrm{d}} = -\gamma \boldsymbol{V}$, where the drag coefficient $\gamma = 6\pi\eta_v R$ is given by Stoke's law and η_v is the dynamic viscosity of the fluid, in addition to random thermal fluctuations $\boldsymbol{f}_{\mathrm{t}} = \sqrt{2D_{\mathrm{t}}}\boldsymbol{\eta}$ that we take to be white and of unknown constant intensity $D_{\mathrm{t}} > 0$. In the overdamped limit where we neglect inertia, Newton's law therefore implies that the forces balance, i.e.

$$0 = \boldsymbol{f}_{\mathrm{d}} + \boldsymbol{f}_{\mathrm{p}} + \boldsymbol{f}_{\mathrm{t}}, \tag{2.98}$$

where $\boldsymbol{f}_{\mathrm{p}} = -\nabla U$ is the force experienced by the particle due to the potential U. A trivial rewriting of eq. (2.98), using the fact that $\dot{\boldsymbol{X}} = \boldsymbol{V}$, then leads to the equation of motion

$$\dot{\boldsymbol{X}} = -\mu \nabla U(\boldsymbol{X}) + \sqrt{2D}\boldsymbol{\eta}, \tag{2.99}$$

where $\mu = 1/\gamma$ and $D = D_{\mathrm{t}}\mu^2$ are referred to as the mobility and diffusion coefficients respectively.

Equilibrium statistical mechanics dictates that in steady-state, the position $\boldsymbol{X}$ of the particle is distributed according to the equilibrium Boltzmann law $p_{\mathrm{ss}} \propto \mathrm{e}^{-U(\boldsymbol{x})/T}$. Combining this with the dynamical description of the particle in eq. (2.98) one straightforwardly derives the *Einstein relation*

$$D = \mu T. \tag{2.100}$$

Indeed, there are many (at least three) ways to see this using only the machinery from sections 2.1 to 2.4. One way is to observe that in equilibrium we may compute the entropic

force

$$\boldsymbol{F}^{\mathcal{S}} = \nabla U/T, \tag{2.101}$$

from which the Einstein relation follows by requiring that the forward and reversed processes obey the same dynamics by taking $\boldsymbol{a}^* = \boldsymbol{a}$ in eq. (2.50). The importance of the Einstein relation should not be understated, as it relates the strength of fluctuations via the diffusion coefficient D and dissipation via the friction $\gamma = 1/\mu$. In particular, it is one example of a more general theorem called the *fluctuation-dissipation theorem* (FDT), which in equilibrium relates correlation functions to the linear response of the system when perturbed by an external drive [23].

Sekimoto cleverly noted in his monograph that there is an intuitive way to introduce heat on the level of individual trajectories for the dynamics in eq. (2.99) which is consistent with the first law [21]. To see this, we simply observe that the total force applied by the fluid bath on the colloid is given by $\boldsymbol{f}_{\mathrm{d}} + \boldsymbol{f}_{\mathrm{t}}$. By reciprocal action, it follows that energy supplied to the bath from the colloid along a realisation $\boldsymbol{X}$ is

$$\mathcal{Q}[\boldsymbol{X}] = -\int_{-\tau}^{\tau} \mathrm{d}t\, (\boldsymbol{f}_{\mathrm{d}} + \boldsymbol{f}_{\mathrm{t}}) \circ \dot{\boldsymbol{X}}, \tag{2.102}$$

which we identify as the heat dissipated by the colloid along the trajectory $\boldsymbol{X}$.

Again, the Stratonovich discretisation is employed in eq. (2.102). To see why, note that the force balance in eq. (2.98) implies that $\mathcal{Q} = \int_{-\tau}^{\tau} \mathrm{d}t\, \boldsymbol{f}_{\mathrm{p}} \circ \dot{\boldsymbol{X}}$. By identifying $\boldsymbol{F}^{\mathcal{Q}} = \boldsymbol{f}_{\mathrm{p}}/T$, we therefore find that the stochastic entropy production $\Delta\mathfrak{s}^{\mathrm{res}}[\boldsymbol{X}] = \mathcal{Q}/T$ corresponds to the entropy production in the reservoir due to heat flow as promised in section 2.2. Moreover, a straightforward computation will show that $\mathcal{Q} = -\Delta E$, where $\Delta E \equiv U(\boldsymbol{X}(\tau)) - U(\boldsymbol{X}(-\tau))$ is the change in internal energy of the particle. Thus, since $\Delta\mathfrak{s}_{\mathrm{ss}} = 0$ in equilibrium, both the Clausius relation

$$\Delta\mathfrak{s}_{\mathrm{ss}}^{\mathrm{sys}} = -\mathcal{Q}/T, \tag{2.103}$$

as well as the first law

$$\Delta E = T\Delta\mathfrak{s}_{\mathrm{ss}}^{\mathrm{sys}} \tag{2.104}$$

follow from eq. (2.54).

Our aim now will be to extend these considerations to the non-equilibrium scenario where we include also an active force $\boldsymbol{f}_{\mathrm{a}} = \gamma v_0 \boldsymbol{e}$ of constant strength, where $\boldsymbol{e}$ is a spherical diffusion with rotational diffusivity D_r and v_0 is some bare self-propulsion speed. This is a standard model often referred to as an *active Brownian particle* (ABP) [77, 78]. In the

overdamped limit, the force balance now reads

$$0 = \boldsymbol{f}_\mathrm{d} + \boldsymbol{f}_\mathrm{a} + \boldsymbol{f}_\mathrm{p} + \boldsymbol{f}_\mathrm{t}, \tag{2.105}$$

giving us the dynamics

$$\dot{\boldsymbol{X}} = v_0 \boldsymbol{e} - \mu \nabla U(\boldsymbol{X}) + \sqrt{2D}\boldsymbol{\eta}. \tag{2.106}$$

For the ABP, there is no FDT that allows us to identify D with a temperature as in eq. (2.100). Nonetheless, one may still argue as before that $\mathcal{Q}$ in eq. (2.102) represents the energy dissipated by the colloid to the fluid bath [29, 34]. From eq. (2.105), however, we now obtain

$$\Delta E = -\mathcal{Q}[\boldsymbol{X}] + \mathcal{W}_\mathrm{a}[\boldsymbol{X}], \tag{2.107}$$

where the *active work* $\mathcal{W}_\mathrm{a}$ is defined by

$$\mathcal{W}_\mathrm{a}[\boldsymbol{X}] = \gamma v_0 \int_{-\tau}^{\tau} \mathrm{d}t\, \boldsymbol{e} \cdot \dot{\boldsymbol{X}}, \tag{2.108}$$

which is independent of discretisation since $\left[\dot{e}_\alpha, \dot{X}_\beta\right] = 0$.

Moreover, for the ABP, one may show that $\boldsymbol{F}^{\mathcal{Q}} = \big(\mu(\boldsymbol{f}_\mathrm{a} + \boldsymbol{f}_\mathrm{p})/D, \boldsymbol{0}\big)$, where the extra components stem from the $\boldsymbol{e}$ dynamics, and so we may again identify $\Delta\mathfrak{s}^\mathrm{res}$ with the entropy production due to dissipation to the fluid bath, i.e.

$$\Delta\mathfrak{s}^\mathrm{res} = \frac{\mu\mathcal{Q}}{D}. \tag{2.109}$$

This motivates setting $T_0 = \mu/D$ in analogy with eq. (2.103), which has dimensions of a temperature, although there is no guarantee it is the actual temperature of the bath. Despite eq. (2.109), in this case there is no trajectory level Clausius relation since $\Delta\mathfrak{s}_\mathrm{ss} \neq 0$, and so $\Delta\mathfrak{s}^\mathrm{res} = \Delta\mathfrak{s}_\mathrm{ss} - \Delta\mathfrak{s}^\mathrm{sys}_\mathrm{ss}$. To recover a Clausius inequality, one must generalise the above appropriately in order to take a finite-time expectation of eq. (2.109), rather than a steady-state expectation for which $\left\langle \Delta\mathfrak{s}^\mathrm{sys}_\mathrm{ss} \right\rangle = 0$ trivially. Finally, replacing $\mathcal{Q}$ in eq. (2.107) with $\Delta\mathfrak{s}^\mathrm{res}$ in eq. (2.109) and using entropy balance, we obtain

$$\Delta E + T_0 \Delta\mathfrak{s}_\mathrm{ss} = T_0 \Delta\mathfrak{s}^\mathrm{sys}_\mathrm{ss} + \mathcal{W}_\mathrm{a}. \tag{2.110}$$

In contrast with the equilibrium colloid, the active Brownian particle is a truly non-equilibrium model. This can be verified by explicitly computing the total EPR $\dot{\mathcal{S}}$. To this end, we make the simplifying assumption that the particle is free, i.e. $U = \text{const}$. Then, using the

dynamics in eq. (2.106), it is straightforward to compute

$$\langle \mathcal{W}_{\mathrm{a}}[\boldsymbol{X}] \rangle = 2\tau\gamma v_0^2 \tag{2.111}$$

Consequently, dividing eq. (2.109) by 2τ and taking the limit $\tau \to \infty$, it follows that the total EPR $\dot{\mathcal{S}}$ is given by

$$\dot{\mathcal{S}} = \frac{v_0^2}{D}. \tag{2.112}$$

In this case, the $\dot{\mathcal{S}}$ simply quantifies the dissipation of the particle due to its persistent motion against friction.

For the remainder of this thesis, we will be concerned with systems composed of many such active particles. However, rather than interacting via e.g. steric repulsion, they interact via *alignment*. Here we consider, for sake of argument, dynamics for the active force $\boldsymbol{f}_a$ where

$$\dot{e}_\alpha = \left(\delta_{\alpha\beta} - e_\alpha e_\beta\right) \circ \left(\tau_\beta + \sqrt{2D_r}\xi_\beta\right), \tag{2.113}$$

and $\boldsymbol{\tau} \equiv \boldsymbol{\tau}(\boldsymbol{X}, \boldsymbol{e}, t)$ is some torque that we leave unspecified. In this situation, however, one may show that

$$\Delta \mathfrak{s}^{\mathrm{res}} = \frac{\mu \mathcal{Q}}{D} + \frac{1}{D_r} \int_{-\tau}^{\tau} \mathrm{d}t\, \boldsymbol{\tau} \circ \dot{\boldsymbol{e}} \tag{2.114}$$

where, as before, $\mathcal{Q}$ is the dissipation to the fluid bath in eq. (2.102).

The second term in eq. (2.114) represents a type of energy dissipation due to the now 'active' dynamics of $\boldsymbol{e}$, which is influenced by fluctuations $\boldsymbol{\xi}$ that are not necessarily from the fluid bath. These fluctuations could for example emerge from some internal dynamics of the active particle of which we are currently unaware. Moreover, from

$$\dot{\mathcal{S}} = \frac{v_0^2}{D} + \lim_{\tau\to\infty} \frac{1}{2D_r\tau} \int_{-\tau}^{\tau} \mathrm{d}t\, \boldsymbol{\tau} \circ \dot{\boldsymbol{e}}, \tag{2.115}$$

we see explicitly that the new term in eq. (2.114) contributes to the irreversibility of the dynamics, in addition to the constant dissipation v_0^2/D from the persistent motion. To recover an interpretation of this in terms of heat, it could for example be necessary to describe also the coarse-grained internal degrees of freedom of the particle that are neglected in the phenomenological Langevin equation in eq. (2.113). A central theme throughout this thesis, however, will be to demonstrate that this regardless makes sense as an *informatic entropy production rate* [28], which quantifies the irreversibility encoded in (what will be) the interactions represented by $\boldsymbol{\tau}$ and the resulting dynamics.

2.7 Conclusion

In this chapter we have reviewed the stochastic thermodynamics of dynamics described by differential equations driven by white noise. This theory, which has gained significant momentum over only the past few decades, is becoming one of the central pillars of the modern non-equilibrium statistical mechanics [25, 38]. As we have seen, one of the main players in this theory is the stochastic entropy production in eq. (2.53), which can be used to obtain the total EPR $\dot{\mathcal{S}}$ from a temporal average involving only a single (long) realisation of the dynamics (cf. eq. (2.68)), rather than from an ensemble average or knowledge of the probability distribution p. For all dynamics described by SDEs of the form of eq. (2.5), the total EPR quantifies the extent to which the principle of detailed balance in eq. (2.61) breaks down, thus quantifying the irreversibility of the dynamics. In the following chapters, we will seek to understand how this EPR constrains fluctuations in collectively moving systems in which spontaneous currents arise due to the aligning dynamics.

First, in section 2.1, we reviewed some results from stochastic analysis. In particular, we recalled that the stochastic integral depends non-trivially on how it is discretised, and that the Stratonovich and Itô discretisations are related by a spurious drift term. Not only did we see in this chapter that using the quadratic covariation to calculate the spurious drift can be quite helpful, but it may also be used for many of the calculations we present later on in this thesis. Further to this, we also recalled that the FPE in eq. (2.17) describes the evolution of the probability distribution $p(\boldsymbol{x},t)$ of the dynamics $\boldsymbol{X}$.

To construct the stochastic thermodynamics of non-equilibrium systems, we introduced in section 2.2 the entropy balance equation, which expresses the total entropy production rate $\dot{\mathcal{S}}$ as the sum of the entropy production in the system $\dot{\mathcal{S}}^{\mathrm{sys}}$ and the reservoir $\dot{\mathcal{S}}^{\mathrm{res}}$ [24, 25]. In sections 2.3 and 2.4, we then connected this deterministic description in terms of the probability distribution p with a trajectory level formulation of the entropy balance equation, showing in particular that one may make sense of the entropy production on the level of individual realisations of the dynamics [56]. This will be key to our discussions in chapters 5 and 6 when we study the EPR in the *microscopic* and *hydrodynamic Vicsek models*.

En passant, in section 2.5, we introduced what we refer to as spherical diffusions [80], which as we shall see next in chapters 3 and 4 play a central role in the continuous-time formulation of the Vicsek model. Using the stochastic calculus reviewed in section 2.1, we showed how one may straightforwardly derive many useful properties of these by utilising the dual representation of the process in local coordinates and as a process embedded in $\mathbb{R}^d$. Finally, in section 2.6, we reviewed Sekimoto's notion of heat along a stochastic trajectory [21], and how the framework introduced in sections 2.2 to 2.4 combined with this provides us with a trajectory level first law for equilibrium processes.

Chapter 3

Self-propelled polar particles

Pioneered by Vicsek [17], the model which now bears his name has inspired a large body of research on the transition to collective 'flock' motion in active particle systems. Though initially constructed as a discrete-time, continuous-space automaton of particles with polar symmetry in $d = 2$, it has by now been generalised and recast in various forms; in continuous-time [81–84], for spatial dimensions $d \neq 2$ [85, 86], to topological (rather than metric) interactions [87–89], to systems with nematic symmetry [1, 85, 90–94], as well as to include steric repulsion and density-dependent bare self-propulsion speeds [1, 82, 84]. Notwithstanding, significant credit for popularising the model should be attributed to a handful of seminal contributions that still represent the flocking avant-garde. These include the phenomenological hydrodynamic theory of Toner and Tu [18, 19] and development of a systematic coarse-graining approach by Bertin, Droz and Grégoire [95, 96], in addition to the detailed descriptions of the phase diagram of the model (and the related Active Ising model) presented by Grégoire, Chaté, Solon and Tailleur [97–101].

In the following chapter and the next, we take a slightly different approach to the **Vicsek model** (VM) by considering instead an analogous dynamics in continuous time. With that said, the primary intent of this chapter is not to present novel results, although we do present some. Rather, we review here the central components of the theory of the VM, with a perspective that focuses more specifically on how the ρ-T (density-temperature) phase diagram connects with the mean-field theory (MFT) which is exact at infinite density. In addition to introducing the model which will be the focus for the remainder of this thesis, this presentation will serve a twofold purpose. First, in chapter 4, we study a closely related model, which we refer to as the **Active XY model** (AXYM) for reasons that will become apparent. The tools we develop in this chapter will be directly applicable to the AXYM, and moreover, the results we derive here will serve as a useful benchmark in chapter 4. Second,

in chapter 5, we will study the entropy production rate (EPR) in microscopic models of flocking, and will then employ the results we derive here and in chapter 4.

We begin in section 3.1 with a short review of the classical VM, where we present some of the key results from the literature. Then, in section 3.2, we introduce the continuous-time formulation of the VM dynamics, henceforth referred to as the **continuous-time Vicsek model** (CTVM), which has the benefit of being amenable to tools from the Itô-Stratonovich calculus presented in chapter 2. In sections 3.3 and 3.4 we then develop and discuss the implications of the MFT of the CTVM. This is done in two steps, by first deriving the steady-state distribution at infinite density in section 3.3, and subsequently discussing the linear stability of the homogeneous flocking phase. We also discuss in section 3.5 a mapping from the MFT of the CTVM to the Toner-Tu type hydrodynamic model, which is valid in the vicinity of the second order critical point in the microscopic model. Finally, in section 3.6, we present some concluding remarks from this chapter.

3.1 The Vicsek model: onset of collective motion

Originally formulated as a discrete time model in $d = 2$, the **Vicsek model** (VM) [17] considers N particles with polar symmetry in a finite volume $\mathcal{V} = [0, L]^2$ (here endowed with periodic boundaries). Much of its popularity can be attributed to its simplicity and similarity with the XY model known from equilibrium thermodynamics. In a historical context, it is particularly important as one of the early examples in non-equilibrium statistical mechanics of a transition to *collective motion*.

In the following, we will let $\{\boldsymbol{x}_i\}$ denote the set of particle positions and $\{\boldsymbol{e}_i\}$ the set of their *polar directors* (or *spins*), where i is the particle label. On each time-step, positions and spins are updated simultaneously according to

$$\boldsymbol{x}_i(t+\Delta t) = \boldsymbol{x}_i(t) + v_0 \Delta t\, \boldsymbol{e}_i(t), \tag{3.1}$$

$$\boldsymbol{e}_i(t+\Delta t) = \mathcal{R}\left(\boldsymbol{\chi}^{\eta}(t)\right) \hat{\boldsymbol{m}}_i(t), \tag{3.2}$$

where $\Delta t > 0$ is some positive time-increment. In eq. (3.1), which encodes the translational dynamics, the simplifying assumption is made that particles self-propel at constant velocity v_0. Equation (3.2), on the other hand, captures the alignment of spins. Specifically, on each update, spins align with the direction $\hat{\boldsymbol{m}}_i = \boldsymbol{m}_i / m_i$ of the local *polar density* (or *magnetisation density*)

$$\boldsymbol{m}_i = \frac{1}{\pi \ell_0^2} \sum_{j \text{ nn. } i} \boldsymbol{e}_j. \tag{3.3}$$

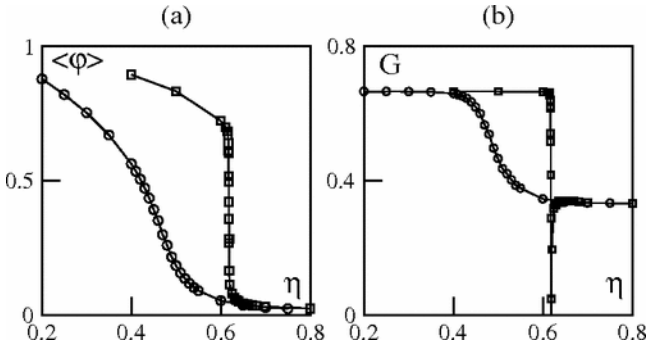


Fig. 3.1 Transition to collective motion in the VM. The figure is taken from [97]. In (a), the order parameter (denoted here by φ) is plotted against the noise intensity η. The transition, which appears continuous for small system sizes (circles), was shown to be discontinuous when the scalar noise was promoted to a 'vectorial noise' (squares). The corresponding Binder cumulant G is shown in (b). In [97], G was also shown to exhibit similar behaviour for the VM with scalar noise for sufficiently large system sizes.

In eq. (3.3) j nn. i denotes a sum over the nearest neighbours j of i with $|\boldsymbol{x}_i - \boldsymbol{x}_j| < \ell_0$, and $\ell_0 \ll L$ is some fixed microscopic interaction length. To avoid any ambiguity in eq. (3.2), we include particle i itself as a nearest neighbour of i. Alignment is further assumed to be imperfect, and particles make an error given by a uniformly distributed random angle $\chi^\eta(t) \sim U(-\eta, \eta)$ where $0 \leq \eta \leq \pi$. This is encoded in eq. (3.2) via the action of the rotation matrix $\mathcal{R}$ on $\hat{\boldsymbol{m}}_i$, where $\mathcal{R}$ is given explicitly by

$$\mathcal{R}(\chi) = \begin{pmatrix} \cos\chi & -\sin\chi \\ \sin\chi & \cos\chi \end{pmatrix}. \tag{3.4}$$

Although the above construction is inherently two-dimensional, the model can fairly straightforwardly be generalised to $d \geq 2$ [81, 85]. In fact, we will see one way this may be achieved in section 3.2.

The achievement of Vicsek et al. in [17] was to observe that the noise intensity η controls the transition from a high-temperature disordered phase to a low-temperature ordered one with collective motion characterised by a spontaneous magnetisation, illustrated in fig. 3.1. Naturally, the order parameter for this transition can be identified with the expected steady-state magnetisation $\langle \bar{m} \rangle$, where the mean global (fluctuating) magnetisation $\bar{\boldsymbol{m}}$ is defined by

$$\bar{\boldsymbol{m}} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{e}_i, \tag{3.5}$$

which attains non-zero values $\langle \bar{m} \rangle > 0$ in the ordered collectively moving phase. Initially, the transition to collective motion in the VM was believed to be second order, with the order

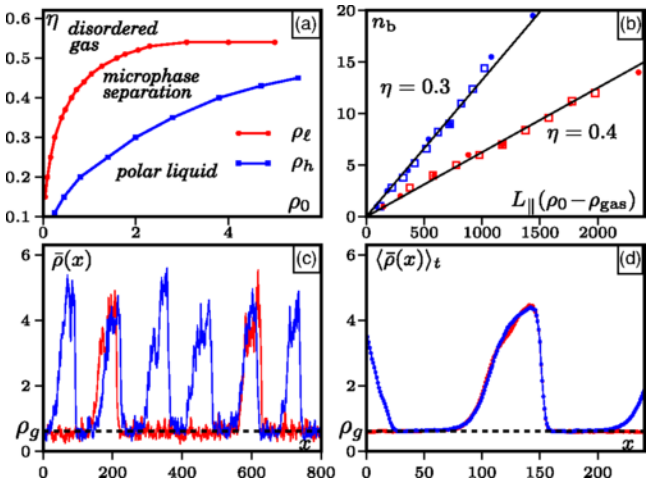


Fig. 3.2 Phase coexistence in the VM. The figure is taken from [99]. In (a), the phase diagram of the VM is shown in the ρ_0-η (density-noise intensity) plane, with the liquid and gas binodals ρ_h and ρ_l (respectively). Importantly, the critical point, where ρ_h and ρ_l meet and the transition becomes second order, is located at $\rho_0 = \infty$. In (b), the number n_b of bands in the coexistence region is shown to scale linearly with the system size. Finally, in (c) and (d), typical banded profiles in the VM are shown, with a well-defined gaseous density ρ_g. For both (c) and (d), the density profile is averaged over the direction perpendicular to the direction of motion, while in (d) an additional temporal average is performed.

parameter varying continuously across the transition, similarly to the order-disorder transition in the Ising model [17]. However, aided by extensive computer simulations, Grégoire et al. demonstrated that the transition is in fact first order [97]. They showed that the Binder cumulant $G = 1 - \langle \bar{m}^4 \rangle / 3 \langle \bar{m}^2 \rangle^2$ drops to negative values at the onset of collective motion (see fig. 3.1), together with the observation of phase coexistence. Importantly, a reason this behaviour remained so elusive is that the model exhibits a large, although finite, correlation length near the onset of order. In consequence, nucleation of a band requires that the system size is larger than the correlation length, while for smaller L the transition would appear continuous [102].

Subsequent analyses addressed more closely the dynamics in the miscibility gap, showing in particular that the VM exhibits *microphase-separation* (MPS) with a smectic arrangement of high-density collectively moving bands traveling against an isotropic gaseous background [99, 100] (see fig. 3.3 for illustrations of the various phases). This work led to the understanding that the phase diagram of the VM, illustrated in fig. 3.2, should rather be considered analogous to the canonical liquid-gas phase diagram in the ρ-T plane. A crucial distinction, however, is that the critical point in the VM is shifted to infinite density. Physically this is rather intuitive, since transitioning to from the gaseous isotropic to the *polar liquid* (PL)

ordered phase requires breaking the continuous rotational symmetry, and so there cannot be a super-critical region [98, 99, 101].

As pointed out by Toner et al. [18, 19], the transition to collective motion in the VM is interesting also because it appears to be in violation with the well-known Mermin-Wagner theorem. This prohibits the breaking of a continuous symmetry and development of long-ranged order in $d \leq 2$. Of course, the caveat here is that the Mermin-Wagner theorem holds only in equilibrium, and so cannot be applied to the VM. This led Toner and Tu to the remarkable feat of developing a phenomenological hydrodynamic theory of the VM, based purely on symmetry considerations, for which exact critical exponents could be obtained in $d = 2$. However, in a reanalysis it was observed that certain terms were missing in their original hydrodynamic expansion, rendering the argument leading to the exact exponents invalid when these were properly accounted for [103].

The following sections will deal more closely with a time-continuous formulation of the VM [81]. By formulating the dynamics in this way, they can be approached using the tools of Itô-Stratonovich calculus, which in turn allow us to provide analytical insight into many of the properties of the VM. Although most of the results presented in this chapter may be found elsewhere in the literature, our emphasis is on the tools we use to derive them and the results that will be important to us in the following chapters. We begin in section 3.2 by formulating the continuous-time analogue of the VM in the form of a Langevin equation. Then, in section 3.3, we investigate the steady-state distribution of the model at infinite density, which allows a series of exact results to be extracted. Finally, in section 3.5, we derive a mapping from the microscopic continuous-time Vicsek dynamics onto a 'Toner-Tu' type hydrodynamic theory, valid at infinite density in the high-temperature limit near the critical point.

3.2 Continuous-time Vicsek dynamics

An apparent downside of the discrete-time Vicsek algorithm presented in eqs. (3.1) and (3.2) is that direct computations are difficult. The standard approach to dealing with the VM analytically is therefore to derive hydrodynamic equations by writing down a collision integral and constructing a Boltzmann equation [85, 87, 92, 95, 96]. In this chapter we aim to describe a different approach, beginning in this section by introducing its continuous-time Langevin analogue which was first proposed in [81]. As will become apparent in chapter 4, this formulation illuminates an interesting feature of the Vicsek interaction, namely that it ignores the local signal strength. Another motivation we have for choosing this approach, is

that it allows us to construct an entropy production rate (EPR) associated with the dynamics, which will be the main topic of discussion in chapter 5.

Before introducing the continuous-time dynamics, we observe that a straightforward generalisation of the VM involves considering a time-dependent *contact* (or *interaction*) kernel $\mathcal{C} = \big(\mathcal{C}_{ij}(t)\big)$ [82] and defining instead the local polar density by

$$\boldsymbol{m}_i^{\mathcal{C}} = \sum_{j=1}^{N} \mathcal{C}_{ij}\boldsymbol{e}_j. \tag{3.6}$$

In this set-up, the standard VM reduces to the case of an isotropic kernel $\mathcal{C}_{ij}(t) = \mathcal{C}_0(|\boldsymbol{x}_j(t) - \boldsymbol{x}_i(t)|)$ with the choice

$$\mathcal{C}_0(r) = \frac{1}{\pi\ell_0^2}\Theta(1 - r/\ell_0), \tag{3.7}$$

where Θ is the Heaviside function. We will mostly be concerned with the subclass of isotropic kernels $\mathcal{C}_0$, and restrict to those that in d dimensions are normalised according to

$$\int_0^\infty \mathrm{d}r\, r^{d-1}\mathcal{C}_0(r) = \frac{\Gamma\left(\frac{d}{2}\right)}{2\pi^{d/2}}, \tag{3.8}$$

where Γ is the Gamma function. Generally, when the interaction distance ℓ_0 is fixed, we will also assume that we may write

$$\mathcal{C}_0(r) = \frac{\Gamma\left(\frac{d}{2}+1\right)}{\pi^{d/2}\ell_0^d}\phi(r/\ell_0), \tag{3.9}$$

where ϕ is independent of any microscopic parameters (other than possibly d). For example, in eq. (3.7), $\phi(r) = \Theta(1-r)$, which also trivially extends to arbitrary d via eq. (3.9). For other bump functions ϕ with $\phi(r) > 0$ for $r < 1$ and $\phi(r) = 0$ otherwise, eq. (3.8) implies that ϕ should be normalised according to

$$\int_0^1 \mathrm{d}r\, r^{d-1}\phi(r) = \frac{1}{d}. \tag{3.10}$$

Finally, note also that eq. (3.8) ensures $\mathcal{C}_{ij}$ has the dimensions of a density as it should, while ϕ is dimensionless.

Other popular choices of interaction kernel include the so-called 'topological' kernels, for which the number K of nearest neighbours is kept fixed rather than the interaction distance [87–89]. These can be formulated by introducing a position-dependent microscopic

interaction range $\ell_0(\boldsymbol{x},[\rho])$ defined implicitly via the equation

$$\int_{B(\boldsymbol{x},\ell_0)} \mathrm{d}\boldsymbol{x}'\,\rho\left(\boldsymbol{x}'\right) = K. \tag{3.11}$$

Here, $\rho(\boldsymbol{x})$ is the local density at position $\boldsymbol{x}$ and $B(\boldsymbol{x},\ell_0) = \{\boldsymbol{x}' \,|\, |\boldsymbol{x}-\boldsymbol{x}'| < \ell_0\}$ is a ball of radius ℓ_0 centered at $\boldsymbol{x}$. It has been argued, on the basis of experiments, that such a choice might more accurately capture the collective dynamics of animal flocks [104]. In any case, the resulting phenomenology does not differ significantly from the VM [89], and so the 'metric' choice will suffice for our purposes.

Deriving the continuous-time dynamics from the VM algorithm in eqs. (3.1) and (3.2) can now be done fairly straightforwardly, and to do this we follow the arguments of [81]. To this end, we begin by considering the VM at zero noise by setting $\eta = 0$. Furthermore, we will denote by $\boldsymbol{e}_i^n = \boldsymbol{e}_i(n\Delta t)$ the evaluation of $\boldsymbol{e}_i$ at time-step $n\Delta t$, as well as by

$$\boldsymbol{e}_i^{n+\frac{1}{2}} = \frac{\boldsymbol{e}_i^n + \boldsymbol{e}_i^{n+1}}{2} \tag{3.12}$$

its interpolation at half-integer time-steps. Equation (3.2) then implies that

$$\frac{e_{i\alpha}^{n+1} - e_{i\alpha}^n}{\Delta t} = \frac{1}{\Delta t}\left(\delta_{\alpha\beta} - e_{i\alpha}^{n+\frac{1}{2}} e_{i\beta}^{n+\frac{1}{2}}\right)\left(\hat{m}_{i\beta}^{\mathcal{C}} - e_{i\beta}^n\right), \tag{3.13}$$

since $\boldsymbol{e}_i^{n+\frac{1}{2}} \cdot \left(\boldsymbol{e}_i^{n+1} - \boldsymbol{e}_i^n\right) = 0$. Clearly, the left-hand side should be replaced by $\dot{\boldsymbol{e}}_i \equiv \mathrm{d}\boldsymbol{e}_i/\mathrm{d}t$ in the limit $\Delta t \to 0$. The right-hand side, however, is not well defined in this limit. To overcome this, one makes the substitution $1/\Delta t \to \nu$ to prevent it from blowing up, leading in the $\Delta t \to 0$ limit to the equation

$$\dot{e}_{i\alpha} = \nu\left(\delta_{\alpha\beta} - e_{i\alpha}e_{i\beta}\right)\hat{m}_{i\beta}^{\mathcal{C}}. \tag{3.14}$$

In this chapter we treat ν as a constant, although one may also consider different choices since the above derivation leaves it unspecified. In fact, in chapter 4, we will consider a model which may be interpreted as having a magnetisation dependent $\nu\left(m_i^{\mathcal{C}}\right) \propto m_i^{\mathcal{C}}$. In any case, ν has the dimension of inverse time, and the *response time* $t_\nu \sim \nu^{-1}$ can be interpreted as the typical time it takes a particle to align with its neighbour. This can also be considered the natural time-step for the dynamics, which should remain finite even as $\Delta t \to 0$ as was implemented in eq. (3.14).

At this stage, we reintroduce noise ad hoc as a spherical diffusion (see chapter 2) with diffusivity D_r, leading to the continuous-time equations

$$\dot{\boldsymbol{x}}_i = v_0 \boldsymbol{e}_i, \tag{3.15}$$

$$\dot{e}_{i\alpha} = \left(\delta_{\alpha\beta} - e_{i\alpha}e_{i\beta}\right) \circ \left(v\hat{m}^{\mathcal{C}}_{i\beta} + \sqrt{2D_r}\xi_{i\beta}\right), \tag{3.16}$$

where $\circ$ denotes that integration should be performed in the Stratonovich sense (see chapter 2). Since the spins $\boldsymbol{e}_i$ are all of unit length, eq. (3.16) can be reduced by a variable transformation to $d-1$ independent equations. Specifically, in $d=2$, we obtain for $\boldsymbol{e}_i = \boldsymbol{e}(\theta_i) \equiv (\cos\theta_i, \sin\theta_i)^T$ that

$$\dot{\theta}_i = v\boldsymbol{e}_i' \cdot \hat{\boldsymbol{m}}^{\mathcal{C}}_i + \sqrt{2D_r}\xi_i, \tag{3.17}$$

where we denote by $\boldsymbol{e}'(\theta) = \mathrm{d}\boldsymbol{e}/\mathrm{d}\theta$, and the $\{\xi_i\}$ are, with slight abuse of notation, independent unit 1-dimensional white noise processes. We refer to eqs. (3.15) and (3.16) (alternatively eq. (3.17) in $d=2$) with an isotropic contact kernel as the **continuous-time Vicsek model** (CTVM).

From here on we restrict to $d=2$ unless explicitly stated otherwise. It will also be convenient to nondimensionalise the equations of motion by measuring length and time in units of the effective diffusion length scale $\ell_D \sim v_0/D_r$ and orientational correlation time $t_D \sim D_r^{-1}$ (see chapter 2) respectively. With this choice, the dynamics can be written

$$\dot{\boldsymbol{x}}_i = \boldsymbol{e}_i, \tag{3.18}$$

$$\dot{\theta}_i = \beta\boldsymbol{e}_i' \cdot \hat{\boldsymbol{m}}^{\mathcal{C}}_i + \sqrt{2}\xi_i, \tag{3.19}$$

where we have defined the 'inverse temperature' $\beta = v/D_r$ (and correspondingly the 'temperature' $T = \beta^{-1}$), which measures the competition between alignment response and orientational decorrelation due to rotational diffusion. The reason we make this choice, rather than choosing t_v as our timescale, is that many of the calculations we perform in this chapter will be in the vicinity of the critical point at infinite density and $\beta = 0$ (see section 3.3 and fig. 3.4).

The CTVM is straightforward to implement numerically, and the phenomenology we observe is identical to that which is reported for the VM. Specifically, decreasing T at fixed average density $\rho_0 = N/L^2$, we first observe a gaseous phase at high $T > T_g(\rho_0)$, MPS in the miscibility gap $T_l(\rho_0) < T < T_g(\rho_0)$ and the PL phase for low $T < T_l(\rho_0)$, where T_g and T_l are the gas and liquid binodals respectively. Similar behaviour is observed when we increase ρ_0 at fixed T. As for the VM, the dynamics in the miscibility gap display traveling

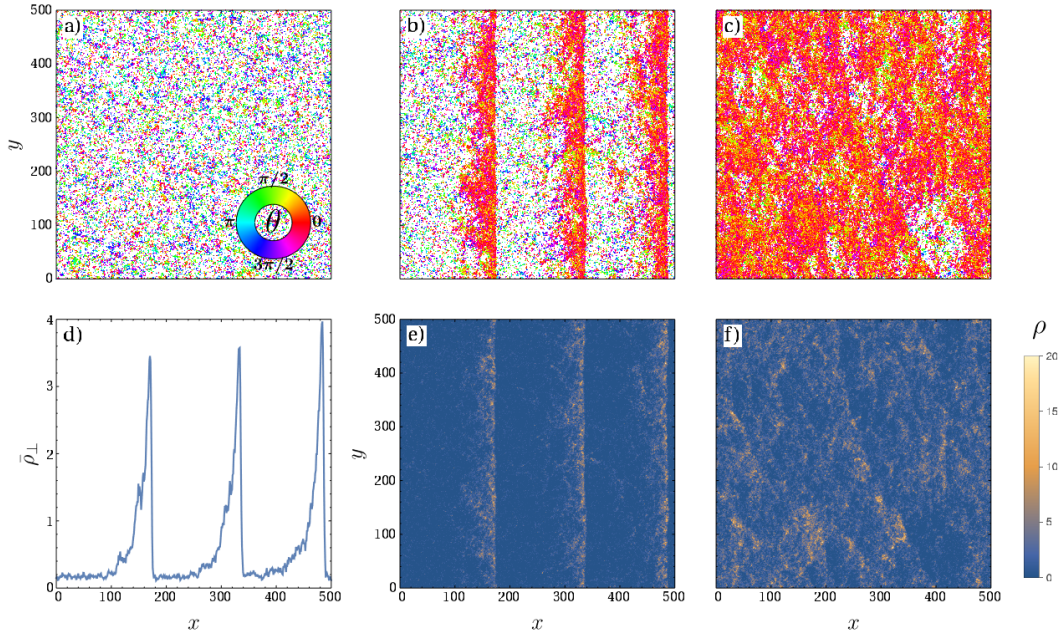


Fig. 3.3 Phases of the CTVM: a) isotropic ($\rho_0 = 0.2$), b) MPS ($\rho_0 = 0.6$) and c) PL ($\rho_0 = 1.6$) phases. In b), particles organise into a smectic arrangement of high-density bands traveling against an isotropic background with a well-defined gaseous density, while in c) the liquid phase fills the simulation domain. Figures e) and f) are, respectively, the density profiles from b) and c), while in d) the density profile from b) is averaged over the spatial direction perpendicular to the motion (denoted by $\bar{\rho}_\perp$). In all three simulations, the isotropic kernel in eq. (3.7) was used with $L = 500$, $\ell_0 = 1$, $v_0 = 1$ and $D_r = 0.2$. Colors indicate the angle θ_i each spin makes relative to the x-axis.

bands, while in the PL region, the liquid phase fills the simulation domain. In fig. 3.3 we show typical snapshots from simulations, including the gaseous, MPS and PL phases. The resulting phase diagram is illustrated in fig. 3.4.

The following sections of this chapter will analyse the dynamics in eqs. (3.18) and (3.19) more closely, and in particular the *mean-field theory* (MFT) of the CTVM which is exact at infinite density. The idea behind this approach is that the structure of the phase diagram of the CTVM in the ρ-T plane is constrained by the limiting dynamics at the boundaries $T = \infty$ and $\rho_0 = \infty$, which should connect continuously with the remainder of the phase diagram. While the dynamics at $T = \infty$ (i.e. $\beta = 0$) is trivial and simply gives an isotropic gas, the MFT at $\rho_0 = \infty$ requires more work. We will see that by solving the MFT, we can find the steady-state homogeneous PL distribution, and the location of the critical point where the binodals T_g and T_l meet and the transition becomes second order. Moreover, by

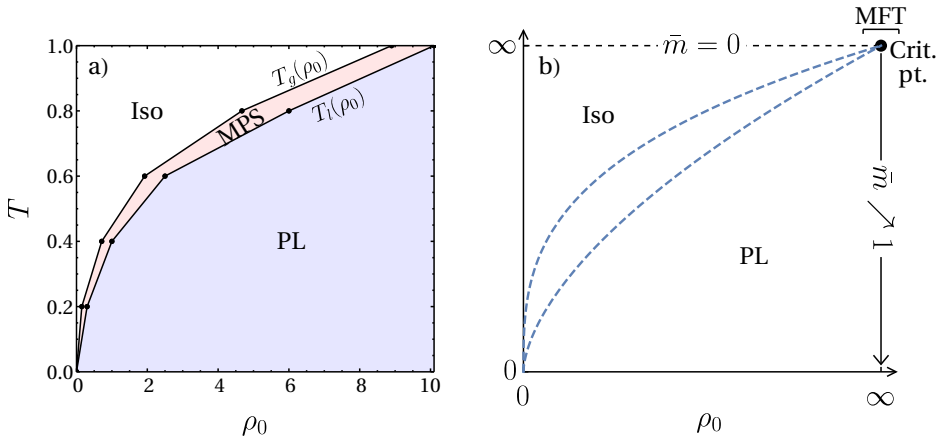


Fig. 3.4 Phase diagram of the CTVM. In a), the phase diagram in the ρ_0-T plane obtained from simulations is plotted, indicating the location of the isotropic, MPS and PL phases. Note that there is no PL phase as $\rho_0 \to 0$, since the system is always isotropic in this limit. $T_g(\rho_0)$ and $T_l(\rho_0)$ denote the gas and liquid binodals respectively. Figure b) is a schematic of the phase diagram that can be inferred from solving the model at the boundaries $T = \infty$ and $\rho_0 = \infty$ (MFT). The prediction of MFT is that: at $\rho_0 = \infty$ the PL phase is stable for all $T > 0$, the critical point should be located at $(\rho_c, T_c) = (\infty, \infty)$, and the order parameter should vary continuously from $0 \to 1$ as T goes from $\infty \to 0$. The shapes of the binodals T_g and T_l (illustrated with blue dashed lines in b)) cannot be inferred from this approach alone and are only drawn for illustrational purposes.

identifying the steady-state PL distribution, we are able to show that it is indeed linearly stable at $\rho_0 = \infty$ for all $T > 0$. The resulting schematic phase diagram obtained from this approach is illustrated in fig. 3.4. Importantly, however, by taking the infinite density limit, the density ρ is *replaced* by a single-particle marginal distribution ϱ. Thus, it is clear from the outset that one *cannot* deduce the structure of the liquid and gas binodals T_l and T_g from this approach alone, although from symmetry reasons there has to be a transition line connecting the origin with the critical point that separates the isotropic and polarly ordered phases at $T = \infty$ and $\rho_0 = \infty$ (respectively). In particular, an analytical argument that correctly predicts the limiting behaviour $T_\alpha \sim \rho_0^{z\alpha}$ for $\alpha = g, l$ and $\rho_0 \gg 1$ in the CTVM still remains elusive.

3.3 Mean-field theory of the CTVM

In this section we construct the MFT of the CTVM at infinite density. When this limit is taken, finite-density fluctuations are suppressed and the dynamics on the hydrodynamic scale becomes deterministic. This has allowed mathematicians to derive an exact hydrodynamic scaling limit for the dynamics of the CTVM at $\rho_0 = \infty$ by constructing the time-dependent solution to a *non-linear* Fokker-Planck equation (FPE), presented below, using a Chapman-

Enskog type expansion and generalised collisional invariants [79, 81, 83, 105]. Here, we only highlight that the steady-state solution to the same FPE may be found without much trouble. This then allows us to compute the order parameter at $\rho_0 = \infty$, along with a microscopic single-particle steady-state dynamics and its statistics.

Before we embark on the calculation at $\rho_0 = \infty$, we first make a note regarding the interaction term in eq. (3.19) and another, although related, 'mean-field limit'. The interaction in the CTVM is non-integrable *in general*, and so from this fact alone the steady-state distribution will not be the Boltzmann distribution for most choices of $\mathcal{C}$. To see this, we let $F_i = \beta \boldsymbol{e}'_i \cdot \hat{\boldsymbol{m}}^{\mathcal{C}}_i$ and observe that for $i \neq j$,

$$\partial_{\theta_j} F_i - \partial_{\theta_i} F_j = \beta \mathcal{C}_{ij} \left(\frac{\delta_{\alpha\beta} - \hat{m}^{\mathcal{C}}_{j\alpha} \hat{m}^{\mathcal{C}}_{j\beta}}{m^{\mathcal{C}}_j} - \frac{\delta_{\alpha\beta} - \hat{m}^{\mathcal{C}}_{i\alpha} \hat{m}^{\mathcal{C}}_{i\beta}}{m^{\mathcal{C}}_i} \right) e'_{i\alpha} e'_{j\beta}. \tag{3.20}$$

For general $\mathcal{C}_{ij}$, the right-hand side does not vanish, meaning that F_i cannot be written as the derivative of a potential. However, a particular example for which the right-hand side of eq. (3.20) *does* vanish is provided by the case of a uniform contact matrix $\mathcal{C}_{ij} = c$, where c is a constant, corresponding to the limit $\ell_0 \to \infty$ for which all particles interact with all other particles. In this case we may write $F_i = -\partial_{\theta_i} u$, and one may straightforwardly show that

$$u = -\beta \hat{\boldsymbol{m}}^{\mathcal{C}} \cdot \sum_{i=1}^{N} \boldsymbol{e}_i, \tag{3.21}$$

where $\boldsymbol{m}^{\mathcal{C}} = c \sum_i \boldsymbol{e}_i$. The steady-state distribution is then simply the Boltzmann distribution, and can be written as

$$p_{\mathrm{ss}}(\{\theta_i\}) = \mathcal{Z}^{-1} \prod_{i=1}^{N} \mathrm{e}^{\beta \boldsymbol{e}_i \cdot \hat{\boldsymbol{m}}^{\mathcal{C}}}, \tag{3.22}$$

where the partition function $\mathcal{Z}$ normalizes p_{ss} to a distribution.

It is not true, however, that the CTVM dynamics become equilibrium in the limit '$v_0 \to 0$', for which the isotropic kernel becomes time-independent, i.e. $\mathcal{C}_{ij} = \mathcal{C}_0(|\boldsymbol{x}_i - \boldsymbol{x}_j|) \to c_{ij}$ for constant c_{ij}, which can be seen from eq. (3.20). The crucial point here, which also holds true more broadly for the models we discuss in this thesis, is that entropy production (or irreversibility) in the CTVM is more closely related to spatial variations in the system than merely the fact that the contact matrix is time-dependent. When $\mathcal{C}$ is uniform as above, $\boldsymbol{m}^{\mathcal{C}}_i = \boldsymbol{m}^{\mathcal{C}}_j \equiv \boldsymbol{m}^{\mathcal{C}}$ for all i, j, and so the particles do not detect any spatial fluctuations in the local magnetisation, leading to equilibrium dynamics. This will all be discussed in much greater detail in chapters 5 and 6, however in what follows now we will see that in a

homogeneous flock (but not in general) the interaction does indeed become integrable and the dynamics effectively equilibrium with a Boltzmann steady-state distribution.

To derive the MFT of the CTVM, we follow [79, 81] and first define the one-particle distribution function

$$p(\boldsymbol{x},\theta) = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \delta(\boldsymbol{x}-\boldsymbol{x}_i)\delta(\theta-\theta_i). \tag{3.23}$$

It may then be proven that in the limit $N \to \infty$,

$$\hat{\boldsymbol{m}}_i^{\mathcal{C}} \to \hat{\boldsymbol{m}}^{\mathcal{C}}(\boldsymbol{x}_i) \equiv \frac{\int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \int_0^{2\pi} \mathrm{d}\theta\, p(\boldsymbol{x},\theta)\mathcal{C}_0(|\boldsymbol{x}-\boldsymbol{x}_i|)\boldsymbol{e}(\theta)}{\left|\int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \int_0^{2\pi} \mathrm{d}\theta\, p(\boldsymbol{x},\theta)\mathcal{C}_0(|\boldsymbol{x}-\boldsymbol{x}_i|)\boldsymbol{e}(\theta)\right|}, \tag{3.24}$$

where the molecular field $\hat{\boldsymbol{m}}^{\mathcal{C}}(\boldsymbol{x})$ depends continuously on the position $\boldsymbol{x}$. Thus, at $N=\infty$, the dynamics in eqs. (3.15) and (3.17) is decoupled and each particle obeys the mean-field dynamics $\dot{\boldsymbol{x}}_i = \boldsymbol{e}_i$ with

$$\dot{\theta}_i = \beta \boldsymbol{e}_i' \cdot \hat{\boldsymbol{m}}^{\mathcal{C}}(\boldsymbol{x}_i) + \sqrt{2}\xi_i. \tag{3.25}$$

Notice in particular that the interaction term now has a type of involutional dependency on the distribution p that the dynamics generate. This means that in order to generate realisations of eq. (3.25), one would first need to compute p.

Following the standard procedure detailed in chapter 2, one may compute the distribution p by solving the FPE

$$\partial_t p + \boldsymbol{e}\cdot\nabla p = \partial_\theta\left(\mathrm{e}^{\beta \boldsymbol{e}\cdot\hat{\boldsymbol{m}}^{\mathcal{C}}}\partial_\theta\left(\mathrm{e}^{-\beta \boldsymbol{e}\cdot\hat{\boldsymbol{m}}^{\mathcal{C}}}p\right)\right), \tag{3.26}$$

associated with the dynamics in eq. (3.25). However, unlike the standard FPEs we treated then, this equation is in fact non-linear, which is a consequence of the relation between $\hat{\boldsymbol{m}}^{\mathcal{C}}(\boldsymbol{x})$ and the distribution p in eq. (3.24). Because of this, we will refer to it as the *non-linear* FPE associated with the mean-field dynamics in eq. (3.25).

Notwithstanding, eq. (3.26) in fact has an exact steady-state solution for any suitably well-behaved isotropic kernel $\mathcal{C}_0$. Indeed, for any constant $\hat{\boldsymbol{M}}$ with $|\hat{\boldsymbol{M}}| = 1$, it is given by

$$p_{\mathrm{ss}}(\boldsymbol{x},\theta) = \frac{\mathrm{e}^{\beta \boldsymbol{e}\cdot\hat{\boldsymbol{M}}}}{2\pi L^2 I_0(\beta)}, \tag{3.27}$$

where I is the modified Bessel function (of the first kind), defined here for integer order $n \in \mathbb{Z}$ by (see e.g. [106])

$$I_n(z) = \frac{1}{\pi}\int_0^{\pi} \mathrm{d}\theta\, \mathrm{e}^{z\cos\theta}\cos(n\theta). \tag{3.28}$$

To see this, we simply compute

$$\hat{\boldsymbol{m}}^{\mathcal{C}}_{\mathrm{ss}}(\boldsymbol{x}) = \frac{\int_{\mathcal{V}} \mathrm{d}\boldsymbol{y} \int_0^{2\pi} \mathrm{d}\theta \, p_{\mathrm{ss}}(\boldsymbol{y},\theta)\mathcal{C}_0(|\boldsymbol{y}-\boldsymbol{x}|)\boldsymbol{e}(\theta)}{\left| \int_{\mathcal{V}} \mathrm{d}\boldsymbol{y} \int_0^{2\pi} \mathrm{d}\theta \, p_{\mathrm{ss}}(\boldsymbol{y},\theta)\mathcal{C}_0(|\boldsymbol{y}-\boldsymbol{x}|)\boldsymbol{e}(\theta) \right|} \tag{3.29}$$

$$= \hat{\boldsymbol{M}}, \tag{3.30}$$

from which it trivially follows that both sides of eq. (3.26) vanish.

Combining eqs. (3.25) and (3.30) it thus follows that the mean-field dynamics in steady-state at infinite density is given by the equilibrium Langevin equation

$$\dot{\theta}_i = -u'_{\mathrm{ss}}(\theta_i) + \sqrt{2}\xi_i, \tag{3.31}$$

where the potential function $u_{\mathrm{ss}} = -\beta \boldsymbol{e} \cdot \hat{\boldsymbol{M}}$. In particular, as stated above, the dynamics of a homogeneous flock in the CTVM is indeed effectively equilibrium, with the Boltzmann steady-state distribution in eq. (3.27).

With eq. (3.27) at hand, the properties of the CTVM at $\rho_0 = \infty$ can be studied more closely. From it, it follows straightforwardly that the mean global magnetisation $\bar{m}$ in eq. (3.5) is given by

$$\bar{m} = \left| \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \int_0^{2\pi} \mathrm{d}\theta \, p_{\mathrm{ss}}(\boldsymbol{x},\theta)\boldsymbol{e}(\theta) \right| \tag{3.32}$$

$$= \frac{I_1(\beta)}{I_0(\beta)}. \tag{3.33}$$

In particular, for large temperatures as $\beta \to 0$,

$$\bar{m} = \frac{\beta}{2} - \frac{\beta^3}{16} + \mathcal{O}\left(\beta^5\right), \tag{3.34}$$

meaning that at the critical point $\beta_c = 0$, the magnetisation $\bar{m} \sim \beta$ predicting a continuous transition as expected. In the opposite limit, $\beta \to \infty$, it follows that $\bar{m} \to 1$. This is illustrated in fig. 3.4.

From the mean-field steady-state dynamics in eq. (3.31) we can also calculate the time-correlation function $C(t) = \langle \boldsymbol{e}_i(0) \cdot \boldsymbol{e}_i(t) \rangle - \langle \boldsymbol{e}_i(0) \rangle^2$ near β_c, and we find that it is given by (see appendix B for the details)

$$C(t) = \mathrm{e}^{-t} \left(1 + \frac{\beta^2}{36} \left(\mathrm{e}^{-3t} - 10 - 6t \right) + \mathcal{O}(\beta^4) \right). \tag{3.35}$$

In a similar vein, we may calculate how the effective diffusion (see chapter 2) is modified due to the forcing term, and we find that

$$D_{\text{eff}}(\beta) = \lim_{\tau\to\infty} \frac{1}{2\tau} \left\langle (\Delta \boldsymbol{x} - \langle \Delta \boldsymbol{x} \rangle)^2 \right\rangle \tag{3.36}$$

$$= 1 - \frac{7}{16}\beta^2 + \mathcal{O}(\beta^4), \tag{3.37}$$

where $\Delta \boldsymbol{x} = \boldsymbol{x}(\tau) - \boldsymbol{x}(0)$. Both results in eqs. (3.35) and (3.37) will be useful to our subsequent discussion of the EPR of the CTVM in chapter 5.

Having now set up the MFT, i.e. the dynamics in eq. (3.25) and the FPE in eq. (3.26), we turn onto a more in-depth analysis of the dynamics on the hydrodynamic scale in this regime. Our primary goal here will be to show that the homogeneous PL distribution in eq. (3.27) is indeed linearly stable. This result is important, because it will serve as a useful benchmark against which to test our findings in chapter 4. Moreover, by introducing the machinery we use here, we can also focus our subsequent discussion in chapter 4. Similar methods to those we will present now have been applied to analogous systems previously [92], although to our knowledge not for the CTVM.

3.4 Linear stability analysis of the homogeneous polar liquid

As described in the previous section, we proceed our analysis of the MFT of the CTVM by investigating the linear stability of the PL steady-state solution in eq. (3.27) to the FPE in eq. (3.26). By showing that this is indeed linearly stable for all $T > 0$, we provide analytical arguments to support that the PL is the true steady-state distribution of the model at $\rho_0 = \infty$, which is essential in order to construct the schematic phase diagram in fig. 3.4. Moreover, from this we develop a method which can be applied again to study similar models in chapter 4.

We first pick $\hat{\boldsymbol{M}} = \hat{\boldsymbol{x}}$ to be along the x-axis, which we may do without loss of generality. Then, linearising the FPE about p_{ss} by writing $p = p_{\text{ss}} + \delta p$ and keeping terms only to linear order in δp, we straightforwardly obtain

$$\partial_t \delta p + \boldsymbol{e} \cdot \nabla \delta p = \partial_\theta \left(\partial_\theta \delta p + \beta \sin\theta \delta p - L^2 \beta \frac{I_0(\beta)}{I_1(\beta)} p_{\text{ss}} \cos\theta \delta m_\perp^{\mathcal{C}} \right). \tag{3.38}$$

In eq. (3.38), we have defined the variation $\delta\boldsymbol{m}^{\mathcal{C}} = \left(\delta m^{\mathcal{C}}_{\parallel}, \delta m^{\mathcal{C}}_{\perp}\right)$ in the molecular field $\boldsymbol{m}^{\mathcal{C}}$, where $\parallel$ and $\perp$ denote the components that are parallel and perpendicular to $\hat{\boldsymbol{M}}$ respectively, as that produced by the perturbation δp, i.e.

$$\delta\boldsymbol{m}^{\mathcal{C}}(\boldsymbol{x}) = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}' \int_0^{2\pi} \mathrm{d}\theta\, \mathcal{C}_0(|\boldsymbol{x}-\boldsymbol{x}'|)\delta p(\boldsymbol{x}',\theta)\boldsymbol{e}(\theta). \tag{3.39}$$

To analyse eq. (3.38), we will proceed by first transforming it to Fourier space in θ. Letting $\delta p_n = \int_0^{2\pi} \mathrm{d}\theta\, \delta p\, \mathrm{e}^{-in\theta}$, it is straightforward to show that

$$\partial_t \delta p_n + \frac{1}{2}\nabla\cdot\begin{pmatrix} \delta p_{n-1} + \delta p_{n+1} \\ i(\delta p_{n+1} - \delta p_{n-1}) \end{pmatrix} = -n^2\delta p_n + \frac{\beta n}{2}(\delta p_{n-1} - \delta p_{n+1}) - c_n(\beta)\delta m^{\mathcal{C}}_{\perp}, \tag{3.40}$$

where the constants c_n are defined by

$$c_n(\beta) = \frac{in\beta}{2\pi I_1(\beta)} \int_0^{2\pi} \mathrm{d}\theta\, \cos\theta\, \mathrm{e}^{\beta\cos\theta - in\theta}. \tag{3.41}$$

The integral in eq. (3.41) can be recognised as a modified Bessel integral, and thus be performed exactly. Using standard tools (see e.g. [106]), one then finds that

$$c_n(\beta) = in\beta\frac{I_n'(\beta)}{I_1(\beta)}. \tag{3.42}$$

Equation (3.40) is more easily handled by considering the real and imaginary parts of the Fourier coefficients δp_n separately, and we denote these by δp_n^R and δp_n^I respectively. After separating the real and imaginary parts of δp_n, we also perform a Fourier transform in the spatial variable $\boldsymbol{x}$, which leads to the coupled set of equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta p^R_{n,\boldsymbol{q}} + \frac{i\boldsymbol{q}}{2}\cdot\begin{pmatrix} \delta p^R_{n-1,\boldsymbol{q}} + \delta p^R_{n+1,\boldsymbol{q}} \\ \delta p^I_{n-1,\boldsymbol{q}} - \delta p^I_{n+1,\boldsymbol{q}} \end{pmatrix} = -n^2\delta p^R_{n,\boldsymbol{q}} + \frac{\beta n}{2}\left(\delta p^R_{n-1,\boldsymbol{q}} - \delta p^R_{n+1,\boldsymbol{q}}\right), \tag{3.43}$$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\delta p^I_{n,\boldsymbol{q}} + \frac{i\boldsymbol{q}}{2}\cdot\begin{pmatrix} \delta p^I_{n-1,\boldsymbol{q}} + \delta p^I_{n+1,\boldsymbol{q}} \\ \delta p^R_{n+1,\boldsymbol{q}} - \delta p^R_{n-1,\boldsymbol{q}} \end{pmatrix} &= -n^2\delta p^I_{n,\boldsymbol{q}} + \frac{\beta n}{2}\left(\delta p^I_{n-1,\boldsymbol{q}} - \delta p^I_{n+1,\boldsymbol{q}}\right) \\ &\quad - n\beta\frac{I_n'(\beta)}{I_1(\beta)}\delta m^{\mathcal{C}}_{\perp,\boldsymbol{q}},\end{aligned} \tag{3.44}$$

where the Fourier coefficients $\delta p^{\alpha}_{n,\boldsymbol{q}} = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \delta p^{\alpha}_n\, \mathrm{e}^{-i\boldsymbol{q}\cdot\boldsymbol{x}}$ for $\alpha = R, I$ (and similarly for $\delta\boldsymbol{m}^{\mathcal{C}}_{\boldsymbol{q}}$).

Now, observe that using eq. (3.39) we may write

$$\delta \boldsymbol{m}_{\boldsymbol{q}}^{\mathcal{C}} = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}' \int_0^{2\pi} \mathrm{d}\theta\, \mathcal{C}_0(|\boldsymbol{x}-\boldsymbol{x}'|)\delta p(\boldsymbol{x}',\theta)\boldsymbol{e}(\theta)\mathrm{e}^{-i\boldsymbol{q}\cdot\boldsymbol{x}} \tag{3.45}$$

$$= \mathcal{C}_{0,\boldsymbol{q}} \begin{pmatrix} \delta p_{1,\boldsymbol{q}}^R \\ -\delta p_{1,\boldsymbol{q}}^I \end{pmatrix}, \tag{3.46}$$

where we have defined, with slight abuse of notation, the Fourier coefficients $\mathcal{C}_{0,\boldsymbol{q}} = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\mathcal{C}_0(|\boldsymbol{x}|)\,\mathrm{e}^{-i\boldsymbol{q}\cdot\boldsymbol{x}}$. Using eq. (3.9) we may further express the coefficients $\mathcal{C}_{0,\boldsymbol{q}}$ in the form of a Hankel transform

$$\mathcal{C}_{0,\boldsymbol{q}} = 2\int_0^1 \mathrm{d}r\, r\phi(r)J_0(\ell_0 q r), \tag{3.47}$$

where J is the Bessel function of the first kind, defined here for integer orders $n \in \mathbb{Z}$ by (see e.g. [106])

$$J_n(z) = \frac{1}{\pi}\int_0^\pi \mathrm{d}\theta\, \cos\left(z\sin\theta - n\theta\right). \tag{3.48}$$

Equation (3.47) simplifies the evaluation of $\mathcal{C}_{0,\boldsymbol{q}}$ for certain choices of ϕ. For example, taking $\phi(r) = \Theta(1-r)$, we find that

$$\mathcal{C}_{0,\boldsymbol{q}} = \frac{2}{\ell_0 q}J_1(\ell_0 q). \tag{3.49}$$

Equations (3.43) and (3.44), combined with eq. (3.46), take the form of a coupled linear system in $\delta p_{n,\boldsymbol{q}}^R$ and $\delta p_{n,\boldsymbol{q}}^I$ for $n \geq 0$. Specifically, we can write

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta \boldsymbol{p}_{\boldsymbol{q}} = \mathcal{L}_{\boldsymbol{q}}\delta \boldsymbol{p}_{\boldsymbol{q}}, \tag{3.50}$$

where $\delta \boldsymbol{p}_{\boldsymbol{q}} = (\delta p_{0,\boldsymbol{q}}^R, \delta p_{1,\boldsymbol{q}}^R, \delta p_{1,\boldsymbol{q}}^I, \ldots)^T$ (note that $\delta p_{0,\boldsymbol{q}}^I = 0$ by construction) and $\mathcal{L}_{\boldsymbol{q}}$ is an infinite matrix. From this it is apparent that the linear stability properties of PL distribution p_{ss} can be determined from assessing the spectrum of the linear operator $\mathcal{L}_{\boldsymbol{q}}$. Specifically, we are interested in the maximal eigenvalue $\lambda_m \equiv \lambda_m(T,\boldsymbol{q})$ of $\mathcal{L}_{\boldsymbol{q}}$ and its behavior as a function of temperature T and the wave vector $\boldsymbol{q}$. If this is always negative, then fluctuations $\delta \boldsymbol{p}_{\boldsymbol{q}}$ will decay in time, meaning that the homogeneous PL is linearly stable.

Since the linearised FPE in eq. (3.38) is not diagonalised in the Fourier basis, computing the spectrum of $\mathcal{L}_{\boldsymbol{q}}$ is, however, still a challenging task analytically. On the other hand, we can fairly straightforwardly perform some numerical checks. Indeed, by introducing an ultraviolet cutoff Λ for the angular modes and defining $\delta \boldsymbol{p}_{\boldsymbol{q}}^{\Lambda} = (\delta p_{0,\boldsymbol{q}}^R, \delta p_{1,\boldsymbol{q}}^R, \delta p_{1,\boldsymbol{q}}^I, \ldots, p_{\Lambda,\boldsymbol{q}}^R, \delta p_{\Lambda,\boldsymbol{q}}^I)^T$ as the truncated vector of angular Fourier coefficients with $|n| \leq \Lambda$, we can study the

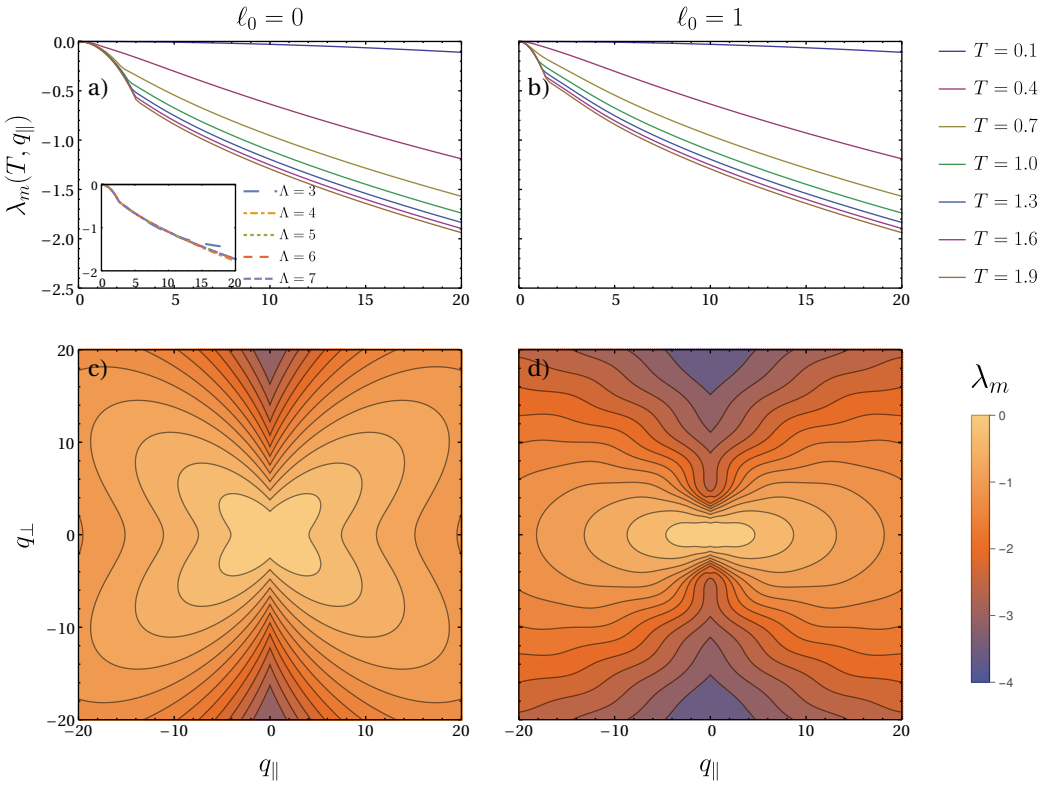


Fig. 3.5 The maximal eigenvalue λ_m of the matrix $\mathcal{L}^{\Lambda}_{\boldsymbol{q}}$ as a function of: a) and b) wave number $q_\parallel$ for different temperatures T at $q_\perp = 0$, and, c) and d) $\boldsymbol{q} = (q_\parallel, q_\perp)$ at fixed temperature $T = 0.4$, where $\parallel$ and $\perp$ denote components parallel and perpendicular to $\hat{\boldsymbol{M}} = \hat{\boldsymbol{x}}$ respectively. In a) and c), the limit of purely local interactions $\ell_0 \to 0$ has been taken, while in b) and d) the same plots are shown for finite $\ell_0 = 1$. In a), the inset shows that convergence to the curves $\lambda_m(T, q_\parallel)$ with the ultraviolet cutoff Λ is quite rapid, and only the curve for $\Lambda = 3$ deviates slightly in the range of $q_\parallel$ shown. For all other plots, the cutoff was set to $\Lambda = 20$.

corresponding finite-dimensional problem

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta \boldsymbol{p}_{\boldsymbol{q}}^{\Lambda} = \mathcal{L}_{\boldsymbol{q}}^{\Lambda}\delta \boldsymbol{p}_{\boldsymbol{q}}^{\Lambda}. \tag{3.51}$$

In particular, the spectrum of $\mathcal{L}_{\boldsymbol{q}}^{\Lambda}$ should coincide with that of $\mathcal{L}_{\boldsymbol{q}}$ in the limit $\Lambda \to \infty$.

In fig. 3.5 we plot the maximal eigenvalue λ_m obtained from calculating the eigenvalues of $\mathcal{L}_{\boldsymbol{q}}^{\Lambda}$ in eq. (3.51) numerically as a function of $\boldsymbol{q}$ for various values of the temperature T. We do this using the isotropic kernel in eq. (3.49) both for finite ℓ_0, and in the limit $\ell_0 \to 0$ where the interactions become perfectly local and $\mathcal{C}_{0,\boldsymbol{q}} \to 1$. In all cases, we find that λ_m is always strictly negative, indicating that the polar liquid is indeed stable for all T. In particular, we do not observe any significant differences in the behaviour of λ_m for finite ℓ_0 or $\ell_0 = 0$. We also observe that that the maximal eigenvalue is a monotonically decreasing function of T within the ranges of $\boldsymbol{q}$ investigated, with $\lambda_m \to 0$ as $T \to 0$, although we are still uncertain whether this has any real physical implications.

A direction we could have pursued at this point, which we intend to do in future work, is to investigate directly the full non-linear FPE, e.g. using numerical tools. Another is to follow the approach of Dean [107], which also captures the finite-density fluctuations by including a noise term in the FPE. For the CTVM, we show in appendix E that the *Dean equation* (DE) takes the form

$$\partial_t p + \boldsymbol{e}\cdot\nabla p = \partial_\theta\left(\partial_\theta p - \beta \boldsymbol{e}'\cdot\hat{\boldsymbol{m}}^{\mathcal{C}} p + \sqrt{2p/N}\,\Xi\right), \tag{3.52}$$

where $\Xi \equiv \Xi(t,\boldsymbol{x},\theta)$ is a mean-zero Gaussian white noise with unit covariance, i.e.

$$\left\langle \Xi(t,\boldsymbol{x},\theta)\Xi(t',\boldsymbol{x}',\theta')\right\rangle = \delta(t-t')\delta(\boldsymbol{x}-\boldsymbol{x}')\delta(\theta-\theta'). \tag{3.53}$$

Equation (3.52) is an equation for the now *fluctuating* empirical one-particle distribution $p(\boldsymbol{x},t) = N^{-1}\sum_{i=1}^{N}\delta(\boldsymbol{x}-\boldsymbol{x}_i)\delta(\theta-\theta_i)$, and can be constructed using Itô's lemma (see chapter 2). From such an approach, one might hope to recover the limiting expressions for the binodals $T_g(\rho_0)$ and $T_l(\rho_0)$. In fact, we will revisit the DE in chapter 5, where we seek a solution as an expansion in large N.

On the other hand, achieving analytical progress with eq. (3.52) beyond linear order will likely be rather difficult, since going beyond linear order even in the deterministic FPE in eq. (3.26) requires substantial work [81]. In section 3.5, we show that one may overcome this by mapping the microscopic dynamics onto a 'Toner-Tu type' hydrodynamic theory by following a standard method from the literature, first introduced in [82]. In addition to

being more accessible with analytical tools, including noise can also be done as for the DE, although this will be added ad-hoc.

3.5 Constructing a hydrodynamic theory from the mean-field CTVM

Of the many hydrodynamic theories that have been derived for the microscopic Vicsek type models, most bear resemblance with that considered initially by Toner and Tu on phenomenological grounds [18, 19]. Here, we will construct an explicit map from the non-linear FPE in eq. (3.26) onto a 'Toner-Tu type' hydrodynamic theory by considering the relaxational dynamics of small perturbations about the steady-state in the vicinity of the critical point at $T_c = \infty$. We stress, however, that all the hydrodynamic equations we derive in this section (and in later sections, see sections 4.4 and 4.5), live on the line $\rho_0 = \infty$ and so we do not expect that they can inform us about band formation. The procedure we follow is standard in the literature [1, 82], and we have only adopted this here for the CTVM. Moreover, exact scaling limits of the CTVM at infinite density have been previously derived from the FPE in the mathematical literature [81, 105], however the dynamical equations they obtain are difficult to analyse with direct tools. The approach we present here is only approximate, although despite this we expect to obtain quantitatively good predictions near T_c.

There are many reasons why we would like to pursue this approach, including introducing the type of hydrodynamic theories usually employed to study the hydrodynamics of flocking models. Subsequent to chapter 4, we will have seen that all the microscopic models of flocking considered in this thesis have effective hydrodynamic descriptions that share many characteristics. This serves to motivate our studies in chapter 6 of the EPR of flocking on the hydrodynamic scale. Furthermore, the theory where only the coefficient of the term that controls the transition to the low-temperature phase is non-constant, can be shown to display all ordered phases that we encounter in this thesis (see chapter 4). Finally, it provides us with a route for making quantitative analytical predictions that are difficult to obtain directly from the non-linear FPE.

We will delay the introduction of the more general class of hydrodynamic Toner-Tu theories we consider until we have completed the derivation of the mapping from the CTVM. Thus, we begin our analysis by introducing the relevant hydrodynamic fields, which will be

the marginal density ϱ and first moment $\boldsymbol{m}$ of the distribution p. Explicitly, we define

$$\varrho(\boldsymbol{x}) = \int_0^{2\pi} \mathrm{d}\theta\, p(\boldsymbol{x},\theta), \tag{3.54}$$

$$\boldsymbol{m}(\boldsymbol{x}) = \int_0^{2\pi} \mathrm{d}\theta\, p(\boldsymbol{x},\theta)\boldsymbol{e}(\theta). \tag{3.55}$$

Importantly, the marginal ϱ should not be confused with the density ρ, since the latter is infinite. Moreover, at infinite density, we do not observe fluctuations due to noise even in infinitesimal boxes of volume $\mathrm{d}\mathcal{V} = \mathrm{d}x\mathrm{d}y$. On the other hand, we can imagine introducing a long wavelength perturbation and study the relaxation of the fields. We will therefore consider introducing a small amplitude perturbation at finite $\beta \ll 1$ of wavelength ℓ_p that is large compared with the microscopic diffusion length scale $\ell_D = 1$ and interaction length ℓ_0, and study the relaxation on timescales large compared with the correlation time $t_D = 1$.

To make progress we use eq. (3.9) to write $\mathcal{C}_0(r) = \phi(r/\ell_0)/(\pi\ell_0^2)$, where ϕ is a function that is independent of any microscopic parameters. From this we may write the local magnetisation density $\boldsymbol{m}^{\mathcal{C}}$ as a local integral of $\boldsymbol{m}$, i.e.

$$\boldsymbol{m}^{\mathcal{C}}(\boldsymbol{x}) = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}' \int_0^{2\pi} \mathrm{d}\theta\, \mathcal{C}_0(|\boldsymbol{x}'-\boldsymbol{x}|)p(\boldsymbol{x}',\theta)\boldsymbol{e}(\theta) \tag{3.56}$$

$$= \frac{1}{\pi\ell_0^2}\int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}'\, \phi(|\boldsymbol{x}'-\boldsymbol{x}|/\ell_0)\boldsymbol{m}(\boldsymbol{x}') \tag{3.57}$$

Now, from our assumptions above, $\nabla \sim \ell_p^{-1} \ll 1$ and so a gradient expansion of $\boldsymbol{m}$ is valid. Writing then

$$\boldsymbol{m}(\boldsymbol{x}') = \boldsymbol{m}(\boldsymbol{x}) + \left(\boldsymbol{x}'-\boldsymbol{x}\right)\cdot\nabla\boldsymbol{m} + \mathcal{O}(\nabla^2) \tag{3.58}$$

we may compute the integral in eq. (3.57) and we find that

$$\boldsymbol{m}^{\mathcal{C}}(\boldsymbol{x}) = \boldsymbol{m}(\boldsymbol{x}) + \mathcal{O}\left((\ell_0/\ell_p)^2\right). \tag{3.59}$$

Combining this result with the mean-field equation (3.25) we thus find that

$$\dot{\theta}_i = \beta\boldsymbol{e}_i'\cdot\hat{\boldsymbol{m}}(\boldsymbol{x}_i) + \sqrt{2}\xi_i + \mathcal{O}\left((\ell_0/\ell_p)^2\right). \tag{3.60}$$

From here on, we will ignore the higher order contributions in ℓ_0/ℓ_p in eq. (3.60), which corresponds to taking the limit $\ell_0 \to 0$.

In appendix A we show from eq. (3.60) that we may consistently expand the FPE in eq. (3.26) to obtain dynamical equations for ϱ and $\boldsymbol{m}$, essentially following [82]. For the

marginal density ϱ we obtain the continuity equation

$$\partial_t \varrho + \nabla \cdot \boldsymbol{m} = 0, \tag{3.61}$$

which is exact to all orders. In particular, it simply expresses the fact that ϱ is conserved and changes locally due to the flow induced by $\boldsymbol{m}$. Further, for the magnetisation density $\boldsymbol{m}$, we obtain to lowest non-trivial order, i.e. $\mathcal{O}(\beta, \nabla)$, that

$$\partial_t \boldsymbol{m} = -a\,(\varrho_c - \varrho)\,\boldsymbol{m} - \frac{1}{2}\nabla\varrho, \tag{3.62}$$

where we have defined

$$a = \frac{\beta}{2m}, \tag{3.63}$$

$$\varrho_c = \frac{2m}{\beta}. \tag{3.64}$$

The way eqs. (3.63) and (3.64) are written may seem slightly convoluted at first, since we have defined coefficients a and ϱ_c that depend on the magnetisation m. As we shall see soon, our motivation for this is to highlight both the similarities and differences between the field theories that we derive and what is generally referred to 'Toner-Tu hydrodynamic theories' in the literature. However, because of the way eqs. (3.63) and (3.64) are written, one should take some care when dealing with it. For example, one may verify that the only constant solutions $(\varrho_0, \boldsymbol{m}_0)$ of eqs. (3.61) and (3.62) have $\varrho_0 = 1/L^2$ and mean magnetisation $\bar{m} = m_0/\varrho_0$ given by

$$\bar{m} = \frac{\beta}{2}, \tag{3.65}$$

consistently with eq. (3.34), and not $m = 0$ as a superficial inspection of eq. (3.62) might lead one to first believe.

Unfortunately, the lowest order system of equations in eqs. (3.61) and (3.62) is unstable about the PL state in eq. (3.65). To see this, note that stability will require all roots of the cubic polynomial

$$g(\sigma) = \sigma^3 + \sigma^2 + \frac{1}{2}\left(q^2 + i\beta q_\parallel\right)\sigma + \frac{1}{2}q_\perp^2 \tag{3.66}$$

to have negative real part. In eq. (3.66), the wave vector $\boldsymbol{q} = (q_\parallel, q_\perp)$ has been decomposed into components that are parallel and perpendicular (respectively) to $\boldsymbol{m}_0$. To determine where the roots of g are located, we introduce a tool from complex analysis that we will apply again in section 4.6. Since we are working in $d = 2$, characteristic polynomials encountered in the linear stability analysis of hydrodynamic equations of the form of eqs. (3.61) and (3.62) are

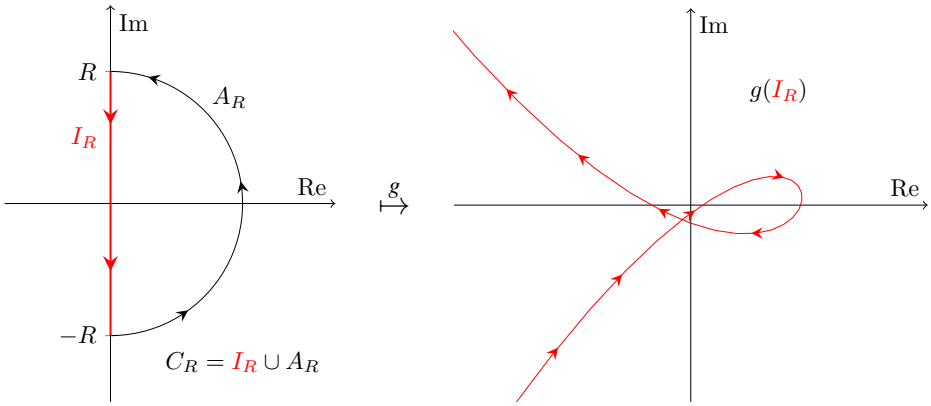


Fig. 3.6 Application of the argument principle to the polynomial g in eq. (3.66). For certain choices of β and wave vector $\boldsymbol{q}$, the loop formed by the image $g(I_R)$ misses the origin, implying that $Z_R \nrightarrow 0$. In the specific case illustrated, $\Delta_{I_R} \arg g = -\pi i$, and so g has one root with positive real part.

generally cubics with complex coefficients, rendering the argument principle quite useful (for an introduction to this method, see e.g. [108] chapter 5, theorem 5.1.4, or [109] chapter 5). It states that the number Z_R of zeros of g inside the semi-circle contour $C_R = I_R \cup A_R$, where

$$I_R = [-iR, iR], \tag{3.67}$$

$$A_R = \left\{ R\mathrm{e}^{i\theta} : \theta \in [-\pi/2, \pi/2] \right\}, \tag{3.68}$$

is given by the change in the argument of g as we trace C_R counterclockwise, thus

$$Z_R = \frac{1}{2\pi i} \oint_{C_R} \mathrm{d} \log g \tag{3.69}$$

$$= \frac{\Delta_{C_R} \arg g}{2\pi i}, \tag{3.70}$$

which is illustrated in fig. 3.6. By determining Z_R in the limit $R \to \infty$ we may therefore deduce the number of zeros of g in the half-plane $\{\mathrm{Re}\, \sigma > 0\}$, i.e. those with positive real part.

On the arc A_R, clearly $g\left(R\mathrm{e}^{i\theta}\right) = R^3 \mathrm{e}^{3i\theta} + \mathcal{O}(R)$ and so

$$\Delta_{A_R} \arg g = \log g\left(R\mathrm{e}^{\pi i/2}\right) - \log g\left(R\mathrm{e}^{-\pi i/2}\right) \tag{3.71}$$

$$= 3\pi i + \mathcal{O}\left(R^{-1}\right). \tag{3.72}$$

To compute Z_R in the limit $R \to \infty$, we are therefore left with having to find the change in the argument of g along I_R. Now, stability clearly requires that

$$\lim_{R\to\infty} \Delta_{I_R} \arg g = -3\pi i, \tag{3.73}$$

so that $Z_R \to 0$. Equation (3.73) is equivalent to the statement that the image $g(I_R)$ wraps around the origin in a *clockwise* fashion, exactly one and a half times. However, along I_R we may decompose g into its real and imaginary parts as

$$\operatorname{Re} g(iy) = -y^2 - \frac{\beta q_\parallel}{2} y + \frac{q_\perp^2}{2}, \tag{3.74}$$

$$\operatorname{Im} g(iy) = -y^3 + \frac{q^2}{2} y, \tag{3.75}$$

from which it is fairly straightforward to find a choice of β and $\boldsymbol{q}$ for which the image of I_R under g is a contour of the form shown in fig. 3.6, where the 'loop' misses the origin. In particular, since eq. (3.73) is not satisfied for such a curve, $\lim_{R\to\infty} Z_R \neq 0$.

Because of this, we proceed in appendix A to also compute the next order corrections to eq. (3.62). We find that to cubic order in β and ∇,

$$\partial_t \boldsymbol{m} + \lambda \left(\boldsymbol{m}\cdot\nabla\right)\boldsymbol{m} = -a\left(\varrho_c - \varrho\right)\boldsymbol{m} + \nu\nabla^2\boldsymbol{m} + \frac{\kappa_1}{2}\nabla\left(m^2\right) - \kappa_2 \boldsymbol{m}\left(\nabla\cdot\boldsymbol{m}\right) - \frac{1}{2}\nabla\varrho\,. \tag{3.76}$$

At this order we obtain additional terms, and the coefficient ϱ_c in eq. (3.76) is modified compared with its previous expression in eq. (3.64), while a is still given by eq. (3.63). Specifically, we find that

$$\lambda_{\alpha\beta} = \frac{\beta}{4m}\left(\frac{3}{4}\delta_{\alpha\beta} - \hat{m}_\alpha \hat{m}_\beta\right), \tag{3.77}$$

$$\varrho_c = \left(\frac{2}{\beta} + \frac{\beta}{4}\right) m, \tag{3.78}$$

$$\nu = \frac{1}{16}, \tag{3.79}$$

$$\kappa_1 = \frac{\beta}{4m}, \tag{3.80}$$

$$\kappa_2 = \frac{5\beta}{16m}. \tag{3.81}$$

Note that here, λ is matrix valued (and symmetric) and so component-wise the lambda-term in eq. (3.76) reads $\lambda_{\alpha\beta} m_\gamma \nabla_\gamma m_\beta$, where repeated Greek indices are implicitly summed over.

Again, as a check of consistency, the constant homogeneous solutions $(\varrho_0, \boldsymbol{m}_0)$ of the hydrodynamic equations (3.61) and (3.76) can be computed. From eq. (3.54) we find that $\varrho_0 = 1/L^2$ as before, while from eq. (3.76) we find that $\bar{m} = m_0/\varrho_0$ is given by

$$\bar{m} = \frac{\beta}{2} - \frac{\beta^3}{16} + \mathcal{O}\left(\beta^5\right) \tag{3.82}$$

consistently with eq. (3.34). Finally, we note that from our results in section 3.4, the hydrodynamic equations eqs. (3.61) and (3.76) should be stable about the PL state, and so we will not perform any further analysis here.

Note that, as before, we have written eq. (3.76) in a way in which the coefficients have a nontrivial dependence on m. Our motivation for this is to compare the structure of this equation with what we refer to as the **hydrodynamic Vicsek model** (HVM) class. In this thesis, this class comprises all hydrodynamic theories for a density ρ and magnetisation $\boldsymbol{m}$ that can be expressed in the form

$$\partial_t \rho + w \nabla \cdot \boldsymbol{m} = 0, \tag{3.83}$$

$$\partial_t \boldsymbol{m} + \lambda \left(\boldsymbol{m} \cdot \nabla\right) \boldsymbol{m} = -a\left(\rho_c - \rho\right) \boldsymbol{m} - b m^2 \boldsymbol{m} + \nu \nabla^2 \boldsymbol{m} + \frac{\kappa_1}{2} \nabla\left(m^2\right) - \kappa_2 \boldsymbol{m} \left(\nabla \cdot \boldsymbol{m}\right) - w_1 \nabla \rho + \boldsymbol{\eta}, \tag{3.84}$$

where $\boldsymbol{\eta}$ is a mean-zero Gaussian white noise with covariance

$$\left\langle \eta_\alpha(\boldsymbol{x}, t) \eta_\beta(\boldsymbol{x}', t') \right\rangle = 2D\delta(\boldsymbol{x} - \boldsymbol{x}')\delta(t - t'). \tag{3.85}$$

In eqs. (3.83) and (3.84), we allow all coefficients w, λ, a, ρ_c, b, ν, κ_1, κ_2, w_1 and D to be functions of ρ and $\boldsymbol{m}$, and moreover allow the average density $\rho_0 = L^{-2} \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \rho$ to take arbitrary values. The noise term $\boldsymbol{\eta}$ in eq. (3.84), which is added ad-hoc, is meant to capture finite-density fluctuations, which is also reflected by our use of ρ over ϱ.

It is not immediately clear, however, whether one can say much in general about the dynamics described by such a large class of theories. However, we find that presenting the equations in this form caters for some means of comparison of the structure of the equations we derive here and in subsequent chapters. Also, in section 4.4 we will study the subspace of models in which the coefficients w, λ, a, ρ_c, b, ν, κ_1, κ_2, w_1 and D are constants, which still leaves a crucial non-linearity in the linear term controlling the transition to the low-temperature phase, and in particular show that it supports both MPS and PL phases encountered for the CTVM. Interestingly, this theory also captures a *polar cluster* phase, which most of our discussion in chapter 4 will revolve around.

3.6 Conclusion

With this, we conclude our discussion of the phenomenology and phase diagram of the CTVM. In chapter 5, we will return to discuss this model in the context of entropy production, which will be the main focus for the latter half of this thesis. Here, we have introduced and studied a continuous-time formulation of the VM algorithm, i.e. eqs. (3.15) and (3.16), both numerically and analytically, and in particular shown that it can be solved exactly at infinite density.

From simulations we found that the phase diagram of the CTVM can be considered equivalent to that already known for the VM. In particular, the model displays an isotropic phase for large temperatures $T > T_g(\rho_0)$, a smectic arrangement of polarised bands (MPS) in the miscibility gap $T \in (T_g(\rho_0), T_l(\rho_0))$, and a PL phase that fills the entire domain for low temperatures $T < T_l(\rho_0)$, where T_g and T_l are the gas and liquid binodals respectively, and ρ_0 the average density. Although we did not perform explicit checks here, we believe there are good reasons to assume that other features, such as the giant number fluctuations and lever-rule, known from the VM will remain the same for the CTVM. Regardless, these aspects are not immediately relevant to our agenda in this thesis, and so we have silently ignored them here.

Our analytical work focused on the construction and assessment of the mean-field dynamics at $\rho_0 = \infty$, provided in eq. (3.25), and subsequently solving for the steady-state single-particle Boltzmann distribution p_{ss} in eq. (3.27). By combining the mean-field dynamics and the Boltzmann distribution, we also constructed the mean-field steady-state dynamics for a single particle in eq. (3.31) at infinite density, which allowed us to compute the order parameter $\bar{m}$ in eq. (3.34), correlation function C in eq. (3.35), and effective diffusion D_{eff} in eq. (3.37) as expansions in small inverse temperature β. Importantly, by computing $\bar{m}$ we demonstrated explicitly that the model has a second order critical point at $(\rho_c, T_c) = (\infty, \infty)$.

In sections 3.4 and 3.5 we investigated further the stability of, and dynamical relaxation to, the homogeneous PL steady-state in the CTVM. In particular, in section 3.4 we demonstrated a numerical procedure that can be employed in order to investigate the linear stability of the PL in the microscopic model, which will be employed again in chapter 4. Unsurprisingly, this showed that the PL distribution is a linearly stable solution to the FPE in eq. (3.26). Then, in section 3.5, we derived a mapping from the CTVM onto a Toner-Tu type hydrodynamic theory in the HVM class, valid at $\rho_0 = \infty$ in the vicinity of the critical point. There are still questions related to this model that we believe should be investigated further. In particular, it is still unclear whether including a noise term that accounts for finite-density fluctuations recovers the dynamics in the miscibility gap. Moreover, it would be interesting to ask

whether the deterministic dynamics possess banded solutions, which has been found for similar theories [82].

While many of the results presented here can be found elsewhere in the literature, our presentation illuminates some features of the model that we believe are of broader interest, as well as some that will be useful to us in the remaining chapters. Still, we are unaware of any previous work which has addressed explicitly the stability of the homogeneous PL distribution in the CTVM, or derived the hydrodynamic theory in section 3.5 (although the procedure we employed is standard in the literature [82]). In any case, we emphasize that the important aspect here is the methodology we have employed, rather than the specific results, and the fact that we will use the material presented as a benchmark against which to test our results in chapter 4. By comparing the mean-field theories of the CTVM and the *Active XY model*, we will be able to gain some insight into the kind of Vicsek-like interactions that are able to sustain the PL against fluctuations. In particular, in chapter 4, we will show that the PL breaks down and is replaced by a polar cluster phase when the 'effective interaction strength' is unbounded.

Chapter 4

The Active XY model

Closely related to the continuous-time Vicsek model (CTVM) discussed in chapter 3 is another model of collective motion, henceforth referred to as the **Active XY model** (AXYM), which has received some attention previously in the literature [82, 83, 110]. Farrell *et al.* [82] were the first to discuss the model, although in the context of pattern formation with a density-dependent self-propulsion speed $v_0(\rho)$. A major achievement of their contribution was to provide a formal approach for deriving 'Toner-Tu type' hydrodynamic theories from Langevin models such as the AXYM – a procedure we outlined in chapter 3 for the CTVM. The hydrodynamic equations obtained for the AXYM by Farrell *et al.* were subsequently studied extensively using both analytical and numerical tools at constant motility in $d = 1$ [100], and in particular shown to support solitonic banded solutions. Subsequently, a more rigorous approach was taken by Degond *et al.* [83], which in fact lead to the derivation of exact hydrodynamic scaling limits of the model at infinite density.

However, not until more recently were simulations of the microscopic model studied in closer detail at constant motility by Chepizhko *et al.* in [110]. Rather surprisingly, it was found that the phenomenology of ordered phase of the AXYM in fact differs quite significantly from the traditional Vicsek model (VM) dynamics [17, 97, 99]. Instead of displaying the hallmark microphase-separation (MPS) and polar liquid (PL) phases, particles in the low-temperature ordered phase of the AXYM were observed to aggregate into dense localised *polar clusters* (PCs), surrounded by an isotropic gas.

In this chapter, we will show explicitly using the tools introduced in chapter 3 that the homogeneous PL is always linearly unstable in the AXYM, and present novel analytical and numerical results obtained for the hydrodynamic theory in $d = 2$ [36]. By comparing the AXYM to the CTVM in chapter 3, it can be viewed as having an unbounded *local* inverse temperature β_i that depends on the local density and magnetisation. Because of this, we argue instead that the PC phase arises because particles in high-density regions are trapped

by a local freezing effect, in which the local interaction strength dominates orientational decorrelation from rotational diffusion. We also gain further physical insight into this by studying the full hydrodynamic model of Farrell *et al.* in $d = 2$ and comparing it to a vectorial model A (in the Halperin-Hohenberg classification [111]). In particular, we show that the instability leading to the formation of polar clusters can be related to a term in the free energy which favours increasing the magnetisation locally.

Our discussion begins in section 4.1 by introducing the AXYM and its phenomenology. Then, in section 4.2, we again employ a mean-field theory (MFT) approach at infinite density, in analogy with our treatment of the CTVM in chapter 3, and show that the mean-field Fokker-Planck equation (FPE) has both an isotropic and a homogeneous PL solution at low temperatures. In section 4.3, however, we show that this solution is always linearly unstable. We proceed to investigate this further in section 4.4, where we consider the subspace of theories in the hydrodynamic Vicsek model (HVM) class (introduced in section 3.5) in $d = 2$ where only the coefficient of the linear term that controls the transition to the low-temperature phase is non-constant, and show in particular that the AXYM can be mapped onto 'a single point' in this space. We also present a detailed linear stability analysis of this subspace of theories, which is deferred to section 4.6. Finally, in section 4.5, we show that we may recover a stable PL phase by making the local inverse temperature locally bounded, before presenting our conclusions and final remarks in section 4.7.

4.1 Locally extensive and non-extensive alignment interactions

Since the introduction of the original VM [17], several extensions and generalisations of the model has appeared in the literature [1, 81–89]. One of the first formulations to appear in continuous time is a model we will refer to as the **Active XY model** (AXYM), initially proposed by Farrell *et al.* [82]. Like the CTVM which we studied in chapter 3, it can be considered an extension of the VM, albeit with a local response time which depends on the local magnetisation. Furthermore, it is intriguing as it builds on even closer analogies with the XY-model known from equilibrium statistical mechanics. In this section, we will introduce the AXYM, and see in particular that its phenomenology at low temperatures in fact differs quite drastically from the hallmark microphase-separation (MPS) and polar liquid (PL) dynamics seen in the VM.

Analogously to the XY-model known from equilibrium statistical mechanics [112], in the AXYM we consider a collection $\{\boldsymbol{e}_i\}$ of N spins in a volume $\mathcal{V} = [0, L]^2$ whose interactions

derive from the potential

$$u_{\mathrm{XY}} = -\frac{1}{2\rho_0}\sum_{i,j=1}^{N} \mathcal{C}_{ij}\boldsymbol{e}_i\cdot\boldsymbol{e}_j. \tag{4.1}$$

In eq. (4.1), $\rho_0 = N/L^2$ is the average density of particles in the system, and $\mathcal{C}_{ij} \equiv \mathcal{C}_0(|\boldsymbol{x}_j - \boldsymbol{x}_i|)$ is an isotropic contact matrix (see chapter 3) that depends on the positions $\{\boldsymbol{x}_i\}$ of the particles. The function $\mathcal{C}_0(r)$ is taken to satisfy the normalisation criteria specified in section 3.2, and depends in particular on the interaction distance ℓ_0 via eq. (3.9). Note also that, as in chapter 3, the normalisation of $\mathcal{C}_0$ implies that it has dimensions of a density, and so ρ_0 appears in eq. (4.1) to make u_{XY} dimensionless. Moreover, we do not include an explicit coupling strength in eq. (4.1), since this may be absorbed by the temperature (to be introduced momentarily).

In contrast with the XY-model, however, the spins are not fixed on a lattice and are freely traveling at a constant motility v_0. Thus, in $d = 2$, the equations of motion we consider in the AXYM are given by

$$\dot{\boldsymbol{x}}_i = v_0\boldsymbol{e}_i, \tag{4.2}$$

$$\dot{\theta}_i = -\nu\partial_{\theta_i}u_{\mathrm{XY}} + \sqrt{2D_r}\xi_i, \tag{4.3}$$

where, as in section 3.1, the interaction strength ν and rotational diffusivity D_r are taken to be constant. For general d, the spin dynamics can be expressed by

$$\dot{e}_\alpha = \left(\delta_{\alpha\beta} - e_\alpha e_\beta\right)\circ\left(\nu\frac{m^{\mathcal{C}}_{i\beta}}{\rho_0} + \sqrt{2D_r}\xi_{i\beta}\right), \tag{4.4}$$

where $\boldsymbol{\xi}_i$ is, with slight abuse of notation, a d-dimensional unit white noise, and the local magnetisation density $\boldsymbol{m}^{\mathcal{C}}_i$ is defined by

$$\boldsymbol{m}^{\mathcal{C}}_i = \sum_{j=1}^{N}\mathcal{C}_{ij}\boldsymbol{e}_j, \tag{4.5}$$

in analogy with our treatment of the CTVM in chapter 3.

Despite the fact that the interaction term in eq. (4.9) derives from the potential u_{XY}, the steady-state distribution of the AXYM is *not* the equilibrium Boltzmann distribution. In fact, as we shall see explicitly in chapter 5, the AXYM is a true non-equilibrium model. At this point, one may nonetheless verify that the Fokker-Planck equation (FPE, see chapter 2) for

the distribution $p(\{\boldsymbol{x}_i, \theta_i\})$, i.e.

$$\partial_t p + v_0 \sum_{i=1}^{N} \boldsymbol{e}_i \cdot \nabla_i p = D_r \sum_{i=1}^{N} \partial_{\theta_i} \left(\mathrm{e}^{-\beta u_{\mathrm{XY}}} \partial_{\theta_i} \left(\mathrm{e}^{\beta u_{\mathrm{XY}}} p \right) \right), \tag{4.6}$$

is not solved by the equilibrium Boltzmann distribution

$$p_{\mathrm{ss}} \propto \mathrm{e}^{-\beta u_{\mathrm{XY}}}, \tag{4.7}$$

where we have defined the inverse temperature $\beta = v/D_r$. Indeed, it is easy to see that this is broken by the advective term in eq. (4.6), and so eq. (4.7) would only be a solution in the limit $v_0 \to 0$ where all spins have fixed positions in space, *or* in the limit $\ell_0 \to \infty$ where $\mathcal{C}_{ij} \to c$ for constant c and all spins interact with all others. This distinguishes the AXYM from the CTVM, which as we saw in section 3.3 is only an equilibrium model in the limit $\ell_0 \to \infty$ at fixed average density ρ_0.

For convenience, we may again (see chapters 2 and 3) measure time in units of the orientational correlation time $t_D \sim D_r^{-1}$ and effective diffusion length scale $\ell_D \sim v_0/D_r$, which we adopt in the following. This leads, in $d = 2$, to the nondimensionalised equations of motion

$$\dot{\boldsymbol{x}}_i = \boldsymbol{e}_i, \tag{4.8}$$

$$\dot{\theta}_i = \beta \boldsymbol{e}_i' \cdot \frac{\boldsymbol{m}_i^{\mathcal{C}}}{\rho_0} + \sqrt{2}\xi_i. \tag{4.9}$$

Crucially, the difference between eqs. (3.19) and (4.9) is that now the interaction term has a fluctuating strength, via $\boldsymbol{m}_i^{\mathcal{C}}$ which is no longer normalised to unit norm. Our task in the following will therefore be to investigate how the dynamics in eqs. (4.8) and (4.9) is different from that studied in chapter 3, and how it depends on the temperature $T = \beta^{-1}$ in addition to the average density ρ_0.

Simulations reveal that T controls the transition from a disordered isotropic phase at large $T > T_{\mathrm{cl}}(\rho_0)$ to an ordered phase at low $T < T_{\mathrm{cl}}(\rho_0)$, where $T_{\mathrm{cl}}(\rho_0)$ is a density-dependent threshold temperature. However, unlike the CTVM, we do not observe a miscibility gap with the typical MPS dynamics, nor do we reach a stable PL phase. Instead, we observe localised high-density regions of particles with strong polar order, that we refer to as *polar clusters* (PCs), immersed in a high-temperature isotropic gas. At fixed T, we find that by varying the average density ρ_0 we are always in the PC phase for temperatures T lower than a critical temperature $T_c = 1/2$, while for larger T the system is isotropic except at very low densities.

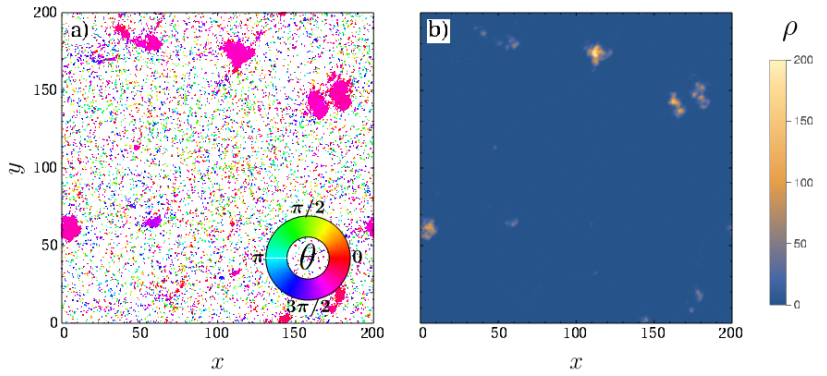


Fig. 4.1 Polar cluster phase of the AXYM: both panels a) and b) show the same snapshot from a realisation of the dynamics in eq. (4.9). In a) particles are colored based on the angle θ_i the spins make relative to the x-axis, while b) shows the corresponding density plot. The centers of the clusters are very dense relative to comparable MPS realisation of the CTVM, effectively trapping particles within them. This may be a condensation phenomenon. Simulation parameters used: $v_0 = 1$, $\nu = 0.4$, $D_r = 1$, $\rho_0 = 1$, $\ell_0 = 1$ and $L = 200$ at time-step $\Delta t = 6.25 \cdot 10^{-4}$.

In fig. 4.1, we show typical realisations of the dynamics in the PC region, while the resulting phase diagram is illustrated in fig. 4.2.

To explain this behaviour we note that, importantly, the interaction term in eq. (4.9) is *locally extensive*. By this we mean that if one introduces an inhomogeneity which doubles the local density $\rho_i^{\mathcal{C}}$ (at fixed ρ_0) around particle i, defined here by

$$\rho_i^{\mathcal{C}} = \sum_{j=1}^{N} \mathcal{C}_{ij}, \tag{4.10}$$

then the strength of the local alignment is effectively doubled. For some physical systems, one would expect this to be a property of the interaction. For example, in a suspension of active particles with magnetic dipole moments, the local interaction should be derivable from a potential and will therefore be extensive. On the other hand, in biological flocks there is no a priori reason that the alignment should be sensitive to the local signal strength. Consequently, the CTVM might be better suited to capture collective motion in living systems, since each particle is oblivious to whether there are few or many particles within its neighbourhood to align with.

In light of the local extensivity, one way to view the AXYM is as a CTVM dynamics with a fluctuating *local* inverse temperature $\beta_i = \beta m_i^{\mathcal{C}}/\rho_0$ by writing eq. (4.9) as

$$\dot{\theta}_i = \beta_i \boldsymbol{e}_i' \cdot \hat{\boldsymbol{m}}_i^{\mathcal{C}} + \sqrt{2}\xi_i, \tag{4.11}$$

where $\hat{\boldsymbol{m}}_i^{\mathcal{C}} = \boldsymbol{m}_i^{\mathcal{C}}/m_i^{\mathcal{C}}$ is the unit vector pointing in the direction of the local magnetisation density $\boldsymbol{m}_i^{\mathcal{C}}$. Indeed, comparing eq. (4.11) with the spin dynamics of the CTVM in eq. (3.19), we see that we could have obtained the AXYM also by arguing that the temperature should have been dependent on $m_i^{\mathcal{C}}$.

From eq. (4.11), we see that when the local density $\rho_i^{\mathcal{C}}$ is large, the local temperature $T_i = \beta_i^{-1}$ becomes small, i.e. $T_i \sim 1/\rho_i^{\mathcal{C}} \ll 1$. In particular, in the ordered phase, density fluctuations that lead to spatial inhomogeneities in the local temperature in turn increases the local interaction strength in regions where the local density $\rho_i^{\mathcal{C}}$ is large, and decreases it in regions where $\rho_i^{\mathcal{C}}$ is small. Consequently, in regions where the polar order is strong, the density will tend to increase since the interaction strength dominates the diffusion, and thus locally *freezing* the spin dynamics. Unlike typical coarsening dynamics in the miscibility gap that lead to band formation, as in the CTVM, the inverse temperature does not saturate in dense regions, so we hypothesise that these clusters may instead be an example of *condensation*.

Of course, we would not expect this to occur if T_i was bounded from below by some $T_0 > 0$. In such a scenario, we could increase ρ_0 until the temperature was uniform across the simulation domain, effectively being back in the CTVM set-up. A natural question one might ask therefore, is whether adding volume exclusion to the model would counteract the formation of clusters, since this would place an upper bound on the local density $\rho_i^{\mathcal{C}}$. Here we have opted not to pursue this question, and leave an investigation into this for future work. We will, however, in section 4.5, consider a model closely related to the AXYM in which the interaction is modified to be locally non-extensive, which also places a bound on the local temperature T_i, and for which we *do* observe stable MPS and PL phases.

Because of the cluster formation in the AXYM in the low-temperature phase, large-scale simulations are rather expensive to perform. Even for finite-ranged interaction kernels, the number of nearest neighbours of each particle in a cluster becomes very large, meaning that for each time-step one has to loop over a significant number of particles to calculate the local aligning force. In addition, large collections of these loops involve the same particles, and so cannot be treated concurrently. Furthermore, because the effective local interaction strength then also becomes very large, the simulation becomes highly sensitive to the time-step.

We continue our discussion of the AXYM in section 4.2 by constructing a mean-field theory (MFT) using an analogous procedure to that which was adopted in section 3.3 for

the CTVM. There we will see that, despite the above, a homogeneous PL solution can be constructed to the mean-field FPE of the AXYM. However, in section 4.3, we show that this is indeed always linearly unstable.

4.2 Mean field theory of the AXYM

In analogy with our treatment of the CTVM in chapter 3, we will again in this section appeal to the fact that a mean-field dynamics and 'non-linear FPE' is an exact representation of the AXYM dynamics at $\rho_0 = \infty$, as shown by Bolley *et al.* [79]. We find two steady-state solutions to the FPE, namely the isotropic and ordered PL solutions, which were first obtained by Degond *et al.* [83]. In contrast to the CTVM, however, the ordered solution of the AXYM only exists in a finite temperature range, yielding a finite critical temperature $T_c \equiv T_{\mathrm{cl}}(\infty)$. Naïvely, one might therefore expect that the phase diagram of the AXYM should be equivalent to that of the CTVM, albeit with a shifted critical point. *However*, as we will see in section 4.3, the PL solution is in fact never linearly stable for any temperature $T < T_c$. Because of this, we conclude that the ordered region of the phase diagram of the model is not a PL region.

To derive the MFT of the AXYM, we will follow the procedure outlined in section 3.3 for the CTVM. First, we begin by introducing the molecular field $\boldsymbol{m}^{\mathcal{C}}(\boldsymbol{x})$ as the angular moment

$$\boldsymbol{m}^{\mathcal{C}}(\boldsymbol{x}) = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}' \int_0^{2\pi} \mathrm{d}\theta \, C_0(|\boldsymbol{x}-\boldsymbol{x}'|)\boldsymbol{e}(\theta)p(\boldsymbol{x}',\theta) \tag{4.12}$$

of the one-particle distribution function $p(\boldsymbol{x},\theta) = \lim_{N\to\infty} N^{-1}\sum_{i=1}^{N}\delta(\boldsymbol{x}-\boldsymbol{x}_i)\delta(\theta-\theta_i)$. Thus, by taking the limit $N\to\infty$ of the angular dynamics in eq. (4.9), it follows that at infinite density

$$\dot{\theta}_i = L^2\beta\boldsymbol{e}_i'\cdot\boldsymbol{m}^{\mathcal{C}}(\boldsymbol{x}_i) + \sqrt{2}\xi_i, \tag{4.13}$$

implying that the dynamics become decoupled. This MFT is, as expected from inspection of eq. (4.9), different from what we had before in eq. (3.25) since now the molecular field has a fluctuating norm. Also, since the molecular field $\boldsymbol{m}^{\mathcal{C}}$ depends on the distribution p, the FPE for the mean-field AXYM given by

$$\partial_t p + \boldsymbol{e}\cdot\nabla p = \partial_\theta\left(\mathrm{e}^{L^2\beta\boldsymbol{e}\cdot\boldsymbol{m}^{\mathcal{C}}}\partial_\theta\left(\mathrm{e}^{-L^2\beta\boldsymbol{e}\cdot\boldsymbol{m}^{\mathcal{C}}}p\right)\right) \tag{4.14}$$

is again non-linear in analogy with the mean-field CTVM in section 3.3, however the precise form of the diffusive contribution on the right-hand side differs from eq. (3.26) since the molecular field has a non-unit norm.

The steady-state solutions of eq. (4.14) also differ from those obtained for the CTVM. For the AXYM, one looks for a solution of eq. (4.14) of the form

$$p_{\mathrm{ss}}(\boldsymbol{x},\theta) = \frac{\mathrm{e}^{\beta \boldsymbol{e}\cdot\boldsymbol{M}}}{2\pi L^2 I_0(\beta M)}. \tag{4.15}$$

Specifically, eq. (4.15) differs from the corresponding distribution in eq. (3.27) for the CTVM in that we a priori allow the norm $M = |\boldsymbol{M}|$ of the constant vector $\boldsymbol{M}$ to take any finite value, rather than restrict to those with unit norm. It follows that for any isotropic kernel $\mathcal{C}_0$, the local magnetisation density is given by

$$\boldsymbol{m}^{\mathcal{C}}_{\mathrm{ss}}(\boldsymbol{x}_i) = \frac{1}{L^2}\frac{I_1(\beta M)}{I_0(\beta M)}\hat{\boldsymbol{M}}. \tag{4.16}$$

Thus, substituting eq. (4.16) back into eq. (4.14), we see that p_{ss} is a solution if and only if the global magnetisation $\boldsymbol{M}$ satisfies the self-consistency condition

$$M = \frac{I_1(\beta M)}{I_0(\beta M)}. \tag{4.17}$$

The self-consistency equation in eq. (4.17) is well known from the mean-field equilibrium XY-model [113]. In particular, it has a disordered solution $M = 0$, and an additional ordered solution with $M > 0$ for low temperatures $T < T_c$, where the critical temperature $T_c = 1/2$. In the low temperature regime $T < T_c$, the ordered solution is monotonic and satisfies $M \nearrow 1$ as $T \searrow 0$ and $M \searrow 0$ as $T \nearrow T_c$. Near T_c, it can be expanded in small $0 < T_c - T \ll 1$ as

$$M = 2(T_c - T)^{1/2} + o\left((T_c - T)^{1/2}\right). \tag{4.18}$$

Thus, computing the mean global magnetisation under p_{ss} is trivial, and it follows that near the critical point,

$$\bar{m} = M \sim (T_c - T)^{1/2} \tag{4.19}$$

with a critical exponent that is different from the one obtained for the CTVM (cf. eq. (3.34)).

Despite the existence of the ordered solution, we will see next in section 4.3 that it is not a linearly stable solution of the FPE. Because of this, we cannot infer what the ordered region illustrated in the phase diagram in fig. 4.2 should be from this approach alone. However, combined with our observations from simulations of the dynamics, we conclude that there cannot be a PL region. We have, however, been able to observe 'band-like' realisations, although these are only present for very large densities and small system sizes. Once the system size L is increased, the system quickly returns to the PC phase. As we will show

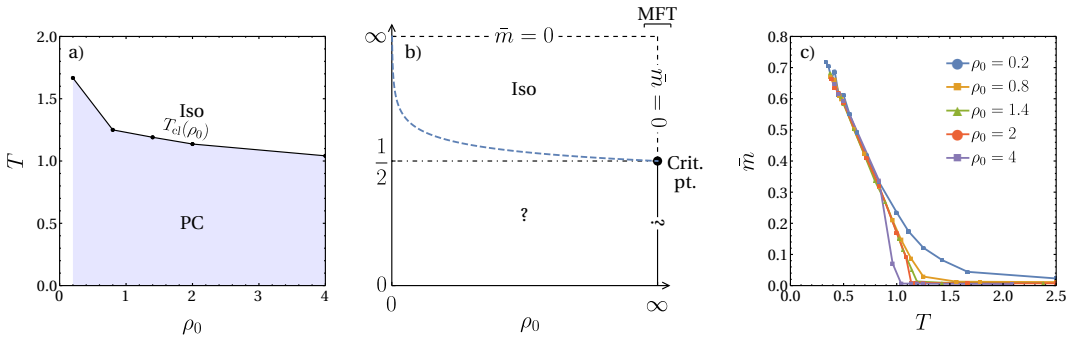


Fig. 4.2 Phase diagram of the AXYM. In panel a), the phase diagram obtained from simulations is shown, while b) is an illustration of the mean-field results. In particular, since the PL steady-state is unstable in the MFT of the AXYM, we cannot yet determine the stable ordered steady-state from MFT alone. In b), the blue dashed line indicates what we believe the critical line $T_c(\rho_0)$ should look like, although it has no analytical basis and an analytical estimate of $T_{\mathrm{cl}}(\rho_0)$ for $\rho_0 \gg 1$ remains elusive. Finally, the data points in a) are estimated from the plot of $\bar{m}$ against T shown in panel c).

next in section 4.3, this can be explained by the fact that the PL solution is destabilised by finite-wavelength modes that cannot be excited for sufficiently small system sizes. We also note that the shape of the critical line $T_{\mathrm{cl}}(\rho_0)$ separating the isotropic and ordered phases, as inferred from simulations, is different for the AXYM compared to the CTVM. On the other hand, this may also simply be a result of the fact that the limit $\rho_0 \to 0$ is not the same for the AXYM and CTVM, which for the former is singular.

4.3 Breakdown of the polar liquid

Investigating the linear stability of the PL solution p_{ss} in eq. (4.15) to the FPE in eq. (4.14) can be done straightforwardly using the tools developed for the CTVM in section 3.4. Thus, the calculation we present here will to a large extent mirror the one presented there. Importantly, however, we will show that in contrast to the CTVM, the homogeneous PL is never linearly stable in the AXYM. In addition, we will see that the instability is caused by a finite-wavelength instability, so that when the system size L is sufficiently small, the unstable modes cannot be excited.

First we pick $\boldsymbol{M} = M\hat{\boldsymbol{x}}$ to be along the x-direction, which we may do without loss of generality. Thus, linearising the FPE in eq. (4.14) about p_{ss}, by writing $p = p_{\mathrm{ss}} + \delta p$ and

keeping terms only to linear order, we obtain

$$\partial_t \delta p + \boldsymbol{e}\cdot\nabla\delta p = \partial_\theta\left(\varphi\sin\theta\,\delta p - L^2\beta p_{ss}\boldsymbol{e}'\cdot\delta\boldsymbol{m}^{\mathcal{C}} + \partial_\theta\delta p\right). \tag{4.20}$$

In eq. (4.20) we have defined $\varphi = \beta M(\beta)$, in addition to the variation $\delta\boldsymbol{m}^{\mathcal{C}}$ in the field $\boldsymbol{m}^{\mathcal{C}}$ as that produced by the perturbation δp, i.e.

$$\delta\boldsymbol{m}^{\mathcal{C}}(\boldsymbol{x}) = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}' \int_0^{2\pi} \mathrm{d}\theta\, C_0(|\boldsymbol{x}-\boldsymbol{x}'|)\delta p(\boldsymbol{x}',\theta)\boldsymbol{e}(\theta). \tag{4.21}$$

Proceeding as in section 3.4, we transform eq. (4.20) to Fourier space in the angular variable θ. From this, we obtain for the Fourier coefficients $\delta p_n = \int_0^{2\pi}\mathrm{d}\theta\,\delta p\,\mathrm{e}^{-in\theta}$ the equation

$$\partial_t\delta p_n + \frac{1}{2}\nabla\cdot\begin{pmatrix}\delta p_{n-1}+\delta p_{n+1}\\ -i(\delta p_{n-1}-\delta p_{n+1})\end{pmatrix} = -n^2\delta p_n + \frac{\varphi n}{2}(\delta p_{n-1}-\delta p_{n+1}) - \boldsymbol{c}_n\cdot\delta\boldsymbol{m}^{\mathcal{C}}, \tag{4.22}$$

where the constants $\boldsymbol{c}_n$ are defined by

$$\boldsymbol{c}_n = \frac{in\beta}{2\pi I_0(\varphi)}\int_0^{2\pi}\mathrm{d}\theta\,\boldsymbol{e}'\,\mathrm{e}^{\varphi\cos\theta - in\theta}. \tag{4.23}$$

Moreover, the integral can be performed exactly by identifying it as a modified Bessel integral [106], and we find that

$$\boldsymbol{c}_n = \begin{pmatrix}-n^2 I_n(\varphi)/I_1(\varphi)\\ in\beta I_n'(\varphi)/I_0(\varphi)\end{pmatrix}. \tag{4.24}$$

Already at this stage, we notice a slight difference between eq. (3.40) for the Fourier modes of δp in the CTVM and eq. (4.22). Whereas eq. (3.40) only coupled $\delta m_\perp^{\mathcal{C}}$ with δp_n, eq. (4.22) now couples both the parallel and perpendicular fluctuations of $\delta\boldsymbol{m}^{\mathcal{C}} = \left(\delta m_\parallel^{\mathcal{C}}, \delta m_\perp^{\mathcal{C}}\right)^T$.

As in section 3.4, we treat the real and imaginary parts of the Fourier coefficients δp_n separately and denote these by δp_n^R and δp_n^I respectively. After performing a Fourier transform in the position $\boldsymbol{x}$, subsequent to separating δp_n^R and δp_n^I, we obtain the coupled set

of equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta p^R_{n,\boldsymbol{q}} + \frac{i\boldsymbol{q}}{2}\cdot\begin{pmatrix}\delta p^R_{n-1,\boldsymbol{q}} + \delta p^R_{n+1,\boldsymbol{q}}\\ \delta p^I_{n-1,\boldsymbol{q}} - \delta p^I_{n+1,\boldsymbol{q}}\end{pmatrix} = -n^2\delta p^R_{n,\boldsymbol{q}} + \frac{\varphi n}{2}(\delta p^R_{n-1,\boldsymbol{q}} - \delta p^R_{n+1,\boldsymbol{q}}) - \boldsymbol{c}^R_n\cdot\delta\boldsymbol{m}^{\mathcal{C}}_{\boldsymbol{q}}, \tag{4.25}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta p^I_{n,\boldsymbol{q}} + \frac{i\boldsymbol{q}}{2}\cdot\begin{pmatrix}\delta p^I_{n-1,\boldsymbol{q}} + \delta p^I_{n+1,\boldsymbol{q}}\\ \delta p^R_{n+1,\boldsymbol{q}} - \delta p^R_{n-1,\boldsymbol{q}}\end{pmatrix} = -n^2\delta p^I_{n,\boldsymbol{q}} + \frac{\varphi n}{2}(\delta p^I_{n-1,\boldsymbol{q}} - \delta p^I_{n+1,\boldsymbol{q}}) - \boldsymbol{c}^I_n\cdot\delta\boldsymbol{m}^{\mathcal{C}}_{\boldsymbol{q}}, \tag{4.26}$$

for the Fourier coefficients $\delta p^R_{n,\boldsymbol{q}}$ and $\delta p^I_{n,\boldsymbol{q}}$. In eqs. (4.25) and (4.26), we also denote by $\boldsymbol{c}^R_n$ and $\boldsymbol{c}^I_n$ the real and imaginary components of $\boldsymbol{c}_n$ respectively.

Using the fact that $\delta\boldsymbol{m}^{\mathcal{C}}_{\boldsymbol{q}} = \mathcal{C}_{0,\boldsymbol{q}}\left(\delta p^R_{1,\boldsymbol{q}}, -\delta p^I_{1,\boldsymbol{q}}\right)^T$ (cf. eq. (3.46) in section 3.4), where $\mathcal{C}_{0,\boldsymbol{q}} = \int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\,\mathcal{C}_0(|\boldsymbol{x}|)\mathrm{e}^{-i\boldsymbol{q}\cdot\boldsymbol{x}}$, we may thus again express eqs. (4.25) and (4.26) as a coupled linear system involving only the functions $\delta p^R_{n,\boldsymbol{q}}, \delta p^I_{n,\boldsymbol{q}}$ for $n \geq 0$, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta\boldsymbol{p}_{\boldsymbol{q}} = \mathcal{L}_{\boldsymbol{q}}\delta\boldsymbol{p}_{\boldsymbol{q}}, \tag{4.27}$$

where $\delta\boldsymbol{p}_{\boldsymbol{q}} = (\delta p^R_{0,\boldsymbol{q}}, \delta p^R_{1,\boldsymbol{q}}, \delta p^I_{1,\boldsymbol{q}}, \ldots)^T$. Our aim will be to show that the maximal eigenvalue $\lambda_m \equiv \lambda_m(T,\boldsymbol{q})$ of $\mathcal{L}_{\boldsymbol{q}}$ attains positive values for all $T < T_c$ and wave vectors $\boldsymbol{q}$ within a finite range, meaning that long-wavelength fluctuations increase exponentially in time and destabilise the PL phase. This can be done numerically by following the same procedure as for the CTVM. That is, we study the corresponding finite-dimensional problem of eq. (4.27) by introducing an ultraviolet cutoff Λ for the angular modes and defining $\delta\boldsymbol{p}^{\Lambda}_{\boldsymbol{q}} = (\delta p^R_{0,\boldsymbol{q}}, \delta p^R_{1,\boldsymbol{q}}, \delta p^I_{1,\boldsymbol{q}}, \ldots p^R_{\Lambda,\boldsymbol{q}}, \delta p^I_{\Lambda,\boldsymbol{q}})^T$ as the truncated vector of angular Fourier components with $|n| \leq \Lambda$.

As shown in fig. 4.3, the spectrum of $\mathcal{L}_{\boldsymbol{q}}$ with $\mathcal{C}_{\boldsymbol{q}}$ given by eq. (3.47) does indeed contain positive values for all temperatures $T < T_c$ in the ranges of $\boldsymbol{q}$ investigated, as well as for both choices of $\ell_0 = 0$ and $\ell_0 = 1$. Moreover, for all temperatures investigated, we find that for wave vectors $\boldsymbol{q}$ with norm $q > q_0(T)$, the spectrum of $\mathcal{L}_{\boldsymbol{q}}$ is strictly negative. This allows us to define a fastest growing mode $\lambda^*_m(T) = \sup_{\boldsymbol{q}}\lambda_m(T,\boldsymbol{q})$, and we find that the supremum is realised for a finite temperature-dependent wave vector $\boldsymbol{q}^*(T) = \left(q^*_{\parallel}(T), 0\right)$ which is parallel to $\boldsymbol{M}$. In fig. 4.5, we also plot $q^*_{\parallel}(T)$ as a function of the temperature T. This defines a length scale $\ell^*(T) \sim 1/q^*_{\parallel}(T)$, which could be compared in future work with the typical linear size of clusters, although we do not pursue this further here. In addition, one may also define a second temperature-dependent length scale $L_0(T) \sim 1/q_0(T)$, such that for system sizes

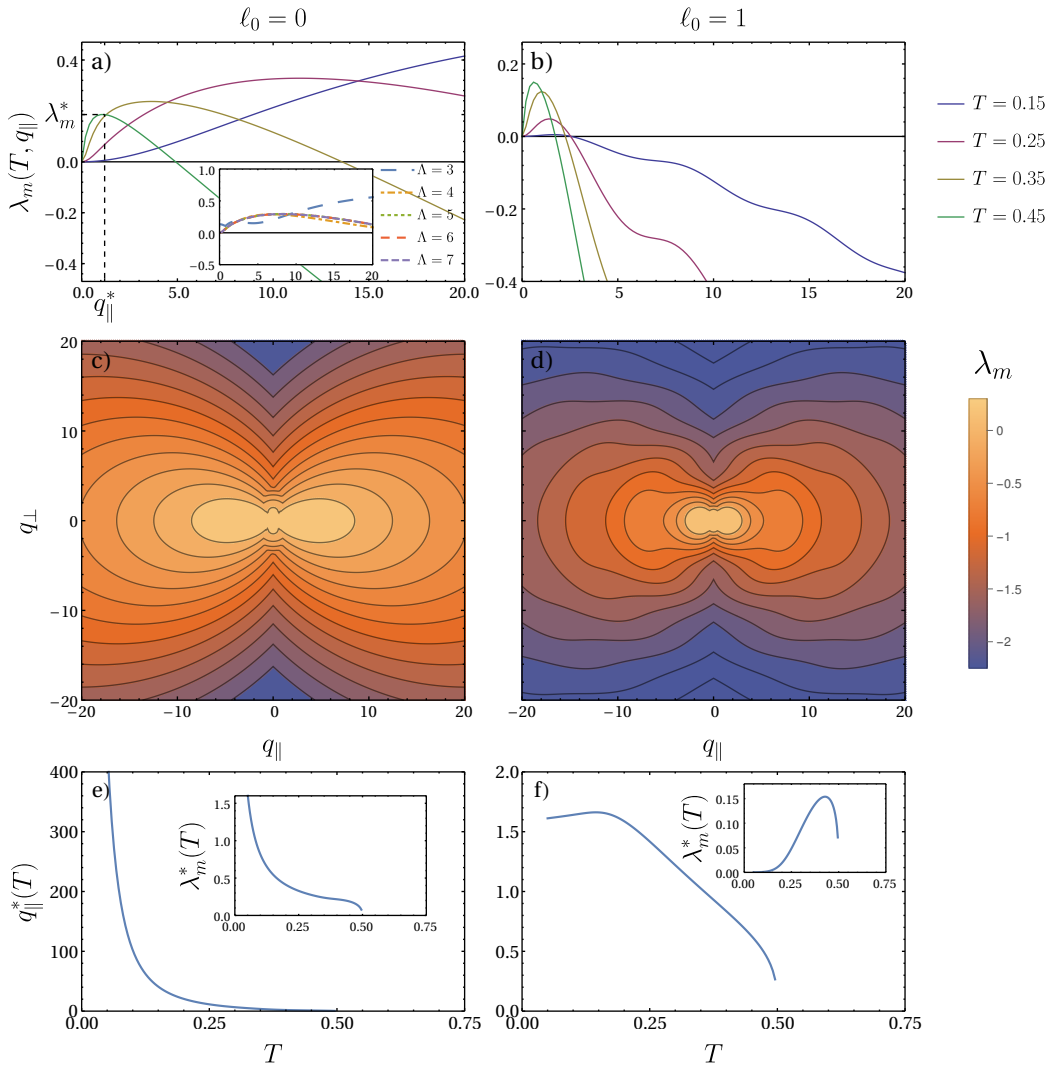


Fig. 4.3 The maximal eigenvalue λ_m of the matrix $\mathcal{L}_{\boldsymbol{q}}$ as a function of: a) and b) wave number $q_\parallel$ for different temperatures T at $q_\perp = 0$, and, c) and d) $\boldsymbol{q} = (q_\parallel, q_\perp)$ at fixed temperature $T = 0.4$. In panels e) and f), the wave number $q_\parallel$ which realises the supremum $\lambda_m^*(T) = \sup_{\boldsymbol{q}} \lambda_m(T, \boldsymbol{q})$ is shown as a function of temperature, with the insets showing $\lambda_m^*(T)$ as a function of T. For the column a), c) and e) the stability problem was solved in the limit of purely local interactions $\ell_0 \to 0$, while for b), d) and f) we chose a finite interaction range $\ell_0 = 1$. The plots demonstrate that λ_m attains positive values for all temperatures $T \leq 1/2$, so that in consequence the polar liquid solution p_{ss} is always unstable. In a), the inset shows that convergence to the curves $\lambda_m(T, q_\parallel)$ with the ultraviolet cutoff Λ is quite rapid. For all other plots, the cutoff was set to $\Lambda = 20$.

$L < L_0(T)$, the unstable modes cannot be excited. This explains why one might observe a seemingly stable PL phase (or bands), when the system size is sufficiently small.

Finally, we also make a brief note that in the case where the wave vector is parallel to $\boldsymbol{M}$, i.e. $\boldsymbol{q} = q_\parallel \hat{\boldsymbol{x}}$, the equations for $\delta p^R_{n,\boldsymbol{q}}$ and $\delta p^I_{n,\boldsymbol{q}}$ decouple and we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta p^R_{n,\boldsymbol{q}} = -n^2\left(\delta p^R_{n,\boldsymbol{q}} - \mathcal{C}_{0,\boldsymbol{q}}\frac{I_n(\varphi)}{I_1(\varphi)}\delta p^R_{1,\boldsymbol{q}}\right) + z_{n,q_\parallel}\delta p^R_{n-1,\boldsymbol{q}} - z^*_{n,q_\parallel}\delta p^R_{n+1,\boldsymbol{q}}, \tag{4.28}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta p^I_{n,\boldsymbol{q}} = -n^2\delta p^I_{n,\boldsymbol{q}} + n\beta\mathcal{C}_{0,\boldsymbol{q}}\frac{I'_n(\varphi)}{I_0(\varphi)}\delta p^I_{1,\boldsymbol{q}} + z_{n,q_\parallel}\delta p^I_{n-1,\boldsymbol{q}} - z^*_{n,q_\parallel}\delta p^I_{n+1,\boldsymbol{q}}, \tag{4.29}$$

where $z_{n,q_\parallel} = (\varphi n - iq_\parallel)/2$. One may check numerically, using a truncation procedure analogous to that used above, that the eigenvalues associated with eq. (4.29) are always dominated by those of the linear operator defined by eq. (4.28). To compute the maximal eigenvalue λ_m of $\mathcal{L}_{\boldsymbol{q}}$ for $q_\perp = 0$ it therefore suffices to consider only eq. (4.28). This may also provide a starting point for calculating analytical estimates of $q^*_\parallel(T)$, which we aim to pursue in future work.

In the next section, we will investigate a subspace of models in the hydrodynamic Vicsek model (HVM) class introduced in chapter 3 where all couplings are constant, with the exception of the quadratic term of the local free energy, which has been the subject of some discussion in the literature [1, 36, 82]. In particular, we will see that in $d = 2$ all three ordered phases seen thus far in this thesis, i.e. the MPS, PL and PC phases [36], can be realised by varying the parameters of this theory. We then relate the AXYM to this theory by following the approach outlined in section 3.5, which maps the AXYM onto a single point in the PC region of the phase diagram.

4.4 The hydrodynamic Vicsek model

Going beyond linear order in the FPE in eq. (4.14) analytically is rather difficult. However, in the vicinity of the critical point at $(\rho_0, T) = (\infty, 1/2)$, we can follow the standard procedure outlined in section 3.5 to derive a mapping onto a model in the HVM class, for which standard field-theoretic tools are more readily applicable. Unlike our discussion in section 3.5, we will focus less on the derivation of this mapping here since we describe the general method in appendix A, and moreover the explicit derivation for the AXYM may be found elsewhere in the literature [82]. Instead, we will show that the AXYM is mapped onto a model where all coefficients are constants, with the important exception of the linear term that controls the transition to the low-temperature phase which depends on the local density ρ. By analysing this restricted subspace of the HVM class, we show that 'the AXYM point' lies in the PC

region of a phase diagram that also contains both the MPS and PL phases. Also, importantly, since most of the remaining discussion in this thesis will focus on this restricted subspace of the HVM class, we will refer to this theory simply as 'the HVM'.

Our discussion in this section will be subdivided into two parts. First, we recall the HVM class from section 3.5, and proceed to investigate its dynamics when the couplings we defined there (which we repeat below) are constants. Importantly, we retain the property that the prefactor of the linear term (in the magnetisation $\boldsymbol{m}$) depends on the density, which is crucial to observe the formation of bands [89, 100]. To investigate this dynamics, we combine simulations with a linear stability analysis, from which we are able to infer a phase diagram for the model. For the second part, we introduce the aforementioned mapping from the AXYM to the HVM theory space.

In the HVM class we consider a density ρ which is advected by the flow induced locally by the magnetisation density $\boldsymbol{m}$. Specifically, recalling our definition in section 3.5, we consider the dynamical equations

$$\partial_t \rho + w\nabla\cdot\boldsymbol{m} = 0, \tag{4.30}$$

$$\partial_t \boldsymbol{m} + \lambda\left(\boldsymbol{m}\cdot\nabla\right)\boldsymbol{m} = -a\left(\rho_c-\rho\right)\boldsymbol{m} - bm^2\boldsymbol{m} + \nu\nabla^2\boldsymbol{m} + \frac{\kappa_1}{2}\nabla\left(m^2\right) - \kappa_2\boldsymbol{m}\left(\nabla\cdot\boldsymbol{m}\right) - w_1\nabla\rho + \boldsymbol{\eta}, \tag{4.31}$$

where $\boldsymbol{\eta}$ is a mean-zero Gaussian white noise with covariance

$$\left\langle \eta_\alpha(\boldsymbol{x},t)\eta_\beta(\boldsymbol{x}',t')\right\rangle = 2D\delta_{\alpha\beta}\delta(\boldsymbol{x}-\boldsymbol{x}')\delta(t-t'). \tag{4.32}$$

Here, we will restrict to the case where all couplings w, λ, a, ρ_c, b, ν, κ_1, κ_2, w_1 and D are constants, and where the average density $\rho_0 = L^{-2}\int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\,\rho$ may take arbitrary values, and henceforth we refer to the resulting model for simplicity as 'the HVM'. Notice, however, that as stated above the coefficient $a(\rho_c-\rho)$ of the term linear in $\boldsymbol{m}$ in eq. (4.31) depends explicitly on the density ρ with this choice. The noise term $\boldsymbol{\eta}$ in eq. (3.84), which is added ad-hoc, is meant to capture finite-density fluctuations.

In the following, we will assume that we may set $\kappa \equiv \kappa_1 = \kappa_2$. Clearly, this assumption does not hold for the hydrodynamic theory of the CTVM from chapter 3 in eqs. (3.61) and (3.76). Notwithstanding, as we shall see at the end of this section and in section 4.5, it *does* hold for the hydrodynamic equations we derive for the AXYM and NeAXYM. Because of this, we restrict our discussion here to the subspace with $\kappa_1 = \kappa_2$ of the full theory space of the HVM, which as we shall see still provides useful insight.

Now observe that when all coefficients of the HVM are constants, we may perform a suitable scale transformation of space, time and the fields to get rid of a, b, ν and ρ_c, which

we adopt in the following. The HVM dynamics can then be written as

$$\partial_t \rho + w\nabla\cdot\boldsymbol{m} = 0, \tag{4.33}$$

$$\partial_t \boldsymbol{m} + \lambda\left(\boldsymbol{m}\cdot\nabla\right)\boldsymbol{m} = \left(\rho - 1 - m^2\right)\boldsymbol{m} + \nabla^2\boldsymbol{m} + \kappa\left(\frac{1}{2}\nabla\left(m^2\right) - \boldsymbol{m}\left(\nabla\cdot\boldsymbol{m}\right)\right) - w_1\nabla\rho + \boldsymbol{\eta}, \tag{4.34}$$

where the remaining free parameters are ρ_0, w, w_1, λ, κ, D and the system size L. To investigate the HVM, we have performed numerical simulations and a linear stability analysis of the constant homogeneous solutions, the results from which we present below. In particular, from simulations we are able to observe all four phases of the microscopic models, including the isotropic, MPS, PL and PC phases. By combining our numerical findings with the analytical results from the linear stability analysis we are then able to present a phase diagram for the HVM.

Before we proceed, however, it should be noted that several authors have addressed the effects of different noise statistics on similar field theories, including scalar versus vectorial noise, as well as additive versus multiplicative [1, 99]. In general these have found that the dynamics remain reasonably stable against such modifications, and so for our purposes the additive noise $\boldsymbol{\eta}$ in eq. (4.34) will suffice. In addition, we note that all simulations we present here were performed using a Fourier-Galerkin pseudospectral scheme with semi-implicit time stepping [114], and allowed to relax to the steady-state for several choices of D to ensure stability.

The constant homogeneous solutions $(\rho_0, \boldsymbol{m}_0)$ to eq. (4.34) are the isotropic and polarly ordered states: when $\rho_0 < 1$ the only solution is the isotropic one for which $m_0 = 0$, while for $\rho_0 > 1$ the equation is also solved by the PL state with $m_0 = \sqrt{\rho_0 - 1}$. The isotropic state can be shown to be linearly stable if and only if $\rho_0 < 1$ fairly straightforwardly. The PL, on the other hand, requires more work, and we find that considering the two special cases for which the wave vector $\boldsymbol{q}$ is either strictly parallel or perpendicular to $\boldsymbol{m}_0$ provides us with some insight. Because this computation is rather involved, we defer it to section 4.6 to make the presentation here more accessible, and will therefore only present our main findings now. It should be noted also that we have checked our predictions by solving the full stability problem numerically, and find very good agreement with our analytical results.

The upshot of the linear stability analysis of the PL is that in the miscibility gap $\rho_0 \in (1, \rho_\ell)$, where

$$\rho_\ell(\lambda, w, w_1) = 1 + \frac{1}{2}\frac{w}{\lambda + 2w_1}, \tag{4.35}$$

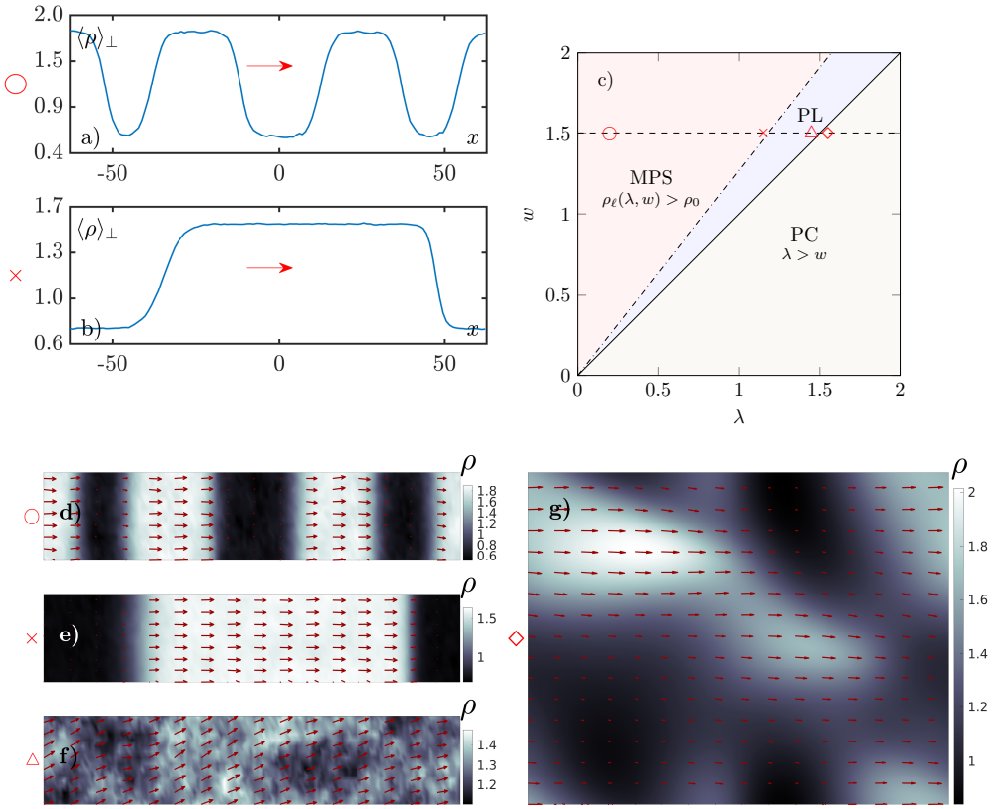


Fig. 4.4 Panels a) - b) and d) - g) illustrate the ordered phases of the HVM in eqs. (4.33) and (4.34): a) - b) show two MPS profiles of the system, where in a) the traveling bands form a smectic arrangement and in b) a single solitonic band travels against an isotropic background. In both plots, $\langle\rho\rangle_\perp$ denotes an instantaneous (in time) average over the direction perpendicular to the motion indicated by red arrows ($\rightarrow$). c) Phase diagram of the model at fixed $\rho_0 = 1.28$, $\kappa = \lambda$ and $w_1 = w/2$, with data points ($\circ$, $\times$, $\triangle$, $\diamond$) corresponding to figures a) - b) and d) - g) (in order of increasing λ). Solid (——) and dash-dotted (— · —) lines correspond to the phase boundaries $w = \lambda$ and $\rho_\ell(\lambda, w) = \rho_0$ respectively, determined from the linear stability analysis. Figures a) - b) and d) - e) display microphase-separation (MPS), where the number of bands is seen to increase as λ is decreased. In figure f) the system is homogeneously polarized, while in g) where $\lambda > w$, both the MPS and homogeneous polar liquid (PL) phases are unstable and localized polar clusters (PC) form.

the PL is unstable to small perturbations. From simulations we observe in this region both spatially inhomogeneous phases seen in the microscopic models, i.e. MPS and PCs, separated by a phase boundary that appears at sufficiently large κ. Insight into this is also provided by the linear stability analysis, from which we find that the PL is linearly stable only when

$$\kappa < 2w_1. \tag{4.36}$$

In particular, when $\kappa > 2w_1$, we observe that PCs are formed both within the spinodal region and for larger ρ_0. We also note that within the region where MPS occurs, simulations indicate that the number of bands increases with decreasing λ.

Now observe that in the special case where $\kappa = \lambda$, $w_1 = w/2$ and the inequality in eq. (4.36) holds, we have that $\rho_\ell \in (5/4, 3/2)$. In particular, when ρ_0 is within this region, all three ordered phases can be realized by varying λ and w. This is illustrated in fig. 4.4, where we plot the resulting phase diagram for fixed $\rho_0 \in (5/4, 3/2)$ in the (λ, w)-plane. In fig. 4.4, we also show typical realisations of the ordered phases.

We may gain further insight into the physics of the HVM, and in particular the nature of the instability that appears at large κ, by appealing to a particularly clever observation made in [1, 35]. In particular, it was found that the HVM dynamics can be written in the form

$$\partial_t \boldsymbol{m} + \lambda \left(\boldsymbol{m}\cdot\nabla\right)\boldsymbol{m} = -\frac{\delta F}{\delta \boldsymbol{m}} + \boldsymbol{\eta}. \tag{4.37}$$

Expressed in this way, the equation is reminiscent of a vectorial model A (in the Halperin-Hohenberg classification [111]), with a self-advection piece, i.e. the λ-term, that explicitly violates time-reversal symmetry (TRS) as it cannot be written as a functional derivative [12] (see also chapter 2). In fact, TRS-violation in the HVM is a slightly more involved issue, and will be the main subject of discussion in chapter 6.

The functional $F \equiv F[\rho, \boldsymbol{m}]$ in eq. (4.37) can be expressed as

$$F[\rho, \boldsymbol{m}] = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \left(f(\rho, \boldsymbol{m}) + \frac{1}{2}\left(\nabla_\alpha m_\beta\right)^2 + \boldsymbol{m}\cdot\nabla\Phi(\rho, \boldsymbol{m}) \right). \tag{4.38}$$

It contains the local 'free-energy density' $f(\rho, \boldsymbol{m})$ which is of standard quartic form (in $\boldsymbol{m}$), i.e.

$$f(\rho, \boldsymbol{m}) = \frac{1}{2}(1-\rho)m^2 + \frac{1}{4}m^4, \tag{4.39}$$

with the notable exception that the coefficient of the quadratic term controlling the transition to the low-temperature phase depends explicitly on the local density ρ. In particular, when $\rho > 1$ locally, f attains the characteristic bi-stable form which marks the transition to the

ordered phase. Additionally, in eq. (4.37) the function Φ is defined by

$$\Phi(\rho, \boldsymbol{m}) = w_1 \rho - \frac{\kappa}{2} m^2, \tag{4.40}$$

and is often referred to as an effective pressure [1, 95, 96, 115]. More importantly, eqs. (4.38) and (4.40) imply that F is reduced when $\boldsymbol{m}$ aligns against gradients in ϱ and towards increasing m. It is this competing effect which culminates in an instability at large κ and leads to the PC phase, since F then favours increasing the magnetisation locally.

With this we conclude our general discussion of the phenomenology of the HVM. As promised, we will now see that the hydrodynamic theory derived by Farrell *et al.* [82] indeed corresponds to a single point in this theory space. In their article, they found, using a method similar to the one we outlined in section 3.5, that the Toner-Tu type hydrodynamic equations for the moments $\varrho = \int_0^{2\pi} \mathrm{d}\theta\, p$ and $\boldsymbol{m} = \int_0^{2\pi} \mathrm{d}\theta\, \boldsymbol{e} p$ that could be derived from the FPE in eq. (4.14) were given by

$$\partial_t \varrho + \nabla \cdot \boldsymbol{m} = 0, \tag{4.41}$$

$$\partial_t \boldsymbol{m} + \lambda \left(\boldsymbol{m} \cdot \nabla\right) \boldsymbol{m} = -a(\varrho_c - \varrho)\boldsymbol{m} - b m^2 \boldsymbol{m} + \frac{1}{16}\nabla^2 \boldsymbol{m} + \kappa \left(\frac{1}{2} \nabla \left(m^2\right) - \boldsymbol{m} \left(\nabla \cdot \boldsymbol{m}\right) \right) - \frac{1}{2} \nabla \varrho \,. \tag{4.42}$$

Moreover, the coefficients ϱ_c, a, b, λ and κ in eqs. (4.41) and (4.42) are given in terms of the temperature T and system size L by

$$\varrho_c = \frac{2T}{L^2}, \tag{4.43}$$

$$a = \frac{L^2}{2T}, \tag{4.44}$$

$$b = \frac{L^4}{8T^2}, \tag{4.45}$$

$$\lambda = \frac{3L^2}{16T}, \tag{4.46}$$

$$\kappa = \frac{5L^2}{16T}. \tag{4.47}$$

Consistency with the microscopic model can straightforwardly be checked from this representation. Looking for constant homogeneous solutions $(\varrho_0, \boldsymbol{m}_0)$ to eqs. (4.41) and (4.42), it follows that $\varrho_0 = 1/L^2$ and $m_0 = 0$ for $T > T_c$ while for $T \leq T_c$ an ordered solution appears

with mean magnetisation

$$\bar{m} = \frac{m_0}{\varrho_0} = 2\sqrt{2T}(T_c - T)^{1/2} \tag{4.48}$$

$$= 2(T_c - T)^{1/2} + o\left((T_c - T)^{1/2}\right), \tag{4.49}$$

consistently with eq. (4.18).

A first notable difference between eqs. (4.43) to (4.47), for the parameters ϱ_c, a, b, λ and κ, and those presented in section 3.5 for the CTVM mapping, is that they do not depend on ϱ or $\boldsymbol{m}$, and are in this sense constant. Moreover, as we will now see, their dependence on L and T is only apparent. Indeed, rescaling $\boldsymbol{m} \to L^2\boldsymbol{m}/T$ and $\varrho \to L^2\,\varrho/T$, we find that the theory in eqs. (4.41) and (4.42) only describes the single point in the HVM theory space with coefficients

$$\varrho_c = 2, \tag{4.50}$$

$$a = \frac{1}{2}, \tag{4.51}$$

$$b = \frac{1}{8}, \tag{4.52}$$

$$\lambda = \frac{3}{16}, \tag{4.53}$$

$$\kappa = \frac{5}{16}. \tag{4.54}$$

This is not a fluke, however, and can be traced back to the FPE in eq. (4.14). Indeed, rescaling $p \to L^2 p/T$ maps the FPE onto an equation with only constant coefficients. The temperature then controls the normalisation of p via $L^2/T = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \int_0^{2\pi} \mathrm{d}\theta\, p$.

By comparing eqs. (4.50) to (4.54) to the (rescaled) HVM in eqs. (4.33) and (4.34), it is straightforward to compute that the AXYM point corresponds to $w = 4\sqrt{2}$, $w_1 = \sqrt{2}$, $\lambda = 3\sqrt{2}/2$, $\kappa = 5\sqrt{2}/2$. Looking back, we thus see that the AXYM indeed breaks the stability requirement in eq. (4.36), and so lies within the PC region of the HVM when $\varrho_0 = L^{-2}\int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\ \varrho = \beta > 2$. Moreover, from our analysis above, we see that the instability is caused by the κ-term in F in eq. (4.38), which favours increasing the magnetisation locally.

Next we will proceed by investigating a model that is closely related to the AXYM, which has the property that the local temperature T_i is bounded from below by the global temperature $T > 0$. With this minor change, we indeed find that the PL phase is stabilised, in agreement with our findings thus far.

4.5 Comparison with the non-extensive Active XY model

One of the key observations made in section 4.1 is that the local temperature $T_i \sim 1/\rho_i^{\mathcal{C}}$ in the AXYM, which can therefore drop arbitrarily close to zero when the local density $\rho_i^{\mathcal{C}}$ becomes large, consequently locally freezing the spin dynamics. Here, we aim to show that by making a simple change to the local temperature which bounds it from below by the global temperature T, so that $T_i \geq T > 0$, a stable polar liquid phase at large densities is recovered. A heuristic argument to support this can be made by observing that when the temperature is bounded in this way, then by sending the average global density ρ_0 to infinity we can make the spatial fluctuations in the local temperature arbitrarily small. Of course, this could not be done for the AXYM, since at large densities the system would still favour configurations where a patch of the system has local temperature $T_i \to 0$.

Importantly, the change we make to the dynamics imply that the interaction is no longer locally extensive, nor can it be derived from a potential. In the following, we therefore refer to this as the **non-extensive Active XY model** (NeAXYM). This has also appeared previously in the literature, in the context of motility-induced phase-separation (MIPS) in the presence of a density-dependent motility [84], although there is limited material discussing it at constant self-propulsion speed. Here we will show that the suppression of the local alignment strength in the NeAXYM, introduced by replacing ρ_0 (AXYM) with the variable local density $\rho_i^{\mathcal{C}}$ (NeAXYM) in the local inverse temperature β_i, is sufficient to stabilise the polar liquid.

For this model, we thus consider a local inverse temperature $\beta_i = \beta m_i^{\mathcal{C}}/\rho_i^{\mathcal{C}}$, which implies that $T_i \geq T$ is bounded from below as promised. Specifically, the angular dynamics in the NeAXYM now reads

$$\dot{\theta}_i = \beta \boldsymbol{e}_i' \cdot \frac{\boldsymbol{m}_i^{\mathcal{C}}}{\rho_i^{\mathcal{C}}} + \sqrt{2}\xi_i. \tag{4.55}$$

Again, this differs from the corresponding eq. (3.19) for the CTVM in that the interaction strength does not have unit norm.

It is straightforward to see that the interaction term in the NeAXYM is again non-integrable, as was the case for the CTVM. Indeed, letting $F_i = \boldsymbol{e}_i' \cdot \boldsymbol{m}_i^{\mathcal{C}}/\rho_i^{\mathcal{C}}$, then it follows that

$$\partial_{\theta_j} F_i - \partial_{\theta_i} F_j = \mathcal{C}_{ij} \left(\frac{1}{\rho_i^{\mathcal{C}}} - \frac{1}{\rho_j^{\mathcal{C}}} \right) \boldsymbol{e}_j' \cdot \boldsymbol{e}_i', \tag{4.56}$$

which for general contact matrices $\mathcal{C}_{ij}$ does not vanish. Note also that, in analogy with the CTVM, eq. (4.55) implies that the limit where the interaction range $\ell_0 \to \infty$ is equilibrium, while the limit where the self-propulsion speed $v_0 \to 0$ is not. As for the CTVM, irreversibility in the NeAXYM is related to spatial fluctuations.

Interestingly, the MFT of the NeAXYM at infinite density shares many features with that of its extensive counterpart. To see this, note that by following the same procedure as before (see sections 3.3 and 4.2), the mean-field dynamics at $\rho_0 = \infty$ can straightforwardly be shown to be

$$\dot{\theta}_i = \beta \boldsymbol{e}_i' \cdot \frac{\boldsymbol{m}^{\mathcal{C}}(\boldsymbol{x}_i)}{\varrho^{\mathcal{C}}(\boldsymbol{x}_i)} + \sqrt{2}\xi_i, \tag{4.57}$$

where the local density field

$$\varrho^{\mathcal{C}}(\boldsymbol{x}) \equiv \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}' \int_0^{2\pi} \mathrm{d}\theta \, \mathcal{C}_0(|\boldsymbol{x} - \boldsymbol{x}'|) p(\boldsymbol{x}', \theta), \tag{4.58}$$

and the molecular field $\boldsymbol{m}^{\mathcal{C}}(\boldsymbol{x})$ is given by eq. (4.12) as before. Thus, writing down the FPE for the dynamics, we again obtain the non-linear equation

$$\partial_t p + \boldsymbol{e} \cdot \nabla p = \partial_\theta \left(\mathrm{e}^{\beta \boldsymbol{e} \cdot \boldsymbol{m}^{\mathcal{C}} / \varrho^{\mathcal{C}}} \partial_\theta \left(\mathrm{e}^{-\beta \boldsymbol{e} \cdot \boldsymbol{m}^{\mathcal{C}} / \varrho^{\mathcal{C}}} p \right) \right). \tag{4.59}$$

The steady-state solution p_{ss} to eq. (4.59) can be computed exactly as before, and is the same as that for the AXYM, i.e.

$$p_{\mathrm{ss}}(\boldsymbol{x}, \theta) = \frac{\mathrm{e}^{\beta \boldsymbol{e} \cdot \boldsymbol{M}}}{2\pi L^2 I_0(\beta M)}, \tag{4.60}$$

where M satisfies the self-consistency condition

$$M = \frac{I_1(\beta M)}{I_0(\beta M)}. \tag{4.61}$$

As before, this should be contrasted with the corresponding eq. (3.27) for the CTVM, where the molecular field has unit norm. Since eq. (4.60) is the same as for the AXYM, the critical temperature is still given by $T_c = 1/2$, which separates an isotropic solution with $M = 0$ for $T > T_c$ and a homogeneous PL solution with $M > 0$ for $T < T_c$. Moreover, the mean magnetisation $\bar{m} \sim (T_c - T)^{1/2}$ near T_c.

Momentarily we will show that, as claimed above, the homogeneous PL distribution of the NeAXYM in eq. (4.60) is indeed a stable solution to the non-linear FPE in eq. (4.59), also for $T < T_c$. Because of this, we briefly mention that by combining eqs. (4.57), (4.60) and (4.61), we can deduce an effective single-particle dynamics under the ordered steady-state distribution at $\rho_0 = \infty$ as for the CTVM in section 3.3. Specifically, it follows that under p_{ss}, the mean-field dynamics is given by

$$\dot{\theta}_i = \varphi(\beta) \boldsymbol{e}_i' \cdot \hat{\boldsymbol{M}} + \sqrt{2}\xi_i, \tag{4.62}$$

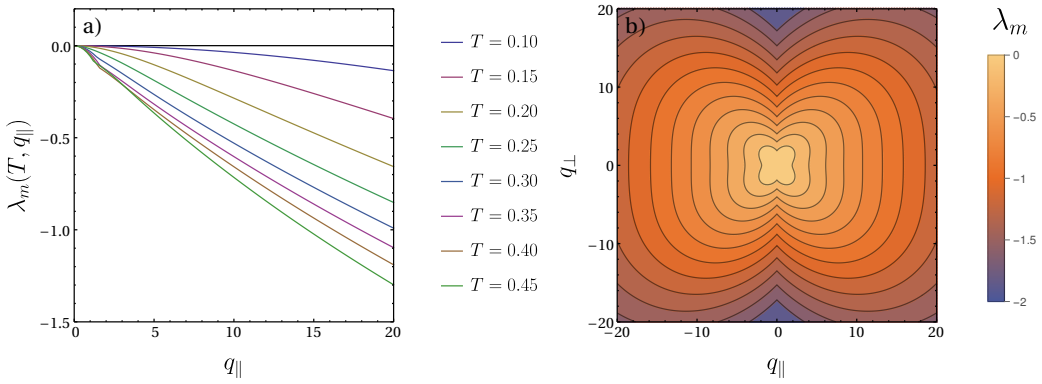


Fig. 4.5 Maximal eigenvalue $\lambda_m(T, \boldsymbol{q})$ in the linearised mean-field NeAXYM at $\ell_0 = 0$ as a function of: a) parallel (to $\boldsymbol{M}$) wave number $q_\parallel$ at $q_\perp = 0$, and b) $\boldsymbol{q} = (q_\parallel, q_\perp)$ at fixed $T = 0.4$. The maximal eigenvalue is always negative for all cases investigated.

where $\varphi(\beta) = \beta M(\beta)$. Interestingly, the dynamics in eq. (4.62) is identical to eq. (3.31) for the CTVM with a modified inverse temperature $\beta \to \varphi(\beta)$. In consequence, our results in eqs. (3.35) and (3.37) for the time-correlation function C and effective diffusion D_{eff} from section 3.3 carry over to the dynamics described by eq. (4.62) in a straightforward manner (to leading order) by replacing $\beta \to (T_c - T)^{1/2}/T_c$.

The linear stability analysis of the solution p_{ss} to the FPE in eq. (4.59) for NeAXYM essentially mirrors that performed for the CTVM and AXYM in sections 3.4 and 4.3, and so we do not present it in as much detail here. We only note that, with $\boldsymbol{M} = M\hat{\boldsymbol{x}}$, the equation for the Fourier coefficients $\delta p^R_{n,\boldsymbol{q}}$ is now given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta p^R_{n,\boldsymbol{q}} + \frac{i\boldsymbol{q}}{2}\cdot\begin{pmatrix}\delta p^R_{n-1,\boldsymbol{q}} + \delta p^R_{n+1,\boldsymbol{q}}\\ \delta p^I_{n-1,\boldsymbol{q}} - \delta p^I_{n+1,\boldsymbol{q}}\end{pmatrix} = -n^2\delta p^R_{n,\boldsymbol{q}} + \frac{\varphi n}{2}(\delta p^R_{n-1,\boldsymbol{q}} - \delta p^R_{n+1,\boldsymbol{q}}) - \boldsymbol{c}^R_n\cdot\delta\boldsymbol{m}^{\mathcal{C}}_{\boldsymbol{q}}, \\ + Mc_{n,\parallel}\delta\varrho^{\mathcal{C}}_{\boldsymbol{q}}, \tag{4.63}$$

where the constants $\boldsymbol{c}_n$ are again given by eq. (4.24), $c_{n,\parallel}$ is the component of $\boldsymbol{c}_n$ parallel to $\boldsymbol{M}$. The variation $\delta\varrho^{\mathcal{C}}_{\boldsymbol{q}}$, which is the response of $\varrho^{\mathcal{C}}(\boldsymbol{x})$ to the perturbation δp of p_{ss}, can be written

$$\delta\varrho^{\mathcal{C}}_{\boldsymbol{q}} = \int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}'\int_0^{2\pi}\mathrm{d}\theta\,\mathcal{C}_0(|\boldsymbol{x}-\boldsymbol{x}'|)\delta p(\boldsymbol{x}',\theta)\mathrm{e}^{-i\boldsymbol{q}\cdot\boldsymbol{x}} \tag{4.64}$$

$$= \mathcal{C}_{0,\boldsymbol{q}}\delta p^R_{0,\boldsymbol{q}}. \tag{4.65}$$

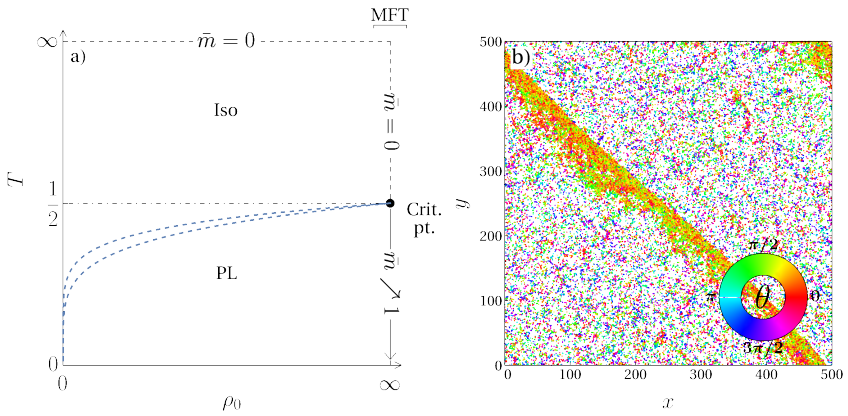


Fig. 4.6 a) Schematic phase diagram of the NeAXYM. The ordered phase can only be observed for T below $T_c = 1/2$. We have also indicated the shapes of the liquid and gas binodals $T_l(\rho_0)$ and $T_g(\rho_0)$ (blue dashed lines), although these cannot be inferred from MFT alone. b) Typical realisation of the banded phase in the NeAXYM in eq. (4.55). Parameters used for the simulation: $L = 500$, $\beta = 0.2$, $\ell_0 = 1$, $\rho_0 = 0.4$

The equation for the coefficients $\delta p^I_{n,\boldsymbol{q}}$, on the other hand, are still given by eq. (4.26). In particular, the only modification to the stability problem for the NeAXYM, compared with the AXYM, is the addition of the final term in eq. (4.63).

We have performed numerical investigations of the stability problem for the NeAXYM using the same procedure as before. In fig. 4.5 we show the results from these, where it is found, as expected, that the maximal eigenvalue $\lambda_m(T,\boldsymbol{q})$ is negative for all values $T < T_c$ and ranges of $\boldsymbol{q}$ checked. We thus conclude that that the PL solution $p_{\rm ss}$ is indeed linearly stable for the NeAXYM.

In addition, we have also performed simulations of the microscopic NeAXYM dynamics in eq. (4.55), and find that its behaviour is very similar to that of the CTVM. Imporantly, however, since the critical temperature is finite for the NeAXYM, we do not observe any ordered phase above T_c. Below, the liquid and gas binodals $T_l(\rho_0)$ and $T_g(\rho_0)$ mark the transitions between, respectively, the liquid and MPS phases, and the gaseous and MPS phases. The realisations of the dynamics in these regions look identical to those for the CTVM, so we refer to fig. 3.3 for illustrations of them. In fig. 4.6 we show the schematic phase diagram for the NeAXYM as inferred from MFT.

Finally, we mention that a mapping onto the HVM class can be constructed also for the NeAXYM, yielding a field theoretic description of the NeAXYM valid on the line $\rho_0 = \infty$.

In appendix A we show, using the same procedure as before, that this is given by

$$\partial_t \varrho + \nabla \cdot \boldsymbol{m} = 0, \tag{4.66}$$

$$\partial_t \boldsymbol{m} + \lambda \left(\boldsymbol{m} \cdot \nabla\right) \boldsymbol{m} = -a(\varrho_c - \varrho)\boldsymbol{m} - b m^2 \boldsymbol{m} + \frac{1}{16}\nabla^2 \boldsymbol{m} + \kappa \left(\frac{1}{2}\nabla\left(m^2\right) - \boldsymbol{m}\left(\nabla \cdot \boldsymbol{m}\right)\right) - w_1 \nabla \varrho, \tag{4.67}$$

where the coefficients depend on ϱ and $\boldsymbol{m}$ and are given by

$$\varrho_c = \frac{2\varrho}{\beta}, \tag{4.68}$$

$$\lambda = \frac{3}{16}\frac{\beta}{\varrho}, \tag{4.69}$$

$$a = \frac{\beta}{2\varrho}, \tag{4.70}$$

$$b = \frac{1}{8}\left(\frac{\beta}{\varrho}\right)^2, \tag{4.71}$$

$$\kappa = \frac{5}{16}\frac{\beta}{\varrho}, \tag{4.72}$$

$$w_{1,\alpha\beta} = \frac{1}{2}\left(1 + \frac{\beta}{4}\frac{m^2}{\varrho^2}\right)\delta_{\alpha\beta} - \frac{\beta}{4}\frac{m_\alpha m_\beta}{\varrho^2}. \tag{4.73}$$

Here, w_1 is matrix valued, and so component-wise the w_1-term in eq. (4.67) reads $w_{1,\alpha\beta}\nabla_\beta \varrho$. Moreover, the constant homogeneous solutions $(\varrho_0, \boldsymbol{m}_0)$ to eqs. (4.66) and (4.67) are again the isotropic with $m_0 = 0$ and for $T < T_c$ also the ordered solution with

$$\bar{m} = 2\sqrt{2T}\left(T_c - T\right)^{1/2} \tag{4.74}$$

as expected.

This concludes our discussion of the phenomenology of the microscopic continuous-time Vicsek models. Next, we include the linear stability analysis of the homogeneous PL for the HVM, from which we only cited the main results in section 4.4.

4.6 Linear stability analysis of the polar liquid in the HVM

The following section completes the derivation of the stability criteria cited in eqs. (4.35) and (4.36). Specifically, we will perform a linear stability analysis of the PL of the HVM in eq. (4.34) by employing the argument principle, which we introduced in section 3.5. Importantly, we are not able to deal with the full stability problem using this method alone,

and consider only wave vectors that are either parallel or perpendicular to the global magnetisation $\boldsymbol{m}_0$, leaving out the cases of oblique wave vectors. However, we have performed numerical checks of the full stability problem, and could find no discrepancies between our results from these and our analytical results for the ranges of parameters investigated.

We begin by writing $\rho = \rho + \delta\rho$ and $\boldsymbol{m} = \boldsymbol{m}_0 + \delta\boldsymbol{m}$ and expand eq. (4.34) to linear order in δ. Then, transforming eq. (4.41) and eq. (4.34) to Fourier space using $\delta\rho_{\boldsymbol{q}} = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \delta\rho(\boldsymbol{x})\mathrm{e}^{-i\boldsymbol{q}\cdot\boldsymbol{x}}$ and $\delta\boldsymbol{m}_{\boldsymbol{q}} = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \delta\boldsymbol{m}(\boldsymbol{x})\mathrm{e}^{-i\boldsymbol{q}\cdot\boldsymbol{x}}$, we find that

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} \delta\rho_{\boldsymbol{q}} \\ \delta m_{\parallel,\boldsymbol{q}} \\ \delta m_{\perp,\boldsymbol{q}} \end{pmatrix} = -\mathcal{L}_{\boldsymbol{q}}\begin{pmatrix} \delta\rho_{\boldsymbol{q}} \\ \delta m_{\parallel,\boldsymbol{q}} \\ \delta m_{\perp,\boldsymbol{q}} \end{pmatrix}, \tag{4.75}$$

where we have defined the linear coupling matrix $\mathcal{L}_{\boldsymbol{q}}$ by

$$\mathcal{L}_{\boldsymbol{q}} = \begin{pmatrix} 0 & iwq_{\parallel} & iwq_{\perp} \\ iw_1 q_{\parallel} - m_0 & \Gamma_{\parallel} + i\lambda m_0 q_{\parallel} & i\kappa m_0 q_{\perp} \\ iw_1 q_{\perp} & -i\kappa m_0 q_{\perp} & \Gamma_{\perp} + i\lambda m_0 q_{\parallel} \end{pmatrix}, \tag{4.76}$$

in addition to the damping coefficients

$$\Gamma_{\parallel} = 2(\rho_0 - 1) + q^2, \tag{4.77}$$

$$\Gamma_{\perp} = q^2. \tag{4.78}$$

In the above, the notation $v_{\parallel}$ and $v_{\perp}$ denote the components of a vector quantity $\boldsymbol{v}$ that are respectively parallel and perpendicular to $\boldsymbol{m}_0$, which we also assume throughout this section.

Rather than solving the full cubic polynomial characteristic equation of $\mathcal{L}$, we will study its roots only for wave vectors that are parallel or perpendicular to $\boldsymbol{m}_0$, i.e. those for which either $q_{\perp} = 0$ or $q_{\parallel} = 0$ respectively. In other words, we will study the roots of the two polynomial equations

$$\det\left(\sigma_{\parallel} - \mathcal{L}(q_{\parallel}, 0)\right) = 0, \tag{4.79}$$

$$\det\left(\sigma_{\perp} - \mathcal{L}(0, q_{\perp})\right) = 0, \tag{4.80}$$

and aim to deduce the conditions under which $\sigma_{\parallel} \equiv \sigma_{\parallel}(q_{\parallel})$ and $\sigma_{\perp} \equiv \sigma_{\perp}(q_{\perp})$ have positive real parts.

Starting with eq. (4.79), one may show that $\sigma_{\parallel}$ solves

$$\left(\sigma_{\parallel} - \Gamma_{\perp} - i\lambda m_0 q_{\parallel}\right) g_{\parallel}(\sigma_{\parallel}) = 0, \tag{4.81}$$

where the polynomial $g_\parallel$ is given by

$$g_\parallel(\sigma) = \sigma^2 - (\Gamma_\parallel + i\lambda m_0 q_\parallel)\sigma + w w_1 q_\parallel^2 + i w m_0 q_\parallel. \tag{4.82}$$

In order to determine the number of roots of $g_\parallel$ in the half-plane $\{\mathrm{Re}\,\sigma > 0\}$ we apply the argument principle (see section 3.5 for an introduction to this method). As before, we let Z_R denote the number of zeros of $g_\parallel$ inside the semi-circle contour $C_R = I_R \cup A_R$, where

$$I_R = [-iR, iR], \tag{4.83}$$

$$A_R = \left\{ Re^{i\theta} : \theta \in [-\pi/2, \pi/2] \right\}, \tag{4.84}$$

and aim to determine Z_R in the limit $R \to \infty$.

Again, it follows from our choice of contour C_R that the change in the argument of $g_\parallel$ along A_R is given simply by

$$\begin{aligned} \Delta_{A_R} \arg(g_\parallel) &= \log g_\parallel(Re^{i\pi/2}) - \log g_\parallel(Re^{-i\pi/2}) \\ &= 2\pi i + \mathcal{O}(R^{-1}). \end{aligned} \tag{4.85}$$

To compute Z_R in the limit $R \to \infty$, we are therefore left with having to find the change in the argument of $g_\parallel$ along I_R. Moreover, stability clearly requires that $Z_R \to 2$ as $R \to \infty$, which, from eq. (4.85), can only occur when

$$\lim_{R\to\infty} \Delta_{I_R} \arg(g_\parallel) = 2\pi i. \tag{4.86}$$

This occurs if and only if the image $g_\parallel(I_R)$ wraps around the origin once, as illustrated in fig. 4.7, or equivalently that the winding number of $g_\parallel(I_R)$ about the origin is 1.

To investigate when this occurs, we decompose $g_\parallel(iy)$ where $y \in [-R, R]$ into its real and imaginary parts, thus

$$\mathrm{Re}\; g_\parallel(iy) = -y^2 + \lambda m_0 q_\parallel y + w w_1 q_\parallel^2, \tag{4.87}$$

$$\mathrm{Im}\; g_\parallel(iy) = -\Gamma_\parallel y + w m_0 q_\parallel. \tag{4.88}$$

Firstly, from eq. (4.88) we see that we must require $\Gamma_\parallel > 0$, or the winding number of $g_\parallel(I_R)$ could only be 0 or -1 (recall that we are tracing the line segment $I_R = [-iR, iR]$ from iR to $-iR$ since C_R is traced counterclockwise). This is ensured so long as $\rho_0 > 1$, which we assume in the following. Furthermore, from eq. (4.87), it follows that the quadratic $\mathrm{Re}\; g_\parallel(iy)$

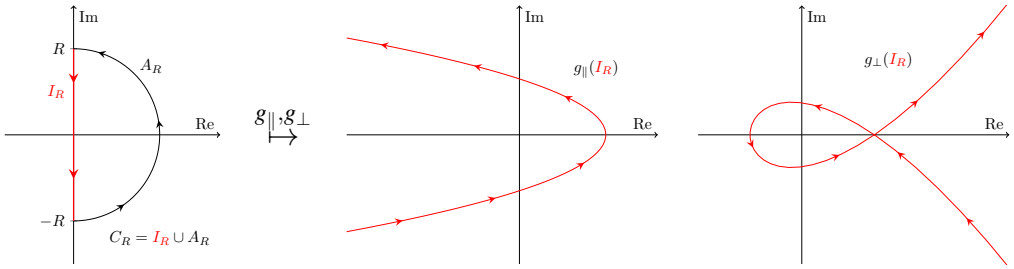


Fig. 4.7 On the semi-circle arc A_R, the polynomials $g_\parallel \sim R^2 e^{2i\theta}$ and $g_\perp \sim R^3 e^{3i\theta}$. Hence, in the limit $R \to \infty$, the number of zeros of $g_\parallel$ and $g_\perp$ in C_R is fully determined by the winding number of $g_\parallel(I_R)$ and $g_\perp(I_R)$ (respectively) about the origin.

has two distinct real roots for all $q_\parallel \neq 0$, given by

$$y_\pm = \frac{q_\parallel}{2}\left(\lambda m_0 \pm \sqrt{(\lambda m_0)^2 + 4ww_1}\right). \tag{4.89}$$

Thus, we only need to require that $\mathrm{Im}\, g_\parallel(iy_+) < 0$ and $\mathrm{Im}\, g_\parallel(iy_-) > 0$. One may show by standard means that this occurs if and only if

$$\rho_0 > 1 + \frac{1}{2}\frac{w}{\lambda + 2w_1}. \tag{4.90}$$

In conclusion therefore, it follows that all the roots of the characteristic equation of $\mathcal{L}$ are positive only when eq. (4.90) is satisfied.

Proceeding with eq. (4.80), i.e. the characteristic equation of $\mathcal{L}$ for wave vectors that are *perpendicular to* $\boldsymbol{m}_0$, we apply the argument principle once more. In this case, the roots $\sigma_\perp$ solve the cubic polynomial equation

$$g_\perp(\sigma_\perp) \equiv \sigma_\perp^3 - c_2\sigma_\perp^2 + c_1\sigma_\perp - c_0 = 0, \tag{4.91}$$

where the coefficients c_2, c_1 and c_0 are given by

$$c_2 = \Gamma_\parallel + \Gamma_\perp, \tag{4.92}$$
$$c_1 = \Gamma_\parallel\Gamma_\perp + q_\perp^2(ww_1 - \kappa^2 m_0^2), \tag{4.93}$$
$$c_0 = wq_\perp^2\left(w_1\Gamma_\parallel - \kappa m_0^2\right). \tag{4.94}$$

By considering the image of C_R under $g_\perp$, we see that along the semi-circle arc A_R, the change in the argument of $g_\perp$ is $3\pi i + \mathcal{O}(R^{-1})$. Thus, in this case we must require that the winding number of $g_\perp(I_R)$ is $\frac{3}{2}$ in order to have $Z_R \to 3$.

Clearly, the winding number of $g_\perp(I_R)$ can be $\frac{3}{2}$ if and only if $g_\perp(I_R)$ wraps around the origin in the way illustrated in fig. 4.7. Thus, to uncover the conditions under which this occurs, we again decompose $g_\perp(iy)$ into its real and imaginary parts:

$$\text{Re } g_\perp(iy) = c_2 y^2 - c_0, \tag{4.95}$$

$$\text{Im } g_\perp(iy) = -y^3 + c_1 y. \tag{4.96}$$

Assuming that $\rho_0 > 1$, we have $c_2 > 0$ which is necessary in order for the winding number of $g_\perp(I_R)$ to be positive. Furthermore, from eq. (4.95) we see that we must require $c_0 > 0$ so that the quadratic Re $g_\perp(iy)$ has two distinct real roots. This holds if and only if

$$\kappa < 2w_1. \tag{4.97}$$

Similarly, from eq. (4.96) we deduce that we must have $c_1 > 0$, or equivalently

$$\kappa^2 < 2 + \frac{ww_1}{\rho_0 - 1}, \tag{4.98}$$

so that the cubic Im $g_\perp(iy)$ has three distinct real roots. Finally, requiring that Re $g_\perp(\pm i\sqrt{c_1}) > 0$ one straightforwardly verifies that we are indeed guaranteed that $Z_R \to 3$. This last condition is equivalent to having

$$c_2 c_1 - c_0 > 0. \tag{4.99}$$

One may show that eq. (4.99) holds if and only if either

$$\kappa^2 < 3 + \frac{ww_1}{2(\rho_0 - 1)} \qquad \text{and} \qquad 2m_0^2(2-\kappa^2) + w\kappa > 0 \tag{4.100}$$

or

$$\left(2m_0^2(3-\kappa^2) + ww_1\right)^2 < 8m_0^2\left(2m_0^2(2-\kappa^2) + w\kappa\right) \tag{4.101}$$

For all simulations we perform, the conditions in eqs. (4.98) and (4.99) are satisfied. Because of this, we will only be concerned with the stability requirement in eq. (4.97). It is also worth highlighting once more that, despite the rather involved analysis undertaken here, we have not solved the full cubic characteristic equation of $\mathcal{L}$ and thus cannot claim to have fully identified all necessary and sufficient conditions for linear stability. Still, we have performed a numerical verification for several choices of the parameters of the theory, and find no discrepancies between our analytical predictions and the numerically estimated stability boundaries.

4.7 Conclusion

With this we conclude our investigation of the AXYM. As we have seen, this can be understood as a model of collective motion that can be related to the CTVM studied in chapter 3, albeit with a magnetisation-dependent local temperature. Our main conclusion is that the AXYM does not display an MPS or PL phase in contrast with the CTVM. Rather, the ordered phase of the model is a novel PC phase, that may be a condensation phenomenon, in which particles aggregate into dense clusters with strong polar order immersed in a high-temperature gaseous phase. This behaviour is striking also because several variations of the standard Vicsek dynamics has been investigated previously, and the consensus seems to be that the phase diagram of the model is reasonably stable against such modifications. An important counter-example is of course provided by the case of fore-aft asymmetric alignment rules [116], which yield a banded liquid type dynamics at low temperatures. In chapter 5, we will investigate the *entropy production rate* (EPR) of the CTVM and AXYM, and we will see there that the emergence of the PC phase has rather drastic consequences also in this context.

Many of the calculations we have performed in this chapter mirror those performed for the CTVM in chapter 3, in particular our treatment in sections 4.2, 4.3 and 4.5. In section 4.2, we showed that the AXYM has, in analogy with the CTVM, a homogeneous PL mean-field distribution at infinite density, albeit with a finite critical temperature T_c and different critical exponent 'β' for the scaling of the order parameter with temperature at the critical point. Simulations, on the other hand, did not yield a PL phase for temperatures $T < T_c$. As described above, we instead observe the PC phase for temperatures T lower than a density-dependent critical temperature $T_{\mathrm{cl}}(\rho_0)$. However, we still believe that $T_{\mathrm{cl}}(\rho_0) \sim 1/2$ for $\rho_0 \gg 1$ and thus connects with the critical point of the MFT.

To explain this behavior, we first performed a linear stability analysis of the AXYM in section 4.3. In contrast with the CTVM in section 3.4, we found here that a finite wavelength instability destabilises the homogeneous PL in the AXYM. This explains the apparent discrepancy between the MFT introduced in section 4.2 and our results from simulations. Furthermore, this allowed us to identify a temperature-dependent fastest growing mode $q_{\parallel}^{*}(T)$, which in future work can be compared with the typical linear size of clusters to test our predictions. For wave numbers q larger than a temperature-dependent threshold value $q_0(T)$, the linear stability analysis also predicts that the PL is linearly stable. In particular, this explains why one can seemingly stabilise a PL phase when the system size L is sufficiently small.

In section 4.4 we then proceeded to investigate the model we referred to as 'the HVM'. There we showed that the HVM supports all three ordered phases, i.e. MPS, PL and PC,

by simulating the dynamics and comparing this with a stability analysis of the PL phase in section 4.6. Moreover, it was shown that one may construct a mapping from the microscopic AXYM onto to a single point in the theory space of the HVM, and that this is indeed located in the PC region.

Interestingly, the HVM also provided us with further physical insight into the nature of the instability that engenders the PC phase, by linking the dynamics to a relaxational model-A type dynamics with an effective free-energy functional F. In particular, F contains an effective pressure term Φ, which can be reduced locally by increasing the magnetisation density. When this effect is strong enough, in the sense of eq. (4.36), the PL phase is destabilised and PCs are favoured by F.

Finally, in section 4.5, we also studied a non-extensive version of the AXYM, that we refer to as the NeAXYM. For this model, the local inverse temperature is bounded from above, and we found that the PL phase is stabilised at large densities both from numerical simulations as well as from a linear stability analysis of the corresponding MFT. In contrast with the CTVM, however, the NeAXYM has a critical point at infinite density and finite temperature $T_c = 1/2$.

Next we embark on the task of investigating the sources of irreversibility in the microscopic continuous-time models of collective motion considered in this chapter and chapter 3. We begin in chapter 5 with a discussion of numerical and analytical results related to the EPR in the microscopic CTVM, AXYM and NeAXYM. Our aim here will be to show that there is an intimate link between the EPR at large densities, where fluctuations are small, and discrete symmetry violations by the steady-state dynamics. In chapter 6, we follow this observation up with a more in-depth study of the EPR in the HVM and a related hydrodynamic theory of polar active matter at low noise.

Chapter 5

Entropy production of collective motion in microscopic polar active matter

Recently, considerable attention has been granted to understanding and quantifying the breaking of detailed balance and production of entropy in active matter systems [12, 28, 31, 34–36, 58, 117, 118]. One reason for this interest is that these systems break time-reversal symmetry (TRS) on the scale of each individual particle [1–3, 26], which distinguishes them from boundary driven systems such as sheared suspensions, or systems connected to multiple particle, heat or volume reservoirs. Consequently, a central question in the study of active matter concerns understanding how this microscopic violation of TRS has global consequences, and in particular to what extent it manifests in large scale currents in the system.

In active systems where self-propelled but otherwise isotropic particles interact via volume exclusion, the **entropy production rate** (EPR) can in general be expressed in terms of the so-called active work (see chapter 2), which quantifies the extent to which the self-propulsive force is producing motion. Over the past few years, a rather intriguing picture describing how fluctuations in the entropy production, or active work, can be related to global properties of the steady-state dynamics has been formed. Astonishingly, the upshot of this work is that the events that produce atypically large values of the stochastic entropy production can be associated with collective motion, and thus the formation of global currents, even if the dynamics a priori assumes no aligning interactions [117–120].

Although there are some contributions that also address the entropy production in polar active systems [30, 32, 35, 36], significantly less is known in comparison with the current state of the art of their scalar counterparts. Here we present preliminary results, which are yet to be published, from our investigations on the EPR in active systems interacting via alignment. We will consider all three models discussed in chapters 3 and 4, namely the

continuous-time Vicsek model (CTVM), Active XY model (AXYM) and non-extensive Active XY model (NeAXYM). Our goal will be to demonstrate that there is an intimate link between the spontaneous formation of a global order and current, in the form of collective motion, and the behaviour of the EPR at high densities where fluctuations are small. This establishes a link between the global violation of TRS and the microscopic one and, in particular, constrains fluctuations on both the hydrodynamic and microscopic scales.

We begin in section 5.1 by discussing the EPR in the CTVM, and present our results from simulations of the dynamics. We argue, in particular, that the *EPR per unit volume* (EPRV) and *EPR per particle* (EPRP) quantify the irreversibility on respectively the hydrodynamic and microscopic scales, and that they correspondingly constrain density and spin fluctuations. Then, in section 5.2, we show that the EPRP can be calculated exactly *at* infinite density in mean-field. To take account of finite-density fluctuations, we show in section 5.3 that the EPRV can be expanded in large average density ρ_0 by constructing a perturbative solution to the Dean equation [107] (DE, see section 3.4 of chapter 3) of the CTVM. Finally, in section 5.4 we consider how our results extend to the AXYM and NeAXYM before presenting our concluding remarks in section 5.5.

5.1 Informatic entropy production in the continuous-time Vicsek model

Collective motion in the Vicsek model represents a paradigmatic example of a true non-equilibrium steady-state, and is characterised by the spontaneous formation of a net global current [1, 2]. Importantly, the global violation of time-reversal symmetry (TRS) seen in such systems arises solely from a *microscopic* one, rather than from some externally imposed drive, by which we mean that total entropy production is a sum of the contributions from the dynamics of each individual active particle. Here, and through section 5.3, we exploit this correspondence to show that the properties of the total **entropy production rate** (EPR) in the microscopic CTVM places significant restrictions on both the hydrodynamic density and microscopic spin fluctuations. Moreover, from the symmetry properties of the global steady-state dynamics, we can make inferences about the non-equilibrium nature of each individual active particle, a point we also develop further in chapter 6.

Recall that in the CTVM, we consider (in d dimensions) a system of N active polar particles with positions $\{\boldsymbol{x}_i\}$ and spins $\{\boldsymbol{e}_i\}$ in a volume $\mathcal{V} = [0,L]^d$ that obey the coupled

Langevin equations

$$\dot{\boldsymbol{x}}_i = \boldsymbol{e}_i, \tag{5.1}$$

$$\dot{e}_{i\alpha} = \left(\delta_{\alpha\beta} - e_{i\alpha}e_{i\beta}\right) \circ \left(\beta \hat{m}^{\mathcal{C}}_{i\beta} + \sqrt{2}\xi_{i\beta}\right), \tag{5.2}$$

where $\circ$ denotes that we are employing the Stratonovich discretisation for the stochastic integral (see chapter 2). In eq. (5.2), the local magnetisation density $\boldsymbol{m}^{\mathcal{C}}_i$ is given by

$$\boldsymbol{m}^{\mathcal{C}}_i = \sum_{j=1}^{N} \mathcal{C}_{ij}\boldsymbol{e}_j, \tag{5.3}$$

while $\hat{\boldsymbol{m}}^{\mathcal{C}}_i = \boldsymbol{m}^{\mathcal{C}}_i / m^{\mathcal{C}}_i$ is the direction of the local magnetisation, and $\mathcal{C}_{ij} = \mathcal{C}_0(|\boldsymbol{x}_j - \boldsymbol{x}_i|)$ is an isotropic contact matrix (see chapter 3) that depends on the microscopic interaction length ℓ_0 via eq. (3.9). In what follows, it will also be convenient to consider the Itô form for the spin dynamics in eq. (5.2), explicitly

$$\dot{e}_{i\alpha} = -(d-1)e_{i\alpha} + \left(\delta_{\alpha\beta} - e_{i\alpha}e_{i\beta}\right)\left(\beta \hat{m}^{\mathcal{C}}_{i\beta} + \sqrt{2}\xi_{i\beta}\right), \tag{5.4}$$

which can be derived using the tools in chapter 2. Moreover, we recall that eq. (5.2) can be reduced to a representation in $d-1$ dimensions by considering $\boldsymbol{e}_i$ as a process on the unit $d-1$ dimensional sphere $\mathbb{S}^{d-1}$ in local coordinates. Specifically, in $d = 2$, we obtain

$$\dot{\theta}_i = \beta \boldsymbol{e}'_i \cdot \hat{\boldsymbol{m}}^{\mathcal{C}}_i + \sqrt{2}\xi_i, \tag{5.5}$$

where $\boldsymbol{e}_i \equiv \boldsymbol{e}(\theta_i)$ and $\boldsymbol{e}'_i = \boldsymbol{e}'(\theta_i)$. Finally, minding the slight abuse of notation, we also note that the two unit white noise terms in eqs. (5.2) and (5.5) can be related by $\xi_i = \boldsymbol{e}'_i \cdot \boldsymbol{\xi}_i = \boldsymbol{e}'_i \circ \boldsymbol{\xi}_i$.

An apparent issue arises when attempting to follow the procedure from chapter 2 to construct the EPR from eqs. (5.1) and (5.2). Specifically, due to the lack of noise in the equation for the position $\boldsymbol{x}_i$, we naïvely seem to lose the entropy balance equation in eq. (2.30) since the 'diffusion matrix' is not invertible, regardless of whether eq. (5.2) is written out in local coordinates. Because of this, we begin by showing that an entropy balance equation for the CTVM may indeed be recovered, and that to construct the entropy production from the trajectory level description of the dynamics, our definition of time-reversal must additionally include flipping the sign of all spins $\boldsymbol{e}_i$.

For ease of notation, we restrict our discussion in the following to $d = 2$, although the results we derive can be generalised to arbitrary d. First, we observe that the Fokker-Planck equation (FPE) for the distribution $p \equiv p(\boldsymbol{u})$ of the dynamics in eqs. (5.1) and (5.5), where

$\boldsymbol{u} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N, \theta_1, \ldots, \theta_N)$, is given by (see chapter 2)

$$\partial_t p = -\sum_{i=1}^{N} \left(\nabla_i \cdot \boldsymbol{j}_{x,i} + \partial_{\theta_i} j_{\theta,i} \right), \tag{5.6}$$

where ∇_i acts on the spatial variable $\boldsymbol{x}_i$, and the probability currents $\boldsymbol{j}_{x,i}$ and $j_{\theta,i}$ are given by

$$\boldsymbol{j}_{x,i} = \boldsymbol{e}_i p, \tag{5.7}$$

$$j_{\theta,i} = \beta \boldsymbol{e}_i' \cdot \hat{\boldsymbol{m}}_i^{\mathcal{C}} p - \partial_{\theta_i} p. \tag{5.8}$$

Thus, defining the system Shannon entropy $\mathcal{S}^{\mathrm{sys}} \equiv \int_{\mathcal{V}_u} \mathrm{d}\boldsymbol{u}\, p \log(hp)$, where phase space $\mathcal{V}_u = \mathcal{V}^N \times [0, 2\pi]^N$ and h has the dimensions of a phase space volume element, one may show using methods very similar to those presented in chapter 2 that

$$\dot{\mathcal{S}} = \dot{\mathcal{S}}^{\mathrm{sys}} + \dot{\mathcal{S}}^{\mathrm{res}}, \tag{5.9}$$

where $\dot{\mathcal{S}}$ is the total EPR and $\dot{\mathcal{S}}^{\mathrm{res}}$ is the entropy production of the reservoir. Moreover, we have explicit expressions for $\dot{\mathcal{S}}$ and $\dot{\mathcal{S}}^{\mathrm{res}}$ given by

$$\dot{\mathcal{S}} = \sum_{i=1}^{N} \int_{\mathcal{V}_u} \mathrm{d}\boldsymbol{u}\, \frac{j_{\theta,i}^2}{p}, \tag{5.10}$$

$$\dot{\mathcal{S}}^{\mathrm{res}} = \beta \sum_{i=1}^{N} \int_{\mathcal{V}_u} \mathrm{d}\boldsymbol{u} \left(\boldsymbol{e}_i' \cdot \hat{\boldsymbol{m}}_i^{\mathcal{C}} \right) j_{\theta,i}. \tag{5.11}$$

To derive eq. (5.9), the key observation one must make is that the entropy produced by the thermodynamic force $\boldsymbol{F}_{x,i} = -\nabla_i \log(hp)$ is always zero, i.e.

$$\int_{\mathcal{V}_u} \mathrm{d}\boldsymbol{u}\, \boldsymbol{j}_{x,i} \cdot \boldsymbol{F}_{x,i} = 0, \tag{5.12}$$

which follows from eq. (5.7) for the current $\boldsymbol{j}_{x,i}$. Finally then, combining eqs. (5.10) and (5.11), it therefore follows that the total steady-state EPR can be obtained from the temporal average

$$\dot{\mathcal{S}} = \lim_{\tau \to \infty} \frac{\beta}{2\tau} \sum_{i=1}^{N} \int_{-\tau}^{\tau} \hat{\boldsymbol{m}}_i^{\mathcal{C}} \circ \mathrm{d}\boldsymbol{e}_i \tag{5.13}$$

for a realisation of the dynamics, which in fact also extends to $d > 2$.

Interestingly, in order to construct a correspondence between the entropy balance in eq. (5.9) and time-reversal T on the level of individual realisations, we must include a sign

change in our definition of T. That is, we should consider instead

$$\mathrm{T}:\begin{cases} \boldsymbol{x}_i(t) \mapsto \boldsymbol{x}_i(-t), \\ \theta_i(t) \mapsto \theta_i(-t)+\pi, \end{cases} \tag{5.14}$$

so that $\overleftarrow{\boldsymbol{e}}_i = -\boldsymbol{e}_i$, where we denote by $\overleftarrow{\boldsymbol{e}}_i \equiv \boldsymbol{e}(\overleftarrow{\theta}_i)$, while $\overleftarrow{\theta}_i \equiv \mathrm{T}\theta_i$ and $\overleftarrow{\boldsymbol{x}}_i \equiv \mathrm{T}\boldsymbol{x}_i$ as in chapter 2. To see this, recall from chapter 2 that the path formulation of the stochastic dynamics expresses the path transition probability density $\mathcal{P}$ for a forward time trajectory on the time interval $[-\tau,\tau]$ in terms of the Freidlin-Wentzell action (FWA) $A = \sum_{i=1}^{N}\int_{-\tau}^{\tau}\mathrm{d}t\,\mathcal{A}_i$ by setting

$$\mathcal{P} \propto \mathrm{e}^{-A}, \tag{5.15}$$

where for the CTVM

$$\mathcal{A}_i = \begin{cases} \frac{1}{4}\left|\dot{\theta}_i - \beta\boldsymbol{e}_i'\cdot\hat{\boldsymbol{m}}_i^{\mathcal{C}}\right|^2 & \text{if} \quad \dot{\boldsymbol{x}}_i = \boldsymbol{e}_i, \\ \infty & \text{otherwise.} \end{cases} \tag{5.16}$$

In particular, eq. (5.1) for $\boldsymbol{x}$ manifests as *constraint* on the space of observable trajectories, i.e. the trajectories for which $\mathcal{P} > 0$. In words, this simply states that in any realisation of the dynamics, the velocity $\dot{\boldsymbol{x}}$ of the particle always has to be given exactly by $\boldsymbol{e}$. Because of this, 'standard' time-reversal would *not* yield an observable trajectory, since this would be a trajectory for which the velocity and spin point in opposite directions.

Combining eqs. (5.14) to (5.16) one may then show that the EPR in eq. (5.13) can be computed from

$$\dot{\mathcal{S}} = \lim_{\tau\to\infty}\frac{1}{2\tau}\log\frac{\mathcal{P}[\boldsymbol{u}]}{\overleftarrow{\mathcal{P}}[\boldsymbol{u}]}, \tag{5.17}$$

where $\overleftarrow{\mathcal{P}}[\boldsymbol{u}] = \mathcal{P}[\overleftarrow{\boldsymbol{u}}]$, and $\overleftarrow{\boldsymbol{u}} = (\overleftarrow{\boldsymbol{x}}_1,\ldots,\overleftarrow{\boldsymbol{x}}_N,\overleftarrow{\theta}_1,\ldots,\overleftarrow{\theta}_N)$. However, as we saw in section 2.6 in chapter 2, there is in no interpretation of the EPR in eq. (5.13) in terms of a heat flux. Notwithstanding, it still measures the breaking of time-reversal symmetry in the sense of eq. (5.17), and so should be regarded as an *informatic EPR* [28].

By performing simulations of the CTVM dynamics, we have numerically estimated the steady-state EPR $\dot{\mathcal{S}}$ in eq. (5.13) from long sample trajectories that are used to perform the temporal average. In fig. 5.1 we plot the corresponding *EPR per particle* (EPRP) $\dot{s} = \dot{\mathcal{S}}/N$ and *EPR per unit volume* (EPRV) $\dot{\sigma}_0 = \dot{\mathcal{S}}/L^2$. Clearly, these quantities are related by

$$\dot{s} = \frac{\dot{\sigma}_0}{\rho_0}, \tag{5.18}$$

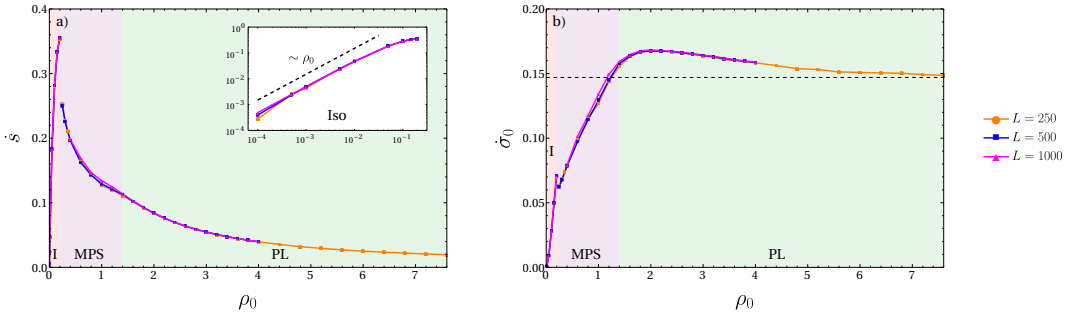


Fig. 5.1 EPR in the CTVM. In panel a), the EPRP $\dot{s}$ is shown as a function of the average density ρ_0 at fixed temperature $T = \beta^{-1}$, while in b), the corresponding EPRV $\dot{\sigma}_0$ is shown. At large densities, the EPRP $\dot{s}$ decays as a power law at rate $\sim 1/\rho_0$, which can be seen by observing that $\dot{\sigma}_0$ approaches a constant value $\dot{\sigma}_0^*$ as $\rho_0 \to \infty$. We also include in the inset in panel a) the scaling of $\dot{s}$ in the limit $\rho_0 \to 0$, which also exhibits power law decay $\dot{s} \sim \rho_0$. In both panels, I, MPS and PL mark, respectively, the isotropic, microphase separation, and polar liquid regions of the phase diagram of the CTVM (see chapter 3).

where $\rho_0 = N/L^2$ is the average density of the system. The key insight is that the EPRP $\dot{s}$ quantifies the extent to which each individual particle violates TRS since the active particles are indistinguishable, while the EPRV $\dot{\sigma}_0$ quantifies the extent of TRS violation by the *global* dynamics. In what follows we will investigate the behaviour of these two quantities in the limits of zero and infinite density.

A first observation one should make from fig. 5.1, is that the EPRP $\dot{s}$ decays to zero as $\rho_0 \to \infty$. This implies that at $\rho_0 = \infty$, the steady-state dynamics of each individual active particle must be equilibrium dynamics. Recall that this is exactly what we observed in section 3.3, where we showed explicitly that in the mean-field limit $\rho_0 \to \infty$ we obtain a Boltzmann distribution for the spin fluctuations, and an effective equilibrium single-particle steady-state dynamics (see eq. (3.31)). In retrospect, therefore, it should not have come as a surprise that the exact steady-state distribution for the CTVM was obtainable at $\rho_0 = \infty$.

Another interesting feature of the EPRP $\dot{s}$ is that it approaches zero in the infinite density limit as a power law, at a rate that can also be inferred from fig. 5.1 by observing that as $\rho_0 \to \infty$ the EPRV $\dot{\sigma}_0$ converges to a finite non-zero value

$$\dot{\sigma}_0^* \equiv \lim_{\rho_0 \to \infty} \dot{\sigma}_0 > 0, \tag{5.19}$$

meaning that $\dot{s} \sim 1/\rho_0$ for $\rho_0 \gg 1$. Notice also that eq. (5.19) indicates that the global dynamics remain irreversible in the limit of infinite density, in contrast with the microscopic

spin dynamics since the EPRP $\dot{s}$ vanishes in this limit as remarked above. In fact, as we shall see in chapter 6, eq. (5.19) implies that the fluctuating hydrodynamics near $\rho_0 = \infty$ cannot be described by effective equilibrium dynamics. However, as will become clear then, this global violation of TRS is in a sense 'weak'; the homogeneous polar liquid (PL) does not violate TRS by itself. Rather, TRS is broken finitely (cf. eq. (5.19)) by any infinitesimal density fluctuation about it. Another way to phrase this is to say that in the PL, *density ripples* diffuse *irreversibly* while *spin fluctuations* decay *reversibly*.

This idea, which we will develop in this chapter and the next, says that the limiting behaviour of $\dot{\sigma}_0$ for high densities can be significantly constrained by determining which discrete symmetries are violated (or not) by the mean global steady-state dynamics [36]. In consequence, going via eq. (5.18), we can pass from statements about the entropy that is produced by the global dynamics, to statements about the entropy production of each individual active particle. In particular, eq. (5.18) in fact provides a way to understand the microscopic TRS violation from a global one. For example, for the homogeneous PL, we will see in chapter 6 that $\dot{\sigma}_0$ must be $< \infty$ at $\rho_0 = \infty$ since the PL breaks T alone but not the combined *parity and time-reversal* operation PT.

Finally, we mention that, as can be observed from the inset in panel a) in fig. 5.1, the EPRP also has a power law decay $\sim \rho_0$ in the limit $\rho_0 \to 0$. The physical explanation for this behaviour is rather intuitive. In the CTVM the dynamics of a single active particle is reversible, which can be seen from eq. (5.13) by taking $N = 1$ and noting that for a single particle $\hat{\boldsymbol{m}}_i^{\mathcal{C}} = \boldsymbol{e}_i$. Thus,

$$\int_{-\tau}^{\tau} \hat{\boldsymbol{m}}_i^{\mathcal{C}} \circ \mathrm{d}\boldsymbol{e}_i = \int_{-\tau}^{\tau} \boldsymbol{e}_i \circ \mathrm{d}\boldsymbol{e}_i \tag{5.20}$$

$$= 0, \tag{5.21}$$

and so the zero density limit of $\dot{s}$ must vanish. The linear scaling then simply reflects the fact that the EPR is dominated by pairwise interactions whose event rate per particle is linear in the density.

5.2 Mean-field theory of the entropy production rate

Using the mean-field theory (MFT) developed in chapter 3, we will see in this section that making analytical predictions of the EPRP in eq. (5.13) in the limit $\rho_0 \to \infty$ can be done without much trouble. However, an important limitation of this approach is that the MFT assumes that the limit of infinite density has *already been taken*. Correspondingly, the MFT of the CTVM describes only the dynamics of a single active particle interacting with an

ambient molecular field. While this is sufficient to calculate the EPRP at infinite density, it cannot be used to derive the limiting value $\dot{\sigma}_0^*$ of the EPRV in eq. (5.19). To obtain this one has to be able to compute the EPRV at finite density first, before taking the limit. In fact, this is a reflection of a deeper insight, namely that in order to account for a finite $\dot{\sigma}_0^*$, one has to include finite-density fluctuations. As we alluded to above, the physical explanation for this is that on the global scale TRS is broken only at the level of density fluctuations in the PL, rather than by the mean global dynamics.

In this section we will begin by demonstrating explicitly that the EPRP $\dot{s} \to 0$ as $\rho_0 \to \infty$ using the MFT from chapter 3. Then, in section 5.3, we continue by showing that the dominant contribution to $\dot{\sigma}_0$ is indeed $\mathcal{O}(1)$ in this limit by including corrections due to finite-density fluctuations.

To compute $\dot{s}$ at infinite density, recall from the MFT of the CTVM in section 3.3 that in the limit $\rho_0 \to \infty$,

$$\hat{\boldsymbol{m}}_i^{\mathcal{C}} \to \hat{\boldsymbol{m}}^{\mathcal{C}}(\boldsymbol{x}_i) \equiv \frac{\int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \int_0^{2\pi} \mathrm{d}\theta \, p(\boldsymbol{x},\theta) \mathcal{C}_0(|\boldsymbol{x}-\boldsymbol{x}_i|) \boldsymbol{e}(\theta)}{\left| \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \int_0^{2\pi} \mathrm{d}\theta \, p(\boldsymbol{x},\theta) \mathcal{C}_0(|\boldsymbol{x}-\boldsymbol{x}_i|) \boldsymbol{e}(\theta) \right|}, \tag{5.22}$$

where the one-particle distribution function p is defined by

$$p(\boldsymbol{x},\theta) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \delta(\boldsymbol{x}-\boldsymbol{x}_i)\delta(\theta-\theta_i). \tag{5.23}$$

The steady-state one-particle distribution p_{ss} of the CTVM may be computed exactly by solving the non-linear Fokker-Planck equation (FPE) in eq. (3.26), and is given explicitly by

$$p_{\mathrm{ss}}(\boldsymbol{x},\theta) = \frac{\mathrm{e}^{\beta \boldsymbol{e}\cdot\hat{\boldsymbol{M}}}}{2\pi L^2 I_0(\beta)}, \tag{5.24}$$

where $\hat{\boldsymbol{M}}$ is a constant vector of unit length and I_0 is the modified Bessel function (of the first kind) of order zero (see eq. (3.28) and [106]). In particular, in steady-state it follows that

$$\hat{\boldsymbol{m}}_i^{\mathcal{C}} \to \hat{\boldsymbol{M}}. \tag{5.25}$$

Now observe that since the particles are indistinguishable, it follows from eq. (5.13) that

$$\dot{s} = \lim_{\tau\to\infty} \frac{\beta}{2\tau} \int_{\tau}^{\tau} \hat{\boldsymbol{m}}_i^{\mathcal{C}} \circ \mathrm{d}\boldsymbol{e}_i \tag{5.26}$$

for any i. Thus, combining eqs. (5.25) and (5.26), and assuming we may swap the order in which the limits are taken, we obtain

$$\dot{s}^* = \lim_{\rho_0 \to \infty} \dot{s} \tag{5.27}$$

$$= \lim_{\tau \to \infty} \frac{\Delta \mathfrak{s}_i^{\mathrm{res}}}{2\tau} = 0 \tag{5.28}$$

where the stochastic EPRP of the reservoir (see chapter 2) $\Delta \mathfrak{s}_i^{\mathrm{res}}$ at $\rho_0 = \infty$ is given by

$$\Delta \mathfrak{s}_i^{\mathrm{res}} = \beta \int_{-\tau}^{\tau} \hat{\boldsymbol{M}} \cdot \mathrm{d}\boldsymbol{e}_i \tag{5.29}$$

$$= \beta \hat{\boldsymbol{M}} \cdot (\boldsymbol{e}_i(\tau) - \boldsymbol{e}_i(-\tau)) \,. \tag{5.30}$$

The expression in eq. (5.30) allows us to in principle fully classify the statistics of the EPRP fluctuations in the CTVM at infinite density. Indeed, combining eq. (5.30) with our results in appendix B for the spin autocorrelation function in mean-field, we obtain for small β the asymptotic expansion

$$\left\langle (\Delta \mathfrak{s}_i^{\mathrm{res}})^2 \right\rangle = \beta^2 \left(1 - \mathrm{e}^{-2\tau}\right) + \frac{\beta^4}{72} \left((29 + 24\tau)\, \mathrm{e}^{-2\tau} - 2\mathrm{e}^{-8\tau} - 27 \right) + \mathcal{O}(\beta^6). \tag{5.31}$$

In particular, in the limit $\tau \to \infty$ we have

$$\lim_{\tau \to \infty} \left\langle (\Delta \mathfrak{s}_i^{\mathrm{res}})^2 \right\rangle = \beta^2 - \frac{3}{8}\beta^4 + \mathcal{O}(\beta^6). \tag{5.32}$$

For future work, we aim to study more closely the fluctuations of the EPR in the CTVM with the aim of understanding the rare events that realise atypical values of the stochastic EPR. In light of this agenda, we believe that these results may be of value. Moreover, this work should aim to confirm these predictions from simulations of the dynamics, which we have opted not to pursue further here.

For the next section, another representation of the EPRP $\dot{s}$ which can be obtained from the above will also be useful. To derive this, observe that the integral in eq. (5.26) can be written in Itô form as a sum of two terms, thus

$$\int_{-\tau}^{\tau} \hat{\boldsymbol{m}}_i^{\mathcal{C}} \circ \mathrm{d}\boldsymbol{e}_i = \int_{-\tau}^{\tau} \hat{\boldsymbol{m}}_i^{\mathcal{C}} \cdot \mathrm{d}\boldsymbol{e}_i + \underbrace{\mathcal{C}_0(0) \int_{-\tau}^{\tau} \mathrm{d}t \, \frac{\left(\boldsymbol{e}_i \cdot \hat{\boldsymbol{m}}_i^{\mathcal{C}}\right)^2}{m_i^{\mathcal{C}}}}_{\Delta} \,. \tag{5.33}$$

The second integral in eq. (5.33), Δ, can be thought of as a self-interaction term. By this we mean that Δ would not enter if we chose a punctured contact kernel with $\mathcal{C}_0(0) = 0$, thereby not including particle i as a nearest neighbour of i itself. Moreover, it is easy to see that Δ is at most $\mathcal{O}(1/\rho_0)$ when $\rho_0 \gg 1$. Indeed, this follows from the observation that $m_i^{\mathcal{C}} \sim \rho_0$ while $\left(\boldsymbol{e}_i \cdot \hat{\boldsymbol{m}}_i^{\mathcal{C}}\right)^2 \leq 1$.

Now, using the Itô form in eq. (5.4) for the spin dynamics, we may write the first integral in eq. (5.33) as

$$\int_{-\tau}^{\tau} \hat{\boldsymbol{m}}_i^{\mathcal{C}} \cdot \mathrm{d}\boldsymbol{e}_i = 2\beta\tau - \int_{-\tau}^{\tau} \mathrm{d}t \left(\beta \boldsymbol{e}_i \cdot \hat{\boldsymbol{m}}_i^{\mathcal{C}} + 1\right) \boldsymbol{e}_i \cdot \hat{\boldsymbol{m}}_i^{\mathcal{C}} + \sqrt{2} \int_{-\tau}^{\tau} \mathrm{d}t\, \xi_i \boldsymbol{e}_i' \cdot \hat{\boldsymbol{m}}_i^{\mathcal{C}}. \tag{5.34}$$

Since the final Itô-stochastic integral in eq. (5.34) is mean zero, it therefore follows by combining eqs. (5.26), (5.33) and (5.34) that

$$\dot{s} = \beta^2 \left(1 - \left\langle \left(\boldsymbol{e}_i \cdot \hat{\boldsymbol{m}}_i^{\mathcal{C}}\right)^2 \right\rangle\right) - \beta \left\langle \boldsymbol{e}_i \cdot \hat{\boldsymbol{m}}_i^{\mathcal{C}} \right\rangle + \beta \mathcal{C}_0(0) \left\langle \frac{\left(\boldsymbol{e}_i \cdot \hat{\boldsymbol{m}}_i^{\mathcal{C}}\right)^2}{m_i^{\mathcal{C}}} \right\rangle, \tag{5.35}$$

valid also at finite ρ_0.

One may confirm that eq. (5.35) indeed gives us the correct result at $\rho_0 = \infty$ by taking the limit to obtain

$$\dot{s}^* = \beta^2 \left(1 - \left\langle \left(\boldsymbol{e}_i \cdot \hat{\boldsymbol{M}}\right)^2 \right\rangle\right) - \beta \left\langle \boldsymbol{e}_i \cdot \hat{\boldsymbol{M}} \right\rangle. \tag{5.36}$$

The remaining expectation values follow readily by again employing the steady-state distribution p_{ss} in eq. (5.24), and we find that

$$\left\langle \boldsymbol{e}_i \cdot \hat{\boldsymbol{M}} \right\rangle = \frac{I_1(\beta)}{I_0(\beta)}, \tag{5.37}$$

$$\left\langle \left(\boldsymbol{e}_i \cdot \hat{\boldsymbol{M}}\right)^2 \right\rangle = 1 - \frac{1}{\beta} \frac{I_1(\beta)}{I_0(\beta)}, \tag{5.38}$$

from which it immediately follows that $\dot{s}^* = 0$ as required.

Starting from the representation of the EPRP in eq. (5.35), we will show next that by including finite-density fluctuations, the dominant contribution to $\dot{\sigma}_0$ in the limit $\rho_0 \to \infty$ is indeed $\mathcal{O}(1)$. More specifically, we will show that the power law decay of $\dot{s}$ could only be of integer power. Combining this with our results from this section, we may conclude that $\dot{s} \sim 1/\rho_0^{\alpha}$ with $\alpha \geq 1$.

5.3 Corrections due to finite-density fluctuations

Having investigated the EPRP in the mean-field limit, we will now turn to investigate the rate at which it decays to zero. We will begin by developing the framework within which the EPRV $\dot{\sigma}_0$ can be studied, and show that the dominant contribution to this is $\mathcal{O}(1)$ for large $\rho_0 \gg 1$. In light of our discussion in the previous two sections, we will do this by systematically incorporating the contributions due to finite-density fluctuations about the homogeneous PL.

At finite density, one way to incorporate noise is to consider instead the Dean equation (DE) [107] rather than the non-linear FPE in eq. (3.26), which we briefly touched on in section 3.4 of chapter 3. This is an equation for the *fluctuating* empirical one-particle distribution function

$$p(\boldsymbol{x},\theta) = \frac{1}{N}\sum_{i=1}^{N}\delta(\boldsymbol{x}-\boldsymbol{x}_i)\delta(\theta-\theta_i). \tag{5.39}$$

In this section, we denote by p only the empirical distribution, and so despite the slight abuse of notation there should be no cause for confusion with eq. (5.23). As mentioned in section 3.4 and shown explicitly in appendix E, p satisfies the DE

$$\partial_t p + \boldsymbol{e}\cdot\nabla p = \partial_\theta\left(\partial_\theta p - \beta\boldsymbol{e}'\cdot\hat{\boldsymbol{m}}^{\mathcal{C}}p + \sqrt{2p/N}\Xi\right), \tag{5.40}$$

where Ξ is a unit white noise and $\hat{\boldsymbol{m}}^{\mathcal{C}} = \boldsymbol{m}^{\mathcal{C}}/m^{\mathcal{C}}$ is the direction of the local magnetisation

$$\boldsymbol{m}^{\mathcal{C}}(\boldsymbol{x}) = \int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}'\int_0^{2\pi}\mathrm{d}\theta\, p(\boldsymbol{x}',\theta)\mathcal{C}_0(|\boldsymbol{x}-\boldsymbol{x}'|)\boldsymbol{e}(\theta) \tag{5.41}$$

as before.

The central result of this section is that, by combining eqs. (5.13), (5.35) and (5.39), one may express the EPR of the CTVM as a spatial integral over an *EPR density* $\dot{\sigma}(\boldsymbol{x})$, i.e.

$$\dot{\mathcal{S}} = \int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\,\langle\dot{\sigma}(\boldsymbol{x})\rangle\,, \tag{5.42}$$

where $\dot{\sigma}$ is given explicitly by

$$\begin{aligned}\dot{\sigma}(\boldsymbol{x}) = N\beta\int_0^{2\pi}\mathrm{d}\theta\left[\beta\left(1-\left(\boldsymbol{e}(\theta)\cdot\hat{\boldsymbol{m}}^{\mathcal{C}}(\boldsymbol{x})\right)^2\right)-\boldsymbol{e}(\theta)\cdot\hat{\boldsymbol{m}}^{\mathcal{C}}(\boldsymbol{x})\right]p(\boldsymbol{x},\theta)\\ +\beta\mathcal{C}_0(0)\int_0^{2\pi}\mathrm{d}\theta\,\frac{\left(\boldsymbol{e}(\theta)\cdot\hat{\boldsymbol{m}}^{\mathcal{C}}(\boldsymbol{x})\right)^2}{m^{\mathcal{C}}(\boldsymbol{x})}p(\boldsymbol{x},\theta).\end{aligned} \tag{5.43}$$

Our goal in what follows will be, using the DE in eq. (5.40), to show that $\dot{\sigma} \sim \mathcal{O}(1)$ for large $N \gg 1$. We will not attempt to do this by brute forcing eq. (5.43), however. Instead, we will show that there cannot be any divergent contribution.

We begin by assuming that for N large, the solution p to eq. (5.40) admits an expansion in small $\varepsilon = N^{-1/2}$, so that we may write

$$p = p_{\mathrm{ss}} + \varepsilon p_1 + \varepsilon^2 p_2 + \mathcal{O}\left(\varepsilon^3\right), \tag{5.44}$$

where p_{ss} is the steady-state distribution in eq. (5.24). The local magnetisation $\boldsymbol{m}^{\mathcal{C}}$ can then correspondingly be expanded as

$$\boldsymbol{m}^{\mathcal{C}} = \frac{1}{L^2}\frac{I_1(\beta)}{I_0(\beta)}\hat{\boldsymbol{M}} + \varepsilon \boldsymbol{m}_1^{\mathcal{C}} + \varepsilon^2 \boldsymbol{m}_2^{\mathcal{C}} + \mathcal{O}\left(\varepsilon^3\right). \tag{5.45}$$

Substituting this back into the DE in eq. (5.40) we obtain a hierarchy of equations for the functions p_k. For example, the first non-trivial equation for p_1 is given explicitly by (taking $\hat{\boldsymbol{M}} = \hat{\boldsymbol{x}}$ for simplicity)

$$\partial_t p_1 + \boldsymbol{e}\cdot\nabla p_1 = \partial_\theta\left(\partial_\theta p_1 + \beta \sin\theta\, p_1 - L^2\beta\frac{I_0(\beta)}{I_1(\beta)} p_{\mathrm{ss}} \cos\theta\, m_{1,\perp}^{\mathcal{C}} + \sqrt{2p_{ss}}\Xi\right), \tag{5.46}$$

where we have decomposed $\boldsymbol{m}_1^{\mathcal{C}} = \left(m_{1,\parallel}^{\mathcal{C}}, m_{1,\perp}^{\mathcal{C}}\right)$ into components that are parallel and perpendicular to $\hat{\boldsymbol{M}}$.

In general, the p_k will solve an inhomogeneous linear equation of the form

$$\partial_t p_k = \mathcal{L} p_k + \Delta_k, \tag{5.47}$$

where the linear operator $\mathcal{L}$, which can be inferred from eq. (5.46), only depends on p_{ss}. The driving terms Δ_k must be derived explicitly for each order, although for $k = 1$ we can see directly from eq. (5.46) that

$$\Delta_1 = \partial_\theta \sqrt{2p_{\mathrm{ss}}}\Xi. \tag{5.48}$$

The formal solution to eq. (5.47) can be constructed in the standard way in terms of the Green's function g of the linear operator $\partial_t - \mathcal{L}$, and is given by

$$p_k(\boldsymbol{x},\theta,t) = \int_{-\infty}^{\infty} \mathrm{d}t' \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}' \int_0^{2\pi} \mathrm{d}\theta' g(\boldsymbol{x},\boldsymbol{x}',\theta,\theta',t,t')\Delta_k(\boldsymbol{x}',\theta',t') \tag{5.49}$$

$$\equiv G[\Delta_k], \tag{5.50}$$

where g is the solution to the Green's equation

$$(\partial_t - \mathcal{L})\, g(\boldsymbol{x},\boldsymbol{x}',\theta,\theta',t,t') = \delta(\boldsymbol{x}-\boldsymbol{x}')\delta(\theta-\theta')\delta(t-t'). \tag{5.51}$$

Note also that in section 3.4 we showed that the spectrum of $\mathcal{L}$ is strictly negative, and so we expect the hierarchy in eq. (5.47) and the solution in eq. (5.49) to be well behaved.

Returning now to the EPR $\dot{\mathcal{S}}$ in eq. (5.42), we see that by combining it with the expansion for p in eq. (5.44) we may correspondingly expand it as

$$\dot{\mathcal{S}} = \frac{\dot{\mathcal{S}}_{-1}}{\varepsilon^2} + \frac{\dot{\mathcal{S}}_{-1/2}}{\varepsilon} + \dot{\mathcal{S}}_0 + \varepsilon\dot{\mathcal{S}}_{1/2} + \mathcal{O}(\varepsilon^2) \tag{5.52}$$

Our goal in the following will be to prove that all terms of odd order in ε (half-integer order in N) in this expansion must necessarily vanish. From eqs. (5.19) and (5.52), and the fact that $\dot{s}^* = 0$, it then immediately follows that for the CTVM the EPRV $\dot{\sigma}_0^*$ at infinite density is given by

$$\dot{\sigma}_0^* = \frac{\dot{\mathcal{S}}_0}{L^2}. \tag{5.53}$$

This result is striking, because it clearly demonstrates that the finite-density fluctuations, which generate the contribution $\dot{\mathcal{S}}_0$ to $\dot{\mathcal{S}}$, are responsible for the non-zero limiting value $\dot{\sigma}_0^*$ of the EPRV. In particular, the MFT developed in section 3.3, which as we recalled in section 5.2 only describes a single particle *at* infinite density, could not account for this result alone, even if it is sufficient to determine the limiting value $\dot{s}^*$ of the EPRP.

To prove that all terms of half-integer order in the expansion in eq. (5.52) must necessarily vanish, we begin by introducing a useful technique for counting powers of ε. When expanding the DE in eq. (3.52), one obtains a bunch of terms, each multiplied by some power of ε which we call the *order* of the term. To simplify our presentation in what follows, we therefore introduce the notation $[\![\cdot]\!]$ to denote this order. For example, the orders of p_k and Δ_k are $[\![p_k]\!] = k$ and $[\![\Delta_k]\!] = k$ respectively. It also makes sense to define $[\![\partial_\theta]\!] = [\![\nabla]\!] = 0$, and similarly for integration, since the derivatives themselves do not carry any factors of ε. In particular, in what follows we will simply avoid writing out derivatives, and so for our purposes a term such as $p_1\partial_\theta p_2$ is just the same as $p_1 p_2$.

It should also be clear that the driving term Δ_k in eq. (5.47) is a linear combination of terms, each of order k, and that it can be written as a functional of p_j for $j < k$ and Δ_1, i.e.

$$\Delta_k \equiv \Delta_k[p_1,\ldots,p_{k-1};\Delta_1]. \tag{5.54}$$

Thus, if a term t in Δ_k contains γ_j factors of p_j and $\delta \in \{0,1\}$ factors of Δ_1, then we must have

$$k = [\![t]\!] \tag{5.55}$$
$$= \left[\!\!\left[\Delta_1^{\delta} p_1^{\gamma_1} \cdots p_{k-1}^{\gamma_{k-1}}\right]\!\!\right] \tag{5.56}$$
$$= \delta + \sum_{j=1}^{k-1} j\gamma_j. \tag{5.57}$$

For example, neglecting derivatives, the only terms that could appear in Δ_2 are $p_1\Delta_1$ and p_1^2.

From eq. (5.54) we also see that the formal solution in eq. (5.49) to the inhomogeneous linear equation in eq. (5.47) allows us to construct p_k iteratively from Δ_1. To do this, we begin by observing that

$$p_1[\Delta_1] = G[\Delta_1]. \tag{5.58}$$

Since G is a linear integral operator, this also implies that the only term that can appear on the right-hand side is Δ_1. Next, we see that

$$p_2[\Delta_1] = G[\Delta_2[p_1[\Delta_1]]]. \tag{5.59}$$

Again, since G is linear, and Δ_2 only contains $p_1\Delta_1$ and p_1^2, p_2 can only contain Δ_1^2. This pattern can clearly be repeated, and one obtains for p_k that

$$p_k[\Delta_1] = G[\Delta_k[p_1[\Delta_1], p_2[\Delta_2], \ldots]], \tag{5.60}$$

implying in particular that p_k only contains the term Δ_1^k.

Piecing everything together, we may finally conclude that, since Δ_1 is mean-zero and Gaussian, $\langle p_k \rangle = 0$ for all odd k by Wick's theorem. Similarly, in the EPR expansion in eq. (5.52) it is easy to convince oneself that $[\![\dot{\mathcal{S}}_{k/2}]\!] = k+2$, and so $\dot{\mathcal{S}}_{k/2}$ contains only terms of order $k+2$. Thus, it follows again from Wick's theorem that $\dot{\mathcal{S}}_{k/2}$ vanishes identically for all odd k.

Before concluding this section, we also mention that by introducing the nematic tensor field $\mathcal{Q}$ as

$$\mathcal{Q}_{\alpha\beta}(\boldsymbol{x}) = \int_0^{2\pi} \mathrm{d}\theta \left(e_\alpha(\theta) e_\beta(\theta) - \frac{1}{2}\delta_{\alpha\beta} \right) p(\boldsymbol{x}, \theta), \tag{5.61}$$

we may significantly simplify the equation for the EPR density $\dot{\sigma}$ in eq. (5.43). Indeed, using this we may perform the integral over θ and write instead

$$\dot{\sigma} = \frac{\varrho}{2}\left(\frac{\mathcal{C}_0(0)}{m^{\mathcal{C}}} + N\beta\right) + \left(\frac{\mathcal{C}_0(0)}{m^{\mathcal{C}}} - N\beta\right)\hat{m}^{\mathcal{C}}_{\alpha}\mathcal{Q}_{\alpha\beta}\hat{m}^{\mathcal{C}}_{\beta} - N\boldsymbol{m}\cdot\hat{\boldsymbol{m}}^{\mathcal{C}}. \tag{5.62}$$

In appendix A we show that the nematic tensor can be expanded in small β and ∇ near the critical point at $(\rho_0, T) = (\infty, \infty)$, where $T = \beta^{-1}$ and that we can write

$$\hat{m}^{\mathcal{C}}_{\alpha}\mathcal{Q}_{\alpha\beta}\hat{m}^{\mathcal{C}}_{\beta} = \frac{\beta}{8}m + \frac{1}{16}\left(\nabla\cdot\boldsymbol{m} - \frac{1}{m}(\hat{\boldsymbol{m}}\cdot\nabla)\,m^2\right) + \text{m.z.} \tag{5.63}$$

to quadratic order in β and ∇, where m.z. denotes a term of zero mean. In particular, using this we have a closure relation which allows us to express the EPR density purely in terms of the local density ϱ and magnetisation $\boldsymbol{m}$.

With this we conclude our discussion of the EPR in the CTVM. Next, we aim to investigate how the principles we have discussed in the previous sections can be applied also to models introduced in chapter 4. In particular, we will see that the emergence of the cluster phase leads to a drastically different scaling for the EPRP and EPRV.

5.4 The isotropic and polar cluster phases

Importantly, the scaling behaviour observed in section 5.1 for the EPRP and EPRV in the CTVM is not generic. Two other important types of behaviour can also be observed for large densities, and in this section we will see that at least one of these can be realised within the Active XY model (AXYM). More specifically, we will discuss the numerical results we have obtained for the EPR in the AXYM in addition to the non-extensive Active XY model (NeAXYM), introduced in chapter 4, and how these correspond with our preceding discussion on the CTVM. In chapter 6 we explore this further, and show that one may in fact relate the types of scaling behaviour to discrete symmetry violations by the steady-state dynamics.

First we recall that in the AXYM and NeAXYM, the angular dynamics can be expressed as a CTVM dynamics with a variable local inverse temperature β_i, i.e.

$$\dot{\theta}_i = \beta_i \boldsymbol{e}_i' \cdot \hat{\boldsymbol{m}}^{\mathcal{C}}_i + \sqrt{2}\xi_i, \tag{5.64}$$

while positions follow $\dot{\boldsymbol{x}}_i = \boldsymbol{e}_i$ as before. In the AXYM, the interaction term derives from a potential $u_{\mathrm{XY}} = \frac{1}{2\rho_0}\sum_{i,j=1}^{N}\mathcal{C}_{ij}\boldsymbol{e}_i\cdot\boldsymbol{e}_j$ and is therefore locally extensive (see chapter 4 for a

precise definition). Consequently, the local inverse temperature is given by $\beta_i = \beta m_i^{\mathcal{C}}/\rho_0$, and one may thus rewrite the angular dynamics in eq. (5.64) as

$$\dot{\theta}_i = \beta \boldsymbol{e}_i' \cdot \boldsymbol{m}_i^{\mathcal{C}} + \sqrt{2}\xi_i. \tag{5.65}$$

Because the local temperature $T_i = \beta_i^{-1}$ can drop arbitrarily close to zero, the ordered phase of the model is no longer the PL. Instead, at low temperatures $T = \beta^{-1}$, we observe a polar cluster (PC) phase in which high density and strongly polarised clusters of particles move in an isotropic gas. In chapter 4 we explained this destabilisation of the polar liquid in the AXYM by a local freezing effect.

In contrast to the AXYM, the phenomenology of the NeAXYM is analogous to that of the CTVM. This is constructed by setting the local inverse temperature to $\beta_i = \beta m_i^{\mathcal{C}}/\rho_i^{\mathcal{C}}$, where the local density $\rho_i^{\mathcal{C}}$ is given by

$$\rho_i^{\mathcal{C}} = \sum_{j=1}^{N} \mathcal{C}_{ij}. \tag{5.66}$$

In particular, the angular dynamics of the NeAXYM can be written as

$$\dot{\theta}_i = \beta \boldsymbol{e}_i' \cdot \frac{\boldsymbol{m}_i^{\mathcal{C}}}{\rho_i^{\mathcal{C}}} + \sqrt{2}\xi_i. \tag{5.67}$$

Crucially, as discussed in section 4.5, for the NeAXYM the local inverse temperature is bounded from below by $T_i \geq T$, which is sufficient to stabilise the polar liquid at low temperatures T. In contrast to the CTVM, however, the critical point at infinite density in the NeAXYM is located at finite temperature $T_c = 1/2$, and the gas and liquid binodals, respectively $T_g(\rho_0)$ and $T_l(\rho_0)$, are both bounded from above by the line $T = T_c$. Thus, for temperatures T above T_c, we only observe an isotropic phase.

As in section 5.1, we may construct the steady-state EPR in the AXYM and NeAXYM from the log ratio of forward and backward path probability densities using the time-reversal operation T defined in eq. (5.14) which also flips the sign of each spin. Applying the definition of the steady-state EPR $\dot{\mathcal{S}}$ in eq. (5.17), we then obtain

$$\dot{\mathcal{S}} = \lim_{\tau\to\infty} \frac{1}{2\tau} \sum_{i=1}^{N} \int_{-\tau}^{\tau} \beta_i \hat{\boldsymbol{m}}_i^{\mathcal{C}} \circ \mathrm{d}\boldsymbol{e}_i \tag{5.68}$$

The only difference between this expression and that for the CTVM in eq. (5.13) is that the constant temperature β becomes a local one, β_i, and appears within the integral. Thus, we expect the methods applied previously can be extended to the AXYM and NeAXYM. In

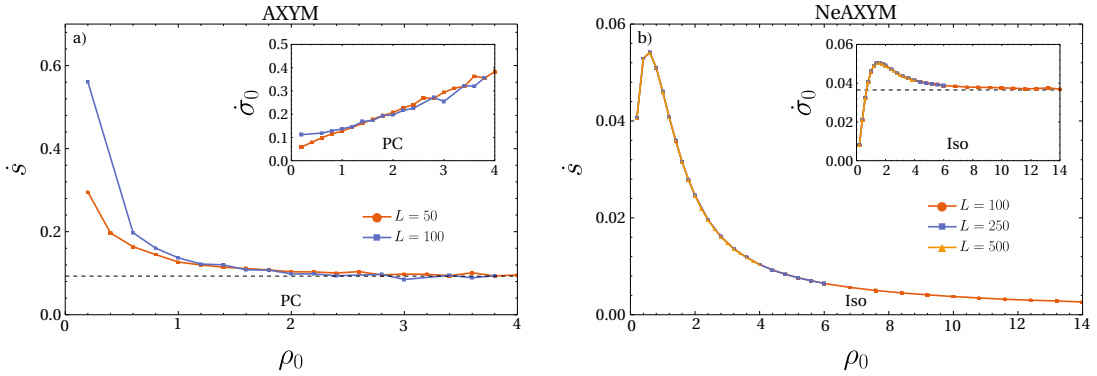


Fig. 5.2 Entropy production rate in the AXYM and NeAXYM. In both panels, the EPR is plotted as a function of the density ρ_0 at fixed temperature. In a), we show results obtained in the PC region of the AXYM, where the EPRP $\dot{s} \to \dot{s}^* > 0$ as $\rho_0 \to \infty$, which is confirmed by the fact that $\dot{\sigma}_0$ follows a linear slope (inset). Similarly, in panel b) we show results from the isotropic region (Iso) of the NeAXYM, which shows that the EPRV (inset) approaches a finite non-zero vale $\dot{\sigma}_0^* > 0$ as $\rho_0 \to \infty$.

particular, by employing the same reasoning which lead us to deduce that $\dot{s}^* = 0$ for the CTVM in eq. (5.28), the same should hold in both the isotropic and PL regions of the AXYM and NeAXYM.

In fig. 5.2, we show results from simulations, where as before, long realisations of the AXYM and NeAXYM dynamics have been utilised to compute the temporal average in eq. (5.68). For the AXYM, the EPR has been calculated numerically as a function of the average density ρ_0 in the ordered PC region of the phase diagram, whereas for the NeAXYM this has been done within the isotropic region, although both at fixed temperature T.

In the PC region of the AXYM, we observe that the EPRP $\dot{s} \sim \mathcal{O}(1)$ for large densities, as can be seen from fig. 5.2. Consequently, the EPRV follows a linear slope $\dot{\sigma}_0 \sim \rho_0$ and $\dot{\sigma}_0^* = \infty$. This indicates that in steady-state at $\rho_0 = \infty$, *both* the microscopic spin *and* global hydro-level dynamics are non-equilibrium. Consequently, we should *not* expect to find a Boltzmann distribution at $\rho_0 = \infty$ for the microscopic dynamics as we did for the CTVM, nor should we for the hydrodynamic theory. In this case, the collective dynamics breaks TRS 'strongly' at infinite density; as we shall see in chapter 6, this means that the mean global dynamics violates TRS, even when finite-density fluctuations are not present.

On the other hand, in the isotropic phase of the NeAXYM we observe that the EPRP $\dot{s} \sim \mathcal{O}(1/\rho_0)$ and consequently that the EPRV $\dot{\sigma}_0 \sim \mathcal{O}(1)$, which is the same as the scaling observed for the PL in the CTVM. We have also performed numerical checks of the isotropic phase of the AXYM (not shown), which produces similar behaviour. As we shall see next

Table 5.1 Classification of the EPR at high densities. Only three general limiting behaviours are possible for the polar active models studied in this thesis, summarised below; either TRS is broken strongly or weakly, or the limiting dynamics are effectively equilibrium. In this table, > 0 means that the EPRP or EPRV is positive and finite, while 'Irr.' and 'Rev.' mean that fluctuations are, respectively, irreversible or reversible near $\rho_0 = \infty$.

	Strong	Weak	Effectively equilibrium
EPRP $\dot{s}^*$	> 0	0	0
EPRV $\dot{\sigma}_0^*$	∞	> 0	0
Phase	PC	PL & Isotropic	Isotropic (HVM)
Spin	Irr.	Rev.	Rev.
Density	Irr.	Irr.	Rev.

in chapter 6, this is interesting, because for the isotropic phase of the hydrodynamic Vicsek model (HVM, see section 4.4) we instead observe 'effective equilibrium' dynamics at small noise (high density). This means that density ripples in the HVM decay reversibly in the isotropic phase, which they do not in the microscopic models, although we are as of yet unable to explain how this discrepancy arises from the coarse-graining. Additionally, we will see in chapter 6 that this notion of effective equilibrium at small noise indeed implies that an effective equilibrium theory can be obtained for the hydrodynamic fluctuations about the isotropic phase in the HVM, and that we obtain a Boltzmann distribution to lowest non-trivial order in the noise coefficient.

Summarising what we have seen so far in this chapter, we now know that there are really only three types of behaviours that can be seen for the EPR in the polar active systems we have studied in this thesis at low noise (summarised in table 5.4): i) both spin and density fluctuations are irreversible, ii) spin fluctuations are reversible and density fluctuations are irreversible, and iii) both spin and density fluctuations are reversible. Also, since the EPR expansion in eq. (5.52) does not contain terms of half-integer order, we cannot have any odd case for which e.g. $\dot{s} \to 0$ while $\dot{\sigma}_0 \to \infty$. In the next chapter, we will connect this to discrete symmetry violations by the mean global dynamics.

5.5 Conclusion

A very interesting feature of the EPR in the polar active systems we have investigated in this chapter, is that the properties of the mean global dynamics places significant constraints on the total entropy produced by the system at high densities [36]. In consequence, since the particles are indistinguishable and contribute the same amount of entropy production,

this also applies to the microscopic dynamics. Remarkably therefore, this allows us to make inferences about the non-equilibrium nature of the microscopic dynamics from the global properties of the system.

We explored this by first introducing the EPR in the CTVM in section 5.1. Here, we saw that even with the constraint on the space of observable trajectories imposed by the fact that positions follow $\dot{\boldsymbol{x}} = \boldsymbol{e}$ exactly, we could construct a consistent theory of entropy production for the model. In order to do this, we had to consider time-reversal in eq. (5.14) with a spin flip, which is different from the 'standard notion' of time-reversal employed in chapter 2. Subsequently, we investigated the EPR numerically, and observed that the EPRV in the CTVM approaches a constant $\dot{\sigma}_0^* > 0$ at large average densities ρ_0. Moreover, by invoking eq. (5.18) relating the EPRP and EPRV, we inferred that this implies that the EPRP $\sim 1/\rho_0$. The physical consequence of these results is that at large densities, spin fluctuations decay reversibly, whereas density fluctuations diffuse irreversibly.

To explain these results, we first determined that the EPRP $\dot{s} \to 0$ as $\rho_0 \to \infty$ analytically using the MFT of the CTVM from chapter 3. Moreover, this calculation allowed us to also infer an exact result for the stochastic entropy production $\Delta\mathfrak{s}_i^{\mathrm{res}}$ at infinite density in eq. (5.30). This expression allows us to in principle fully characterise the fluctuations of $\Delta\mathfrak{s}_i^{\mathrm{res}}$ in this limit, which may be compared and verified in future work with results from simulations.

Importantly, the MFT cannot capture finite-density fluctuations, and so it is unable to account for the observed $\dot{\sigma}_0^* > 0$. To remedy this, we included noise systematically in section 5.3 by means of the Dean equation in eq. (5.40), and performed a perturbative expansion of the EPRV at large densities. From this analysis, we found that the exponent characterising the decay (or divergence) of the EPRV could only take integer values. Thus, since the divergent term of the EPRV in the CTVM vanishes, the dominant contribution is at most $\mathcal{O}(1)$, and stems from the finite-density fluctuations. Moreover, the analysis we have performed here can clearly be seen to hold more generally, a fact we exploit in chapter 6.

Finally, in section 5.4, we discussed how our results obtained for the CTVM may be applied also to the AXYM and NeAXYM. In analogy with the CTVM, we saw that the EPRP must vanish at infinite density in the isotropic and PL phases of these models. In contrast, however, the EPRP approaches a constant positive value $\dot{s}^* > 0$ as $\rho_0 \to \infty$ in the PC phase.

Chapter 6

TRS-violations and entropy production in hydrodynamic theories of polar active matter

Precise identification of time-reversal symmetry (TRS) violations from the large scale dynamics of active matter is not always trivial, as the microscopic motion does not necessarily generate global net currents. For example, in field theories of motility-induced phase separation (MIPS) such as Active Model B (AMB), the absence of steady mass currents renders the steady-state deceivingly similar to equilibrium phase-separation [3, 26]. As we have done in chapter 5, one might therefore ask to what extent this qualitative notion of 'looking like equilibrium' is reflected quantitatively by the entropy production rate (EPR), measuring the extent of TRS violation by the stochastic dynamics. To address this question, we investigate in this chapter the dominant contribution to the EPR for two hydrodynamic theories of polar active systems at low noise, allowing us to distinguish between TRS violation at the levels of fluctuating and mean global dynamics, thus building on our results from chapter 5. In particular, from this analysis we determine the properties of the mean global dynamics which causes TRS violation at zero noise.

In this chapter we observe that dynamics in the small noise regime may be organised into three main classes based on the scaling of the EPR with the strength of local noise fluctuations. Within this scheme, 'strongly' non-equilibrium dynamics is characterised by a diverging EPR in the limit of small noise. The dominant divergent contribution stems from the 'ground-state' dynamics at zero noise, signifying the presence of steady TRS-violating currents that survive in this limit. When the EPR is finite in the limit of small noise, we further classify dynamics as either 'weakly' non-equilibrium or effectively equilibrium, where the latter corresponds to the case where the EPR vanishes in this limit. Note that the EPR can

also vanish on approach to a critical point while maintaining non-equilibrium behaviour but we do not address such cases here [15]. For dynamics of weak or effectively equilibrium type, the ground-state dynamics do not violate TRS and entropy is produced only at the level of fluctuations. However, for effectively equilibrium dynamics, detailed balance is restored at small noise where we recover Boltzmann statistics to lowest nontrivial order, while it is broken by a finite amount for any infinitesimal fluctuation in the weak case.

We will structure this chapter as follows. In section 6.1 we define the EPR of the hydrodynamic Vicsek model (HVM) from chapter 4 via the difference between time-forward and reversed path probability weights and discuss the implications of the magnetisation density changing sign on time-reversal. Here, we also introduce the asymptotic scaling relation for the EPR at small noise, and proceed in sections 6.2 and 6.3 to study its behaviour in the various phases of the model. Section 6.4 then introduces the model we refer to as the **diffusive flocking model** (DFM), in which we consider a more general constitutive equation for the advective current that includes noise and depends nonlinearly on both density and the local magnetisation. This allows us to consider both the case of a time-even and time-odd magnetisation density, and we explore in section 6.5 how this choice changes our results from sections 6.2 and 6.3. In addition, we demonstrate that in order to fully account for the entropy produced due to density currents in the flocking phase, we must also track this advective current. Finally, in section 6.6 we summarise our findings and present our concluding remarks.

6.1 Entropy production rate of the HVM

Even though fluctuations are essential in order to allow the time-reversed trajectory of a stochastic dynamics to be realisable under the time-forward dynamics, a well-defined entropy production rate (EPR) in the limit of vanishing noise strength can often be established [12, 35]. Previous studies have investigated this limit of vanishing noise in field theories of active matter, e.g. for Active model B (AMB) describing motility-induced phase separation (MIPS) on the hydrodynamic scale [12]. Here it was found that the scaling of the steady-state EPR at small noise depends on the phase of the system. For an isotropic system, the EPR in AMB is $\mathcal{O}(D)$, while it is $\mathcal{O}(D^0)$ when phase-separation has occurred, where D quantifies the strength of the noise. Of course, in scalar active systems like AMB, a diverging case such as $\mathcal{O}(1/D)$ is precluded because there are no steady-state currents. More specifically, from classical thermodynamic arguments we expect that in systems with such currents, including for example an electrical current in a wire for which the EPR $\dot{\mathcal{S}} = I^2R/T$ where T is the temperature, I is the current and R the resistance of the wire, exhibit a divergent scaling at

low noise. Because of this, it is interesting that Dadhichi *et al.* [35] noted that in their model of flocking the EPR scales as $\mathcal{O}(D^0)$ in both the homogeneous isotropic and polar liquid (PL) phases. Here we aim to provide some further details on the physics behind these results, and to organise them within a few unifying principles, by studying the EPR in the hydrodynamic Vicsek model (HVM).

For convenience, we recall here the definition of the HVM as discussed in section 4.4, specifically eqs. (4.37) to (4.40). This is a model of collective motion, in which we consider a fluctuating density ρ which is advected locally by the flow induced by the magnetisation density $\boldsymbol{m}$. In particular, we assume as in section 4.4 that the model lives in a periodic domain $\mathcal{V}$ in $d = 2$, and that ρ follows the continuity equation

$$\partial_t \rho = -\nabla \cdot \boldsymbol{J}, \tag{6.1}$$

where for the HVM the current $\boldsymbol{J}$ is given by the constitutive equation

$$\boldsymbol{J} = w\boldsymbol{m}. \tag{6.2}$$

Later, in section 6.4, we will consider a different constitutive equation for the density current, which one might expect more generally if for example the particles also interact via steric repulsion, or if they undergo Brownian motion in addition to self-propulsion due to interactions with an underlying substrate.

Further to this, the dynamics of the magnetisation density $\boldsymbol{m}$ in the HVM is given by

$$\partial_t \boldsymbol{m} + \lambda(\boldsymbol{m}\cdot\nabla)\boldsymbol{m} = -\frac{\delta F}{\delta \boldsymbol{m}} + \boldsymbol{\eta}, \tag{6.3}$$

and is in this form reminiscent of a vectorial model A (in the Halperin-Hohenberg classification [111]), with an additional self-advection piece given by the λ-term [1]. Moreover, in eq. (6.3), the free energy functional F can be expressed as

$$F[\rho,\boldsymbol{m}] = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \left(f(\rho,\boldsymbol{m}) + \frac{1}{2}\left(\nabla_\alpha m_\beta\right)^2 + \boldsymbol{m}\cdot\nabla\Phi \right), \tag{6.4}$$

and contains a local free energy density $f(\rho,\boldsymbol{m})$ of standard quartic form

$$f(\rho,\boldsymbol{m}) = \frac{1}{2}(1-\rho)m^2 + \frac{1}{4}m^4, \tag{6.5}$$

in addition to the local 'effective pressure'

$$\Phi(\rho,\boldsymbol{m}) = w_1\rho - \frac{\kappa}{2}m^2. \tag{6.6}$$

Fluctuations in the model are accounted for by the Gaussian white noise term $\boldsymbol{\eta}$ in eq. (6.3), which is taken to have autocovariance

$$\left\langle \eta_\alpha(\boldsymbol{x},t)\eta_\beta(\boldsymbol{x}',t')\right\rangle = 2D\delta_{\alpha\beta}\delta(\boldsymbol{x}-\boldsymbol{x}')\delta(t-t'), \tag{6.7}$$

so that D quantifies the strength of the noise. Importantly, even with only constant coefficients, we saw in section 4.4 that the theory in eqs. (6.1) and (6.3) captures all phases of the microscopic models investigated in chapters 3 and 4, i.e. the isotropic, microphase-separation (MPS), PL and polar cluster (PC) phases.

We construct the steady-state EPR of the HVM from the path transition probability density on the time interval $[-\tau,\tau]$ in the standard way [74] (see also chapter 2). For the dynamics in eqs. (6.1) and (6.3), this is defined via the Freidlin-Wentzell action (FWA) $\mathcal{A}$, where

$$\mathcal{A}[\rho,\boldsymbol{m}] = \frac{1}{4}\int_{-\tau}^{\tau}\mathrm{d}t\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\left|\partial_t\boldsymbol{m} + \lambda\left(\boldsymbol{m}\cdot\nabla\right)\boldsymbol{m} + \frac{\delta F}{\delta\boldsymbol{m}}\right|^2 \quad \text{if} \quad \partial_t\rho + w\nabla\cdot\boldsymbol{m} = 0, \tag{6.8}$$

and $\mathcal{A}=\infty$ otherwise. The transition probability density $\mathcal{P}[\rho,\boldsymbol{m}]$ of a trajectory $(\rho(t),\boldsymbol{m}(t))_{t\in[-\tau,\tau]}$ is then constructed in the standard way, i.e. by setting

$$\mathcal{P}[\rho,\boldsymbol{m}] \propto \exp\left(-\mathcal{A}[\rho,\boldsymbol{m}]/D\right). \tag{6.9}$$

It is important to note that in this formulation of the stochastic dynamics, eq. (6.1) acts as a constraint which limits the space of observable trajectories (i.e. those with $\mathcal{P}>0$). In order for all observable trajectories to have an observable time-reversal under $\mathcal{P}$, it is necessary therefore that the protocol T we choose for time-reversal involves flipping the local magnetisation, i.e.

$$\mathrm{T}:\begin{cases}\rho(\boldsymbol{x},t) \mapsto \rho(\boldsymbol{x},-t),\\ \boldsymbol{m}(\boldsymbol{x},t) \mapsto -\boldsymbol{m}(\boldsymbol{x},-t).\end{cases} \tag{6.10}$$

Indeed, this ensures that $\mathcal{A}<\infty$ if and only if $\overleftarrow{\mathcal{A}}<\infty$, where $\overleftarrow{\mathcal{A}}[\rho,\boldsymbol{m}] = \mathcal{A}[\overleftarrow{\rho},\overleftarrow{\boldsymbol{m}}]$ while $\overleftarrow{\rho}=\mathrm{T}\rho$ and $\overleftarrow{\boldsymbol{m}}=\mathrm{T}\boldsymbol{m}$, which can be seen directly from eq. (6.8). We may thus define a time-conjugate ensemble to eq. (6.9) by setting

$$\overleftarrow{\mathcal{P}}[\rho,\boldsymbol{m}] = \mathcal{P}[\overleftarrow{\rho},\overleftarrow{\boldsymbol{m}}], \tag{6.11}$$

which is supported on the same constrained space of trajectories as $\mathcal{P}$.

Interestingly, one observes that the functional $F[\rho,\boldsymbol{m}]$ defined in eq. (6.4) is not invariant under T. In particular, under this protocol F decomposes into T-symmetric and T-anti-symmetric contributions F^S and F^A respectively, where

$$\begin{aligned} F^A[\rho,\boldsymbol{m}] &= \frac{1}{2}\left(F[\rho,\boldsymbol{m}] - F[\rho,-\boldsymbol{m}]\right) \\ &= \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \boldsymbol{m}\cdot\nabla\Phi(\rho,\boldsymbol{m}), \end{aligned} \tag{6.12}$$

is the part of F that is odd in $\boldsymbol{m}$, and $F = F^S + F^A$. In fact, because of this we cannot interpret F as a true free-energy since it would clearly have to remain invariant under time-reversal. Moreover, this also implies that the system in eqs. (6.1) and (6.3) is out of equilibrium even when $\lambda = 0$, meaning that the self-advective contribution is not the only explicitly TRS violating component in the equations of motion, which may have been overlooked in previous work on the HVM [1].

Following standard treatments of stochastic thermodynamics (see chapters 2 and 5), we formally define the steady-state entropy production rate $\dot{\mathcal{S}}$ as the (log) ratio between the forward and time-reversed ensembles $\mathcal{P}$ and $\overleftarrow{\mathcal{P}}$ [12, 24, 25]

$$\dot{\mathcal{S}} \equiv \lim_{\tau\to\infty} \frac{1}{2\tau} \log \frac{\mathcal{P}[\rho,\boldsymbol{m}]}{\overleftarrow{\mathcal{P}}[\rho,\boldsymbol{m}]}. \tag{6.13}$$

Note that the definition in eq. (6.13) allows us to consider the EPR pathwise, i.e. as a functional of a trajectory $\dot{\mathcal{S}} \equiv \dot{\mathcal{S}}[\rho,\boldsymbol{m}]$. By construction, this functional satisfies the symmetry $\dot{\mathcal{S}}[\overleftarrow{\rho},\overleftarrow{\boldsymbol{m}}] = -\dot{\mathcal{S}}[\rho,\boldsymbol{m}]$ as can be readily observed from eq. (6.13).

In appendix D we show from eq. (6.8), eq. (6.9) and eq. (6.11) that $\dot{\mathcal{S}}$ can be expressed in integral form as

$$\dot{\mathcal{S}} = D^{-1}\int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \left\langle \frac{w}{2}|\boldsymbol{m}|^2(\nabla\cdot\boldsymbol{m}) - \left(\lambda\,(\boldsymbol{m}\cdot\nabla)\,\boldsymbol{m} + \frac{\delta F^A}{\delta \boldsymbol{m}}\right)\cdot\frac{\delta F^S}{\delta \boldsymbol{m}}\right\rangle. \tag{6.14}$$

We will also view $\dot{\mathcal{S}} \equiv \dot{\mathcal{S}}(D)$ as a function of the noise coefficient D and look to determine the asymptotic scaling

$$\dot{\mathcal{S}}(D) \sim D^{\chi}, \qquad D \ll 1. \tag{6.15}$$

Following an analogous derivation to that which we presented for the CTVM in section 5.3, one may show that $\chi \in \{-1,0,1,\ldots\}$ can only take integer values. As we will argue, this result also makes sense physically. We will see that when $\chi = -1$ the steady ground-state dynamics at $D = 0$ violates detailed balance in a pathwise sense, meaning that macroscopic

irreversible currents are inherent to the dynamics and are not solely observable at the level of fluctuations. On the other hand, when $\chi = 0$ the system is in a sense 'weakly non-equilibrium'. In particular, the deterministic $D = 0$ dynamics does not violate detailed balance, yet it is broken by a finite amount for any infinitesimal fluctuation and is never recovered as we send $D \to 0$. In contrast, when $\chi \geq 1$, the small noise limit is an effectively equilibrium regime where detailed balance is restored. Indeed, we will show that in the isotropic phase of the HVM, where $\chi = 1$, an expansion of the fields in small D allows the lowest order contribution beyond the steady $D = 0$ solution to be mapped onto an equilibrium system of decoupled underdamped harmonic oscillators.

In the following two sections we investigate analytically as well as numerically the scaling in eq. (6.15) in the various phases of the model in eqs. (6.1) and (6.3). We begin by studying the homogeneous isotropic and PL phases in section 6.2, where some analytical progress can be made at the fluctuating level. Subsequently, in section 6.3 we look at the nonlinear MPS and PC regimes.

6.2 Constant homogeneous ground-states

For the computations we present here, we will assume that the steady-state dynamics relaxes onto a 'ground-state' trajectory $(\rho_0(\boldsymbol{x},t),\boldsymbol{m}_0(\boldsymbol{x},t))_{t\in(-\infty,\infty)}$ [1] in the limit $D \to 0$. That is, we assume that the probability distribution over trajectories concentrates on a single path as $D \to 0$, and express this by

$$(\rho,\boldsymbol{m}) \overset{D\to 0}{\to} (\rho_0,\boldsymbol{m}_0), \tag{6.16}$$

where the limit is understood in the almost sure sense. We also assume that $(\rho_0,\boldsymbol{m}_0)$ solves eqs. (6.1) and (6.3) at $D = 0$ and that this limit is unique up to possible degeneracies arising from rotational invariance. Firstly our aim will be to classify the ground-states that satisfy the *pathwise* equilibrium condition

$$\dot{\mathcal{S}}[\rho_0,\boldsymbol{m}_0] = 0. \tag{6.17}$$

In particular, if eq. (6.17) holds, the dynamics must have $\chi > -1$.

Clearly, the pathwise equilibrium ground-states include those that are invariant under T in eq. (6.10), meaning that they satisfy

$$(\rho_0(\boldsymbol{x},t),\boldsymbol{m}_0(\boldsymbol{x},t)) = (\rho_0(\boldsymbol{x},-t),-\boldsymbol{m}_0(\boldsymbol{x},-t)), \tag{6.18}$$

[1] From here on we omit the subscript notation in $(\rho_0(\boldsymbol{x},t),\boldsymbol{m}_0(\boldsymbol{x},t))_{t\in(-\infty,\infty)}$ when talking about a trajectory.

which follows from the fact that $\dot{\mathcal{S}}[\bar{\rho},\bar{\boldsymbol{m}}] = -\dot{\mathcal{S}}[\rho,\boldsymbol{m}]$. The constant homogeneous isotropic state with $\rho_0 = \text{const.}$ and $m_0 = 0$ provides an example of such a state. On the other hand, the PL state with $\rho_0 > 1$ and $m_0 = \sqrt{\rho_0 - 1}$ clearly violates T alone. This is where rotational invariance arises as an important symmetry principle, because it implies that $\mathcal{P}$ (and thus $\dot{\mathcal{S}}$) must be invariant under the parity transformation

$$\mathrm{P}: \begin{cases} \rho(\boldsymbol{x},t) \mapsto \rho(-\boldsymbol{x},t), \\ \boldsymbol{m}(\boldsymbol{x},t) \mapsto -\boldsymbol{m}(-\boldsymbol{x},t), \end{cases} \tag{6.19}$$

which translates to the statement that a flock is equally likely to travel to the left as to the right. Now, if the ground-state trajectory $(\rho_0,\boldsymbol{m}_0)$ is PT-symmetric, i.e. it satisfies

$$(\rho_0(\boldsymbol{x},t),\boldsymbol{m}_0(\boldsymbol{x},t)) = (\rho_0(-\boldsymbol{x},-t),\boldsymbol{m}_0(-\boldsymbol{x},-t)), \tag{6.20}$$

then it follows that it is pathwise equilibrium. Perhaps surprisingly then, one realises that the constant homogeneous PL state in fact is pathwise equilibrium since it satisfies PT. However, this is to be expected: a charged particle gyrating at constant frequency in the plane perpendicular to an imposed magnetic field is certainly in equilibrium (although here T should be replaced by CT to include charge conjugation). Interestingly, these observations also imply that if rotational symmetry is broken a priori, for example if birds in a flock are preferentially travelling towards their nest or some source of food, then $\mathcal{P}$ would no longer be P-invariant and the PL state would have $\chi = -1$. We also note that the fact that $\chi > -1$ in both the homogeneous isotropic and PL states also follows directly from eq. (6.14) by evaluating the integral at constant $(\rho_0,\boldsymbol{m}_0)$.

In order to go beyond the $D = 0$ dynamics, we must take account of fluctuations. We do so by assuming that the fluctuating dynamics admit an expansion in small $\sqrt{D}$, following [12, 35], so that

$$\rho = \rho_0 + \rho_1 D^{1/2} + \mathcal{O}(D), \tag{6.21}$$

$$\boldsymbol{m} = \boldsymbol{m}_0 + \boldsymbol{m}_1 D^{1/2} + \mathcal{O}(D). \tag{6.22}$$

Furthermore, we restrict here to the case where $(\rho_0,\boldsymbol{m}_0)$ is constant and homogeneous. By substituting eqs. (6.21) and (6.22) into the equations of motion in eqs. (6.1) and (6.3) and

collecting terms, we obtain at order $D^{1/2}$

$$\partial_t \rho_1 = -w\nabla\cdot\boldsymbol{m}_1, \tag{6.23}$$

$$\partial_t \boldsymbol{m}_1 + \lambda\,(\boldsymbol{m}_0\cdot\nabla)\,\boldsymbol{m}_1 = -\frac{\delta F_L}{\delta \boldsymbol{m}}[\rho_1,\boldsymbol{m}_1] + \boldsymbol{\eta}_1. \tag{6.24}$$

Here, F_L is the quadratic functional

$$F_L[\rho,\boldsymbol{m}] = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\left(f_L(\rho,\boldsymbol{m}) + \frac{1}{2}(\nabla_\alpha m_\beta)^2 + \boldsymbol{m}\cdot\nabla\Phi_L(\rho,\boldsymbol{m})\right), \tag{6.25}$$

where we have defined the local free energy f_L by

$$f_L(\rho,\boldsymbol{m}) = \frac{a_0}{2}m^2 - \rho\boldsymbol{m}_0\cdot\boldsymbol{m} + (\boldsymbol{m}_0\cdot\boldsymbol{m})^2, \tag{6.26}$$

in addition to the linearised effective pressure

$$\Phi_L = w_1\rho - \kappa\boldsymbol{m}_0\cdot\boldsymbol{m}. \tag{6.27}$$

Moreover, $a_0 = 1-\rho_0+m_0^2$ and $\boldsymbol{m}_0$ satisfies $a_0 m_0 = 0$, while $\boldsymbol{\eta}_1$ is a mean-zero Gaussian white noise with

$$\left\langle \eta_{1\alpha}(\boldsymbol{x},t)\eta_{1\beta}(\boldsymbol{x}',t')\right\rangle = 2\delta_{\alpha\beta}\delta(\boldsymbol{x}-\boldsymbol{x}')\delta(t-t'). \tag{6.28}$$

We may perform a similar procedure in order to obtain an expansion of $\dot{\mathcal{S}}$ from eq. (6.14) in small D of the form

$$\dot{\mathcal{S}}(D) = \dot{\mathcal{S}}_{-1}D^{-1} + \dot{\mathcal{S}}_0 + \dot{\mathcal{S}}_1 D + \mathcal{O}(D^2). \tag{6.29}$$

In particular, recall from section 5.3 that these are the only possible terms that could enter in the expansion of $\dot{\mathcal{S}}$, i.e. that there are no terms of half-integer order in D. Also, from the asymptotic scaling relation in eq. (6.15) and the positivity of the EPR, we know that $\dot{\mathcal{S}}_k = 0$ for $k<\chi$ and that the leading order coefficient $\dot{\mathcal{S}}_\chi > 0$. By explicitly computing this expansion of $\dot{\mathcal{S}}$ to order D^0, one may show that

$$\begin{cases} \chi \geq 1, & \text{isotropic,} \\ \chi = 0, & \text{PL,} \end{cases} \tag{6.30}$$

and so $\chi > -1$ in the homogeneous phases as argued for above. For the explicit calculation of eq. (6.30), we refer to appendix D for the details.

From simulations we further find that in fact $\chi = 1$ in the isotropic phase, as shown in fig. 6.1, although we do not explicitly compute $\dot{\mathcal{S}}_1$. In the PL case, we obtain an explicit expression for $\dot{\mathcal{S}}_0$ given by

$$\begin{aligned}\dot{\mathcal{S}}_0 =& m_0^2(2w_1 - \kappa + \lambda)\int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \langle \rho_1 \partial_\parallel m_\parallel \rangle + m_0(w - 2m_0^2\kappa)\int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \langle m_\parallel \partial_\perp m_\perp \rangle \\ &+ 2m_0\kappa \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \big\langle (\partial_\perp m_\perp)(\nabla^2 m_\parallel) \big\rangle\,. \end{aligned} \tag{6.31}$$

In eq. (6.31) we use subscripts $\parallel$ and $\perp$ to denote components of $\boldsymbol{m}_1$ and ∇ that are parallel and perpendicular to $\boldsymbol{m}_0$ respectively.

Consistently with eq. (6.30), eq. (6.31) implies that $\dot{\mathcal{S}}_0 \sim m_0$ for $m_0 \ll 1$. In fact, this could have been predicted without explicitly performing the systematic expansion of $\dot{\mathcal{S}}$ in small D. Indeed, if we were to imagine expanding the integral expression for $\dot{\mathcal{S}}$ in eq. (6.14) using the series representations eqs. (6.21) and (6.22) at $m_0 = 0$, we see from simple power counting that the only combinations of fields that could possibly appear within the integrand at order D^0 are of the form

$$\langle \rho_2 \rangle, \qquad \langle \nabla \cdot \boldsymbol{m}_2 \rangle, \qquad \langle \rho_1^2 \rangle, \qquad \langle m_1^2 \rangle, \qquad \langle \rho_1 \nabla \cdot \boldsymbol{m}_1 \rangle, \ldots \tag{6.32}$$

Now, the symmetry $\dot{\mathcal{S}}[\bar{\rho}, \overleftarrow{\boldsymbol{m}}] = -\dot{\mathcal{S}}[\rho, \boldsymbol{m}]$ excludes all of ρ_2, ρ_1^2 and m_1^2 from entering, while $\nabla \cdot \boldsymbol{m}_2$ would just integrate to zero over $\mathcal{V}$. For the final average in eq. (6.32), observe that eq. (6.23) implies

$$\langle \rho_1 \nabla \cdot \boldsymbol{m}_1 \rangle = -w^{-1} \langle \rho_1 \partial_t \rho_1 \rangle \tag{6.33}$$

$$= \frac{-1}{2w} \partial_t \langle \rho_1^2 \rangle \tag{6.34}$$

$$= 0. \tag{6.35}$$

Hence, there are in fact no nontrivial contributions that could enter in the expansion of $\dot{\mathcal{S}}$ at order D^0 when $m_0 = 0$, so we must have $\chi \geq 1$ in the isotropic phase.

Since $\dot{\mathcal{S}}$ is $\mathcal{O}(D)$ in the isotropic phase, we in fact recover effective equilibrium in the limit $D \to 0$. To see this, we transform the linearised equations of motion eqs. (6.23) and (6.24) to Fourier space. Throughout this chapter we use the convention that the Fourier coefficients $h_{\boldsymbol{q}}$ of a function $h(\boldsymbol{x})$ are given by

$$h_{\boldsymbol{q}} = \mathcal{V}^{-1} \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, h(\boldsymbol{x}) \exp(-i\boldsymbol{x} \cdot \boldsymbol{q}), \tag{6.36}$$

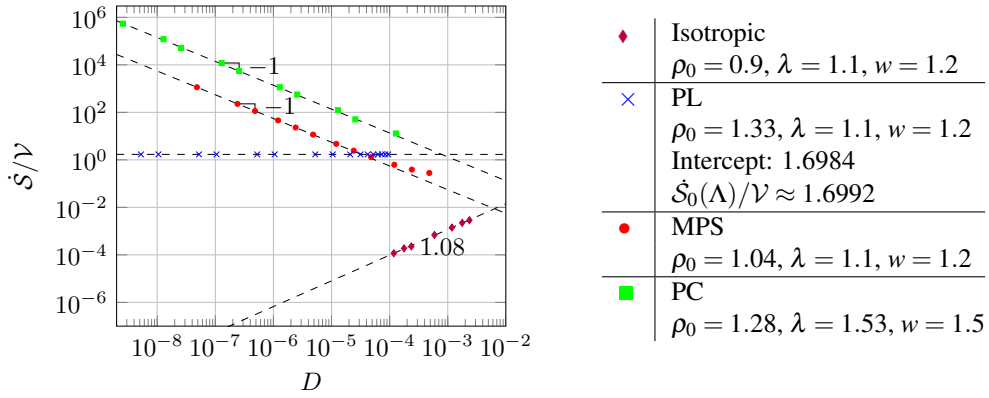


Fig. 6.1 Scaling of the EPR $\dot{\mathcal{S}}$ (normalised by volume $\mathcal{V}$) with the noise coefficient D in the isotropic, PL, MPS and PC regimes. Dashed lines (- - - -) represent the best linear fit to the data from simulations (marked by ♦, ×, ●, ■), with the associated intercept (best estimate of $\lim_{D\to 0}\dot{\mathcal{S}}(D)$ from simulation data) reported in the legend for the PL. The intercept is compared with the numerically evaluated analytical result in eq. (6.47) for $\dot{\mathcal{S}}_0(\Lambda)/\mathcal{V}$, where $\Lambda = 2\pi N/L$, $L = 14\pi$ and $N = 96$.

where we with slight abuse of notation denote by $\mathcal{V} = \mathrm{vol}(\mathcal{V})$. We thus obtain the set of equations

$$\partial_t \rho_{1,\boldsymbol{q}} = -iw\boldsymbol{q}\cdot\boldsymbol{m}_{1,\boldsymbol{q}}, \tag{6.37}$$

$$\partial_t \boldsymbol{m}_{1,\boldsymbol{q}} = -\Gamma(q)\boldsymbol{m}_{1,\boldsymbol{q}} - iw_1\boldsymbol{q}\rho_{1,\boldsymbol{q}} + \boldsymbol{\eta}_{1,\boldsymbol{q}}, \tag{6.38}$$

where we have defined the damping coefficient $\Gamma(q) = a_0 + q^2$ and the noise term $\boldsymbol{\eta}_{1,\boldsymbol{q}}$ is mean zero, Gaussian and white with covariance

$$\langle \eta_{1\alpha,\boldsymbol{q}}(t)\eta^*_{1\beta,\boldsymbol{q}}(t')\rangle = 2\mathcal{V}^{-1}\delta_{\alpha\beta}\delta_{\boldsymbol{q},\boldsymbol{q}'}\delta(t-t'). \tag{6.39}$$

Using the mapping

$$\rho_{1,\boldsymbol{q}} = x_{\boldsymbol{q}} + iy_{\boldsymbol{q}}, \tag{6.40}$$

$$\partial_t \rho_{1,\boldsymbol{q}} = v_{x,\boldsymbol{q}} + iv_{y,\boldsymbol{q}}, \tag{6.41}$$

and setting $\boldsymbol{X}_{\boldsymbol{q}} = (x_{\boldsymbol{q}}, y_{\boldsymbol{q}})$, $\boldsymbol{V}_{\boldsymbol{q}} = (v_{x,\boldsymbol{q}}, v_{y,\boldsymbol{q}})$ we immediately see that these follow standard equilibrium Langevin equations for an underdamped particle in a harmonic potential [22],

$$\dot{\boldsymbol{X}}_{\boldsymbol{q}} = \boldsymbol{V}_{\boldsymbol{q}}, \tag{6.42}$$

$$\dot{\boldsymbol{V}}_{\boldsymbol{q}} = -\Gamma(q)\boldsymbol{V}_{\boldsymbol{q}} - \nabla_{X_{\boldsymbol{q}}} U(\boldsymbol{q}) + \sqrt{2\Gamma(q)T(q)}\boldsymbol{\zeta}_{\boldsymbol{q}}, \tag{6.43}$$

where the potential $U = ww_1 q^2 X_{\boldsymbol{q}}^2/2$. The final degrees of freedom in eq. (6.37), eq. (6.38) are captured by the transverse component $\boldsymbol{V}_{T,\boldsymbol{q}} = (v_{Tx,\boldsymbol{q}}, v_{Ty,\boldsymbol{q}})$ of $\boldsymbol{m}_{1,\boldsymbol{q}}$ with respect to $\boldsymbol{q}$, i.e.

$$v_{Tx,\boldsymbol{q}} + iv_{Ty,\boldsymbol{q}} = -iw\boldsymbol{q}_\perp \cdot \boldsymbol{m}_{1,\boldsymbol{q}}, \tag{6.44}$$

and $\boldsymbol{q}_\perp$ is perpendicular to $\boldsymbol{q}$ with $q_\perp = q$. Again this follows an equilibrium Langevin equation,

$$\dot{\boldsymbol{V}}_{T,\boldsymbol{q}} = -\Gamma(q)\boldsymbol{V}_{T,\boldsymbol{q}} + \sqrt{2\Gamma(q)T(q)}\boldsymbol{\zeta}_{T,\boldsymbol{q}}. \tag{6.45}$$

In eq. (6.43), eq. (6.45) the noise terms $\boldsymbol{\zeta}_{\boldsymbol{q}}$, $\boldsymbol{\zeta}_{T,\boldsymbol{q}}$ are mean zero unit Gaussian white noises, and interestingly the effective temperature T is defined by

$$T(q) = \frac{1}{2\mathcal{V}} \frac{w^2 q^2}{a_0 + q^2}. \tag{6.46}$$

Since the modes $\boldsymbol{V}_{\boldsymbol{q}}$, $\boldsymbol{V}_{T,\boldsymbol{q}}$ are independent for all $\boldsymbol{q}$, the dependence of the effective temperature T on q does not lead to any current in phase space. At higher order in D, however, modes $\rho_{1,\boldsymbol{q}}$ and $\boldsymbol{m}_{1,\boldsymbol{q}}$ are coupled at different wave vectors $\boldsymbol{q}$ via the nonlinear terms in the equation for the magnetisation density in eq. (6.3). In particular, these terms couple heat baths at different temperatures $T(q)$, driving the dynamics at the next order away from equilibrium.

Linear theory also allows us to make quantitative predictions about $\dot{\mathcal{S}}_0$ from eq. (6.31) in the PL phase. Indeed, transforming this to Fourier space we obtain

$$\dot{\mathcal{S}}_0(\Lambda)/\mathcal{V} = \sum_{q\leq\Lambda} \left\langle \left(\boldsymbol{u}_{\boldsymbol{q}}^{\mathrm{pl}}\right)^\dagger \dot{\sigma}_{\boldsymbol{q}}^{\mathrm{pl}} \boldsymbol{u}_{\boldsymbol{q}}^{\mathrm{pl}} \right\rangle = \sum_{q\leq\Lambda} \mathrm{Tr}\left(\dot{\sigma}_{\boldsymbol{q}}^{\mathrm{pl}} C^{\mathrm{pl}}(q)\right), \tag{6.47}$$

where the Hermitian matrix $\dot{\sigma}_{\boldsymbol{q}}^{\mathrm{pl}}$ is given by

$$\dot{\sigma}_{\boldsymbol{q}}^{\mathrm{pl}} = \frac{im_0}{2} \begin{pmatrix} 0 & m_0(2w_1 - \kappa + \lambda)q_\parallel & 0 \\ m_0(\kappa - 2w_1 - \lambda)q_\parallel & 0 & (w - 2(m_0^2 + q^2)\kappa)q_\perp \\ 0 & (2(m_0^2 + q^2)\kappa - w)q_\perp & 0 \end{pmatrix}. \tag{6.48}$$

In addition, in eq. (6.47) we have defined the vector

$$\boldsymbol{u}_{\boldsymbol{q}}^{\mathrm{pl}} = \left(\rho_{1,\boldsymbol{q}}, P_{\parallel,\boldsymbol{q}}, P_{\perp,\boldsymbol{q}}\right)^T \tag{6.49}$$

of Fourier modes, as well as the matrix $C^{\mathrm{pl}} \equiv (C_{ij}^{\mathrm{pl}})$ of equal-time correlators by

$$C_{ij}^{\mathrm{pl}}(q)\delta_{\boldsymbol{q},\boldsymbol{q}'} = \left\langle u_{i,\boldsymbol{q}}^{\mathrm{pl}} \left(u_{j,\boldsymbol{q}'}^{\mathrm{pl}}\right)^* \right\rangle. \tag{6.50}$$

The sum in eq. (6.47) runs over modes with wave numbers smaller than the ultraviolet cutoff Λ, which is introduced since the sum is divergent with $\Lambda \to \infty$. This type of divergence is sometimes seen in field theories of this kind, since they are often derived based on the assumption that they are only valid down to a certain length scale. In appendix D we show from this that $\dot{\mathcal{S}}_0(\Lambda)$ as predicted by the linear theory diverges in the ultraviolet as Λ^2. Although the closed form expressions for the correlators entering in eq. (6.47) are too algebraically involved to report explicitly, they may be calculated straightforwardly by numerical methods. By doing this, we may calculate the corresponding sum in eq. (6.47) and quantitatively compare the results with measurements of $\dot{\mathcal{S}}$ from simulations. In fig. 6.1 we plot the results obtained from simulations, which show good agreement with the analytical results. Plot a) in fig. 6.1 demonstrates that the EPR is $\mathcal{O}(D)$ in the isotropic phase, as well as the predicted $\mathcal{O}(D^0)$ scaling in the PL phase. In particular, both remain finite at fixed Λ as $D \to 0$, with $\dot{\mathcal{S}} \to 0$ in the isotropic phase.

6.3 Nonlinear ground-states

Previous studies have investigated the nonlinear solutions to eqs. (6.1) and (6.3) at $D = 0$, and particularly interesting to our present context are the seminal contributions by Solon *et al.* [99, 100] on the structure of the banded profiles. These are effectively one-dimensional traveling wave solutions that are invariant along the direction perpendicular to the motion. We thus write $\boldsymbol{m}_0 = (m_0, 0)$ without loss of generality, and look for solutions of the form

$$\rho_0(\boldsymbol{x}, t) \equiv \tilde{\rho}_0(x - ct), \tag{6.51}$$

$$m_0(\boldsymbol{x}, t) \equiv \tilde{m}_0(x - ct). \tag{6.52}$$

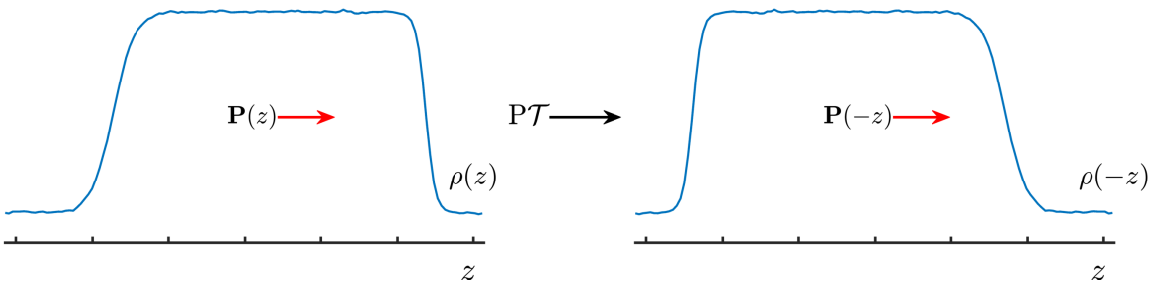


Fig. 6.2 Illustration showing a cross section of a banded profile traveling in the positive x direction (left) and its image under the map PT (right).

Direct substitution then allows us to deduce a set of equations for $\tilde{\rho}_0$ and $\tilde{m}_0$ in terms of the variable $z = x - ct$ given explicitly by

$$\tilde{\rho}_0 = \rho_g + \frac{w}{c}\tilde{m}_0, \tag{6.53}$$

$$\tilde{m}_0'' = -\left(c - \frac{ww_1}{c} - \lambda\tilde{m}_0\right)\tilde{m}_0' - \left(\rho_g - 1 + \frac{w}{c}\tilde{m}_0 - \tilde{m}_0^2\right)\tilde{m}_0, \tag{6.54}$$

where primes denote differentiation with respect to z. Equation (6.54) can be mapped onto a Newton problem for a particle in a potential under the influence of a nonlinear drag, and all stable orbits in the $(\tilde{m}_0, \tilde{m}_0')$ plane with $\tilde{m}_0 \geq 0$ can be uniquely identified with a pair (c, ρ_g). In terms of the stochastic equations in eqs. (6.1) and (6.3), it is assumed that the noise selects the stable steady-state profile (of which there are infinitely many [99, 100]).

Importantly, these solutions to eq. (6.52) break both T and PT-symmetry. Thus, we expect the MPS steady-state to have $\chi = -1$. Indeed, using the traveling wave ansatz in eq. (6.51) and eq. (6.52) we deduce two expressions for $\dot{\mathcal{S}}_{-1} = \lim_{D\to 0} D\dot{\mathcal{S}}(D)$, that are

$$\dot{\mathcal{S}}_{-1}/\mathcal{V} = \frac{1}{2L}\int_{-L}^{L} \mathrm{d}z \left((1-\tilde{\rho}_0)\tilde{m}_0 + \tilde{m}_0^3 - \tilde{m}_0''\right)^2 \tag{6.55}$$

$$= -\frac{\lambda}{2L}\int_{-L}^{L} \mathrm{d}z\, (\tilde{m}_0')^3. \tag{6.56}$$

The first of these is most straightforwardly derived from eq. (D.3) in appendix D by using the ODE eq. (6.54) for $\tilde{m}_0$, and verifies that $\dot{\mathcal{S}}_{-1} \geq 0$ as must be the case. The latter, i.e. eq. (6.56), follows immediately from eq. (6.14) after integrating out total derivatives. Now, from the final expression in eq. (6.56) one observes that $\dot{\mathcal{S}}_{-1}$ vanishes identically for even distributions, i.e. those that satisfy $\tilde{m}_0(z_0 + z) = \tilde{m}_0(z_0 - z)$ for some z_0, which is exactly the PT-symmetry in eq. (6.20). However, the traveling wave profiles are clearly asymmetric with a steeper front than tail, leading in general to the observed $\dot{\mathcal{S}}_{-1} \neq 0$. In fig. 6.2 we illustrate this, and in particular how the banded profile transforms under PT. Finally, a

sanity check also verifies that both expressions in eqs. (6.55) and (6.56) are invariant under P alone ($\tilde{m}_0(z) \to -\tilde{m}_0(-z)$) as they should be since $\dot{\mathcal{S}}$ is, as remarked previously, oblivious to whether the wave is moving left or right (in eq. (6.54) this must be complemented by $c \to -c$).

We include in fig. 6.1 the scaling of the EPR $\dot{\mathcal{S}}$ in both the MPS and PC regimes. As shown, both are strongly non-equilibrium within our classification scheme, with $\dot{\mathcal{S}} \sim D^{-1}$. Although we do not possess explicit PC solutions to the dynamics at $D = 0$, it is straightforward to argue qualitatively that this is what one should expect due to the highly inhomogeneous nature of the clusters. For future work, we aim to investigate this in more detail.

6.4 Entropy production rate of the DFM

Above we found that for the HVM, pathwise violation of detailed balance at ground-state level is the direct result of a PT-symmetry breaking by asymmetric deterministic $D = 0$ profiles. Here we aim to show that in order to pick up a contribution that diverges when $D \to 0$ also from the density current in the homogeneous PL phase, this current needs to be tracked explicitly when computing the EPR. However, in order for this contribution to be finite from the outset (at finite D), we must change the model so as to allow for independent density current fluctuations. We then find as expected that the EPR does indeed diverge as $\mathcal{O}(1/D)$. Moreover, we recover an explicit expression for the pathwise EPR of a constant homogeneous PL state in this case.

In the following, we add a diffusive contribution and noise to the constitutive equation for the density current $\boldsymbol{J}$ in eq. (6.2). Specifically, we consider $\boldsymbol{J} = \boldsymbol{J}_d + \boldsymbol{\xi}$ as in [1], where the noise $\boldsymbol{\xi}$ is mean zero, Gaussian and white with covariance

$$\langle \xi_\alpha(\boldsymbol{x},t)\xi_\beta(\boldsymbol{x}',t')\rangle = 2D_\rho \delta_{\alpha\beta}\delta(\boldsymbol{x}-\boldsymbol{x}')\delta(t-t'). \tag{6.57}$$

Furthermore, we take the deterministic part $\boldsymbol{J}_d$ of the current $\boldsymbol{J}$ to be of the form

$$\boldsymbol{J}_d = w\boldsymbol{m} - \gamma^{-1}\nabla\mu, \tag{6.58}$$

where γ is a constant friction coefficient. Here, μ serves an analogous purpose to the chemical potential known from equilibrium diffusive systems. However, since TRS is broken in this model, there is a priori no reason that it should be the functional variation of some $F[\rho,\boldsymbol{m}]$.

Notwithstanding, we will for simplicity ignore this issue and assume that we may write

$$\mu = \frac{\delta F}{\delta \rho}, \tag{6.59}$$

with the same functional F as that which appears in the equation for $\boldsymbol{m}$. We only make minor changes to F for stability purposes by modifying the local free-energy $f(\rho, \boldsymbol{m})$ to include a quadratic term in ρ, so that now

$$f(\rho, \boldsymbol{m}) = \frac{a_\rho}{2}\rho^2 + \frac{1}{2}(1-\rho)m^2 + \frac{1}{4}m^4. \tag{6.60}$$

In addition we add a square gradient contribution, giving us

$$F[\rho, \boldsymbol{m}] = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \left(f(\rho, \boldsymbol{m}) + \frac{\nu_\rho}{2}|\nabla\rho|^2 + \frac{1}{2}(\nabla_\alpha m_\beta)^2 + \boldsymbol{m}\cdot\nabla\Phi(\rho, \boldsymbol{m}) \right). \tag{6.61}$$

Observe, however, that these new terms do not change the equation for $\boldsymbol{m}$ in eq. (6.3), since they both drop out when considering the functional variation of F with respect to $\boldsymbol{m}$. With this choice, we have that

$$\mu = a_\rho\rho - \nu_\rho\nabla^2\rho - \frac{1}{2}m^2 - w_1\nabla\cdot\boldsymbol{m}. \tag{6.62}$$

Again, eq. (6.58) is motivated by coarse graining, and the diffusive contribution arises for example in cases where interactions such as steric repulsion are included in the microscopic model [1]. Notably, there is a kind of paradigm shift when breaking the local linear relation $\boldsymbol{J} \propto \boldsymbol{m}$, which implies that $\bar{\boldsymbol{m}} = \boldsymbol{m}/\rho$ should no longer be considered the local average direction of the velocity of particles. Physically, this reflects a situation on the microscopic scale where the bare self-propulsion may be thwarted by e.g. repulsive forces, so that mass currents may move against the local polar order. More importantly in the context of entropy production, this means that a trajectory of the system in which $\boldsymbol{J}$ and $\boldsymbol{m}$ do *not* point in the same direction is realisable in the forward time ensemble, since fluctuations alone can now reverse $\boldsymbol{J}$ at fixed $\boldsymbol{m}$ even if highly unlikely.

Numerical integration of the dynamics with $\boldsymbol{J} = \boldsymbol{J}_d + \boldsymbol{\xi}$, hereafter referred to as the **diffusive flocking model** (DFM), allows us to investigate the resulting phase diagram as in section 4.4 of chapter 4. On the other hand, achieving analytical progress to a comparable extent as with the HVM is more difficult. Notably, however, from a linear analysis we do in fact find a finite wavelength instability in the region where $\rho_0 < 1$, in which the coarsening dynamics develop a polar crystalline structure as illustrated in fig. 6.3. Our analysis also provides us with the isotropic-to-crystal phase-boundary, and we find that in the

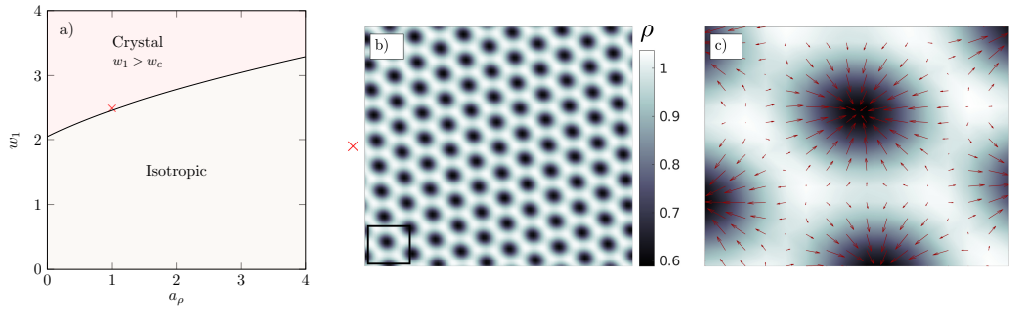


Fig. 6.3 a) Phase diagram of the DFM at fixed $\rho_0 = 0.9$, $\nu_\rho = 1$, $\gamma = 0.5$ and $w_1 = w/2$. When the condition in eq. (6.63) is met, there is a finite range $q \in [q_-, q_+]$ of modes that are unstable to perturbations away from the isotropic state even when $\rho_0 < 1$. The resulting steady-state is a type of polar crystal in which a hexagonal lattice formed by high-density ridges enclose low density valleys, as illustrated in b). Figure c) shows an enlarged part of the plot in b), indicated by a black square, including also the local magnetisation density plotted with red arrows (→). The simulation parameters used in b) are given by $w_1 = 2.5$ and $a_\rho = 1$, corresponding to the data point (×) in a), and in addition $\lambda = \kappa = 1.1$, $D = 10^{-4}$ and $L = 7\pi$.

case $w_1 = w/2$ the system is unstable to perturbations when

$$w_1^2 > w_c^2 \equiv a_\rho + \nu_\rho(1-\rho_0) + 4\gamma\nu_\rho + 2\sqrt{\nu_\rho(1-\rho_0+2\gamma)(a_\rho+2\gamma\nu_\rho)}. \tag{6.63}$$

More specifically, when the inequality in eq. (6.63) holds there is a finite range $q \in [q_-, q_+]$ of modes that are unstable, where the exact expressions for $q_\pm$ are provided in appendix C. In contrast, we do not observe significant changes to the phase diagram in the region where $\rho_0 > 1$. We explain this behaviour by observing that when $\rho_0 < 1$, and the local polar order is weak, the diffusive dynamics is significant while it is overpowered by advective transport when the polar order is strong [1]. At $D = 0$ the state is stationary and has both $\partial_t \rho = 0$ and $|\partial_t \boldsymbol{m}| = 0$. In fact, from simulations we observe that the stronger condition $|\langle \boldsymbol{J}_d \rangle| = 0$ is met, meaning that at $D = 0$ we expect

$$w\boldsymbol{m} = \gamma^{-1}\nabla\mu. \tag{6.64}$$

From simulations we observe that the local polar order is directed such that it points *in* towards low density, as illustrated in fig. 6.3. Equation (6.64) then tells us that the advective transport induced by $\boldsymbol{m}$ is compensated by a reversed 'diffusion' running *up* gradients in ρ.

In the following we also restrict to the case where $D_\rho = D/\gamma$ for simplicity [1], in which case the Freidlin-Wentzell action for the DFM takes the form

$$\mathcal{A}_{\mathrm{DF}}[\rho,\boldsymbol{m}] = \frac{1}{4}\int_{-\tau}^{\tau}\mathrm{d}t\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\left[\gamma\left|\nabla^{-1}\left(\partial_t\rho+\nabla\cdot\boldsymbol{J}_d\right)\right|^2+\left|\partial_t\boldsymbol{m}+\lambda\left(\boldsymbol{m}\cdot\nabla\right)\boldsymbol{m}+\frac{\delta F}{\delta\boldsymbol{m}}\right|^2\right], \tag{6.65}$$

and the path transition density $\mathcal{P}_{\mathrm{DF}}[\rho,\boldsymbol{m}]$ is constructed as before by setting $\mathcal{P}_{\mathrm{DF}} \propto \exp(-\mathcal{A}_{\mathrm{DF}}/D)$. Note that in eq. (6.65), we have defined the inverse gradient operator $\nabla^{-1} = \nabla^{-2}\nabla$, i.e. with gauge choice $|\nabla\times\nabla^{-1}h(\boldsymbol{x})| = 0$ [12]. Crucially, with the added density current fluctuations, we are now free to define time-reversal without flipping the local magnetisation as in eq. (6.10). Specifically, we let

$$\mathrm{T}^{\pm}:\begin{cases}\rho(\boldsymbol{x},t) \mapsto \rho(\boldsymbol{x},-t)\\ \boldsymbol{m}(\boldsymbol{x},t) \mapsto \pm\boldsymbol{m}(\boldsymbol{x},-t)\end{cases} \tag{6.66}$$

and observe that both compositions $\mathcal{A}^{\pm}_{\mathrm{DF}}[\rho,\boldsymbol{m}] = \mathcal{A}_{\mathrm{DF}}[\mathrm{T}^{\pm}\rho,\mathrm{T}^{\pm}\boldsymbol{m}]$ are now well defined on the full space of trajectories. As before, this means that when we define the two time-reversed ensembles to $\mathcal{P}_{\mathrm{DF}}$ by setting $\overleftarrow{\mathcal{P}}^{\pm}_{\mathrm{DF}}[\rho,\boldsymbol{m}] = \mathcal{P}_{\mathrm{DF}}[\mathrm{T}^{\pm}\rho,\mathrm{T}^{\pm}\boldsymbol{m}]$, all trajectories that are observable under $\mathcal{P}_{\mathrm{DF}}$ are also observable under $\overleftarrow{\mathcal{P}}^{\pm}_{\mathrm{DF}}$. Now, comparing the time-forward ensemble with each of these gives rise to two different definitions of the entropy production rate, given by

$$\dot{\mathcal{S}}^{\pm} \equiv \lim_{\tau\to\infty}\frac{1}{2\tau}\log\frac{\mathcal{P}_{\mathrm{DF}}[\rho,\boldsymbol{m}]}{\overleftarrow{\mathcal{P}}^{\pm}_{\mathrm{DF}}[\rho,\boldsymbol{m}]}. \tag{6.67}$$

In section 6.4 we will attempt to understand how this choice of polar signature may alter the scaling of the EPR at low noise [34, 35].

Analogously with our treatment in the previous section, we observe that when $\boldsymbol{m}$ is odd under time-reversal, the functional F splits into even and odd pieces F^S and F^A respectively. However, in our current setting this has further consequences as well, since it also implies that we should not consider μ a chemical potential like quantity either. Indeed, we see that μ splits into contributions

$$\mu = \mu^S + \frac{\delta F^A}{\delta\rho}. \tag{6.68}$$

Furthermore, the deterministic part of the current, $\boldsymbol{J}_d$, also splits into an $\boldsymbol{m}$-like odd piece under time-reversal and a $\nabla\mu^S$-like even piece. That is, we write $\boldsymbol{J}_d = \boldsymbol{J}^S_d + \boldsymbol{J}^A_d$, where we

define

$$\boldsymbol{J}_d^S = -\gamma^{-1}\nabla\mu^S, \tag{6.69}$$

$$\boldsymbol{J}_d^A = w\boldsymbol{m} - \gamma^{-1}\nabla\frac{\delta F^A}{\delta\rho}. \tag{6.70}$$

Consequently, since F does not remain invariant under time-reversal, and therefore neither μ nor $\boldsymbol{J}_d$ either, it could not feature in an equilibrium theory and violates TRS.

With these definitions, we may again deduce explicit integral expressions for the EPRs $\dot{\mathcal{S}}^{\pm}$ by using eqs. (6.65) to (6.67) and the definitions of $\mathcal{P}_{\mathrm{DF}}$, $\overleftarrow{\mathcal{P}}^{\pm}_{\mathrm{DF}}$, and we refer to appendix D for the details. There we show that

$$\dot{\mathcal{S}}^{+} = D^{-1}\int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\left\langle \gamma w^2|K\boldsymbol{m}|^2 - w\boldsymbol{m}\cdot\nabla\mu + \lambda\left[(\boldsymbol{m}\cdot\nabla)\boldsymbol{m}\right]\cdot\left(\lambda\left(\boldsymbol{m}\cdot\nabla\right)\boldsymbol{m} + \frac{\delta F}{\delta\boldsymbol{m}}\right)\right\rangle \tag{6.71}$$

and

$$\dot{\mathcal{S}}^{-} = D^{-1}\int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\left\langle \boldsymbol{J}_d^A\cdot\nabla\mu^S - \left(\lambda(\boldsymbol{m}\cdot\nabla)\boldsymbol{m} + \frac{\delta F^A}{\delta\boldsymbol{m}}\right)\cdot\frac{\delta F^S}{\delta\boldsymbol{m}}\right\rangle \tag{6.72}$$

where K is a matrix operator with entries $K_{\alpha\beta} = \nabla_\alpha^{-1}\nabla_\beta = \nabla^{-2}\nabla_\alpha\nabla_\beta$. Note that in eq. (6.71) we have performed an average over noise histories in order to obtain the given expression, which explains why the symmetry $\dot{\mathcal{S}}^{+}[\mathrm{T}^{+}\rho, \mathrm{T}^{+}\boldsymbol{m}] = -\dot{\mathcal{S}}^{+}$ does not seem to hold pathwise any longer. However, it is recovered when writing the expression out as in eq. (D.14) in appendix D. As in sections 6.2 and 6.3, we proceed to analyse eqs. (6.71) and (6.72) in the low D limit both analytically and numerically. We also carry over the definitions we employed there, in particular defining the exponents $\chi^{\pm}$ via the asymptotic scaling relation $\dot{\mathcal{S}}^{\pm}(D) \sim D^{\chi^{\pm}}$ for $D \ll 1$. As we will see, similar considerations to those made before carry over in a straightforward manner, allowing us to predict the correct scaling in all cases.

6.5 Effect of tracking the local density current

Continuing as in sections 6.2 and 6.3, we look for ground-state trajectories $(\rho_0, \boldsymbol{m}_0)$ that satisfy the pathwise equilibrium condition

$$\dot{\mathcal{S}}^{\pm}[\rho_0, \boldsymbol{m}_0] = 0, \tag{6.73}$$

in order to determine when we should expect $\chi^{\pm} = -1$. Again, it is clear that these include all states that are either $\mathrm{T}^{\pm}$ or $\mathrm{PT}^{\pm}$-symmetric, and so the situation remains unchanged when choosing T^{-}. Indeed, in this case the isotropic $m_0 = 0$ state satisfies both symmetries, while

the PL $m_0 = \sqrt{\rho_0 - 1}$ state is PT^--symmetric only. However, we should also expect a similar situation when choosing the T^+-protocol for time-reversal; since $\boldsymbol{m}_0$ does not flip sign upon time-reversal, any constant trajectory satisfies T^+ alone. Thus we expect $\chi^\pm > -1$ for both the isotropic gas and PL, meaning that there is no clear distinction between the two protocols for the homogeneous phases at ground-state level.

On the other hand, the situation changes quite drastically once the density current dynamics are tracked explicitly. In particular, in doing so, we expect that TRS violation at the $D = 0$ level should become visible from a misalignment of the density current and magnetisation density. To see this, we promote $\boldsymbol{J}$ to a dynamical variable and consider the Freidlin-Wentzell action at this level. That is, we define

$$\mathcal{A}^J_{\mathrm{DF}}[\rho, \boldsymbol{J}, \boldsymbol{m}] = \frac{1}{4}\int_{-\tau}^{\tau} \mathrm{d}t \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \left[\gamma |\boldsymbol{J} - \boldsymbol{J}_d|^2 + \left| \partial_t \boldsymbol{m} + \lambda \left(\boldsymbol{m}\cdot\nabla\right)\boldsymbol{m} + \frac{\delta F}{\delta \boldsymbol{m}} \right|^2 \right], \tag{6.74}$$

if $\partial_t \rho + \nabla \cdot \boldsymbol{J} = 0$ and $\mathcal{A}^J_{\mathrm{DF}} = \infty$ otherwise. Importantly, $\boldsymbol{J}$ now takes the role that $w\boldsymbol{m}$ had previously in sections 6.2 and 6.3, in the sense that it must be odd on time-reversal. Again, this is due to the constraint imposed by the continuity equation which limits the space of observable trajectories under the action eq. (6.74). Time-reversal is thus generalised accordingly by setting

$$\mathrm{T}^\pm_J : \begin{cases} \rho(\boldsymbol{x}, t) \mapsto \rho(\boldsymbol{x}, -t), \\ \boldsymbol{J}(\boldsymbol{x}, t) \mapsto -\boldsymbol{J}(\boldsymbol{x}, -t), \\ \boldsymbol{m}(\boldsymbol{x}, t) \mapsto \pm\boldsymbol{m}(\boldsymbol{x}, -t). \end{cases} \tag{6.75}$$

Pathwise there is now a clear distinction between the two protocols $\mathrm{T}^\pm_J$. Indeed, when $(\rho_0, \boldsymbol{J}_0, \boldsymbol{m}_0)$ is a constant trajectory with both $J_0 > 0$ and $m_0 > 0$, the protocol T^+_J introduces a discrepancy between $\boldsymbol{J}_0$ and $\boldsymbol{m}_0$ that cannot be transformed away by parity. On the other hand, since both $\boldsymbol{J}_0$ and $\boldsymbol{m}_0$ transform the same way under T^-_J, the trajectory $(\rho_0, \boldsymbol{J}_0, \boldsymbol{m}_0)$ remains invariant under PT^-_J as before. This fact is reflected by the EPR computed at the level of the action in eq. (6.74). To see this, we set $\mathcal{P}^J_{\mathrm{DF}} \propto \exp\left(-\mathcal{A}^J_{\mathrm{DF}}/D\right)$ and $\overleftarrow{\mathcal{P}}^{J,\pm}_{\mathrm{DF}}[\rho, \boldsymbol{m}] = \mathcal{P}^J_{\mathrm{DF}}[\mathrm{T}^\pm_J \rho, \mathrm{T}^\pm_J \boldsymbol{m}]$ as before, and define $\dot{\mathcal{S}}^\pm_J$ via the usual definition as in eq. (6.67). In appendix D we show that this computation results in the expression

$$\dot{\mathcal{S}}^+_J = D^{-1} \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \left\langle \gamma w^2 m^2 - w\boldsymbol{m}\cdot\nabla\mu - \lambda\left[\left(\boldsymbol{m}\cdot\nabla\right)\boldsymbol{m}\right]\cdot\left(\lambda\left(\boldsymbol{m}\cdot\nabla\right)\boldsymbol{m} + \frac{\delta F}{\delta \boldsymbol{m}}\right)\right\rangle \tag{6.76}$$

for $\dot{\mathcal{S}}^+_J$, which differs from eq. (6.71) by omission of the factor K inside the first term, while in fact $\dot{\mathcal{S}}^-_J = \dot{\mathcal{S}}^-$ remains unchanged from eq. (6.72).

It is instructive to investigate the difference between the two expressions in eq. (6.71) for $\dot{\mathcal{S}}^+$ and eq. (6.76) for $\dot{\mathcal{S}}_J^+$, and in particular to show that indeed the operator $K \neq \mathrm{Id}$. To this end, we introduce the potential φ as the solution to the Poisson's equation

$$\nabla^2 \varphi = \nabla \cdot \boldsymbol{m}, \tag{6.77}$$

which is unique up to an additive constant (assuming periodic boundaries on $\mathcal{V}$). Defining $\boldsymbol{m}_T \equiv \boldsymbol{m} - \nabla\varphi$ it follows by construction that

$$\boldsymbol{m} = \nabla\varphi + \boldsymbol{m}_T, \tag{6.78}$$

where $\boldsymbol{m}_T$ is solenoidal, i.e. $\nabla \cdot \boldsymbol{m}_T = 0$. Note that in the presence of periodic boundary conditions, this construction is a slight generalisation of the standard Helmholtz decomposition since $\boldsymbol{m}_T$ is not necessarily the curl of a vector potential. Rather, it is a constant $\bar{\boldsymbol{m}}$ plus a curl, where $\bar{\boldsymbol{m}}$ is given by the value of $\boldsymbol{m}$ at zero wave-number, i.e. $\bar{\boldsymbol{m}} = \mathcal{V}^{-1}\int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \boldsymbol{m}$. Still, the decomposition in eq. (6.78) does retain orthogonality in $L^2(\mathcal{V})$, meaning that

$$\int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, m^2 = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, |\nabla\varphi|^2 + \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, m_T^2. \tag{6.79}$$

Thus, observing that by definition we have $K\boldsymbol{m} = \nabla\varphi$, it follows that the difference between $\dot{\mathcal{S}}^+$ and $\dot{\mathcal{S}}_J^+$ is simply

$$\dot{\mathcal{S}}_J^+ - \dot{\mathcal{S}}^+ = \gamma w^2 D^{-1} \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \langle m_T^2 \rangle\,. \tag{6.80}$$

This is consistent with our observation that the PL breaks T_J^+ at ground-state level as remarked above, and in particular it follows that we have

$$\lim_{D\to 0} D\dot{\mathcal{S}}_J^+ / \mathcal{V} = \gamma w^2 m_0^2 \tag{6.81}$$

for the constant homogeneous ground-states. This mechanism by which TRS is broken at $D = 0$ due to $\boldsymbol{J}$ and $\boldsymbol{m}$ having different polar signatures under time-reversal has an analogue for a single active particle on the microscopic level. To see this we make the identifications $\dot{\boldsymbol{x}}\,\delta(\boldsymbol{x}(t) - \boldsymbol{x}) \to \boldsymbol{J}$ and $\boldsymbol{e}\,\delta(\boldsymbol{x}(t) - \boldsymbol{x}) \to \boldsymbol{m}$, where $\boldsymbol{x}(t)$ is the position of the active particle and $\boldsymbol{e}(t)$ denotes its polar orientation and direction of self-propulsion. Similarly to the current $\boldsymbol{J}$, the particle velocity $\dot{\boldsymbol{x}}$ must change sign under time-reversal. The polar orientation $\boldsymbol{e}$ need not, on the other hand, and thus generally leads to a bare entropy production associated with the motility (see section 2.6 of chapter 2).

Continuing as in sections 6.2 and 6.3, we include fluctuations by performing a systematic expansion of the equations of motion and the EPRs $\dot{\mathcal{S}}^{\pm}$ via the integral expressions in

eq. (6.71) and eq. (6.72) in small D. Again, we start by assuming that the ground-state $(\rho_0, \boldsymbol{m}_0)$ is constant and homogeneous. Thus, substituting the expansions in eq. (6.21) and eq. (6.22) into the continuity equation in eq. (6.1) with $\boldsymbol{J} = \boldsymbol{J}_d + \boldsymbol{\xi}$ we obtain to lowest nontrivial order

$$\partial_t \rho_1 = -\nabla \cdot \left(w\boldsymbol{m}_1 - \gamma^{-1} \nabla \frac{\delta F_L}{\delta \rho}[\rho_1, \boldsymbol{m}_1] + \boldsymbol{\xi}_1 \right), \tag{6.82}$$

where now with slight abuse of notation

$$F_L[\rho, \boldsymbol{m}] = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \left(f_L(\rho, \boldsymbol{m}) + \frac{\nu_\rho}{2} |\nabla \rho|^2 + \frac{1}{2} (\nabla_\alpha m_\beta)^2 + \boldsymbol{m} \cdot \nabla \Phi_L(\rho, \boldsymbol{m}) \right). \tag{6.83}$$

In addition, the local free energy f_L is given by

$$f_L(\rho, \boldsymbol{m}) = \frac{a_\rho}{2} \rho^2 + \frac{a_0}{2} m^2 - \rho \boldsymbol{m}_0 \cdot \boldsymbol{m} + (\boldsymbol{m}_0 \cdot \boldsymbol{m})^2, \tag{6.84}$$

while Φ_L remains the same as in eq. (6.27). The noise term $\boldsymbol{\xi}_1$ is a mean zero Gaussian white noise process with covariance

$$\langle \xi_{1\alpha}(\boldsymbol{x}, t) \xi_{1\beta}(\boldsymbol{x}', t') \rangle = 2\gamma^{-1} \delta_{\alpha\beta} \delta(\boldsymbol{x} - \boldsymbol{x}') \delta(t - t'), \tag{6.85}$$

so that the linearised equation eq. (6.82) is indeed independent of D. Note also that there is no change to the linearised equation for the magnetisation density since this is the same for both models under consideration.

Similarly, an expansion of the EPRs $\dot{\mathcal{S}}^{\pm}$ in small D allows us to write

$$\dot{\mathcal{S}}^{\pm}(D) = \dot{\mathcal{S}}^{\pm}_{-1} D^{-1} + \dot{\mathcal{S}}^{\pm}_0 + \dot{\mathcal{S}}^{\pm}_1 D + \mathcal{O}(D^2), \tag{6.86}$$

where $\dot{\mathcal{S}}^{\pm}_k = 0$ for $k < \chi^{\pm}$, and $\dot{\mathcal{S}}^{\pm}_{\chi^{\pm}} > 0$ as before. By explicitly computing this expansion one finds that

$$\begin{cases} \chi^{\pm} = 0, & \text{isotropic,} \\ \chi^{\pm} = 0, & \text{PL.} \end{cases} \tag{6.87}$$

In particular, since the linearised continuity equation eq. (6.82) of the DFM implies that the steady-state expectation $\langle \rho_1 \nabla \cdot \boldsymbol{m}_1 \rangle$ no longer vanishes identically as in eq. (6.33), we cannot any longer expect that $\chi^- > 0$ in the isotropic phase. Since $\chi^{\pm} = 0$ in the isotropic and PL phases, we classify both phases as being weakly non-equilibrium for the DFM. This means that the linearised dynamics of the DFM cannot be mapped onto an equilibrium dynamics for any choice of $\boldsymbol{m}_0$.

As in sections 6.2 and 6.3, we now present a more in-depth analysis of the leading order term in the expansion eq. (6.86) of $\dot{\mathcal{S}}^{\pm}$. Beginning with the isotropic phase where $m_0 = 0$, we find after a Fourier transform that $\dot{\mathcal{S}}_0^{\pm}$ may be written in bilinear form as in eq. (6.47):

$$\dot{\mathcal{S}}_0^{\pm}(\Lambda)/\mathcal{V} = \sum_{0<q\leq\Lambda} \left\langle \left(\boldsymbol{u}_{\boldsymbol{q}}^{\mathrm{iso}}\right)^{\dagger} \dot{\sigma}_{\boldsymbol{q}}^{\pm,\mathrm{iso}} \boldsymbol{u}_{\boldsymbol{q}}^{\mathrm{iso}} \right\rangle \tag{6.88}$$

$$= \sum_{0<q\leq\Lambda} \mathrm{Tr}\left(\dot{\sigma}_{\boldsymbol{q}}^{\pm,\mathrm{iso}} C^{\mathrm{iso}}(q)\right). \tag{6.89}$$

Here, the sum runs over wave vectors $\boldsymbol{q}$, and an explicit dependence in $\dot{\mathcal{S}}_0^{\pm}(\Lambda)$ on the ultraviolet cutoff Λ is introduced in order to study the limit in which it is taken to infinity. We denote the Fourier modes of ρ_1 and $\boldsymbol{m}_1$ by $\rho_{1,\boldsymbol{q}}$ and $\boldsymbol{m}_{1,\boldsymbol{q}}$ respectively, and have defined

$$\boldsymbol{u}_{\boldsymbol{q}}^{\mathrm{iso}} = \left(\rho_{1,\boldsymbol{q}}, m_{L,\boldsymbol{q}}, m_{T,\boldsymbol{q}}\right)^T, \tag{6.90}$$

where $m_{L,\boldsymbol{q}} = \hat{\boldsymbol{q}} \cdot \boldsymbol{m}_{1,\boldsymbol{q}}$ and $m_{T,\boldsymbol{q}} = \hat{\boldsymbol{q}}_{\perp,\boldsymbol{q}} \cdot \boldsymbol{m}_{1,\boldsymbol{q}}$ are respectively the longitudinal and transverse components of $\boldsymbol{m}_1$ with respect to $\hat{\boldsymbol{q}}$, and $\hat{\boldsymbol{q}}_{\perp}$ is perpendicular to $\boldsymbol{q}$ and of unit length. The equal-time steady-state correlation matrix $C^{\mathrm{iso}} \equiv (C_{ij}^{\mathrm{iso}})$ is then defined by

$$C_{ij}^{\mathrm{iso}}(q)\delta_{\boldsymbol{q},\boldsymbol{q}'} = \left\langle u_{i,\boldsymbol{q}}^{\mathrm{iso}} \left(u_{j,\boldsymbol{q}'}^{\mathrm{iso}}\right)^* \right\rangle. \tag{6.91}$$

In addition, the Hermitian matrices $\dot{\sigma}^{\pm,\mathrm{iso}}$ in eq. (6.88) are given by

$$\sigma_{\boldsymbol{q}}^{+,\mathrm{iso}} = \frac{1}{2}\begin{pmatrix} 0 & iwq\Gamma_\rho & 0 \\ -iwq\Gamma_\rho & 2\gamma w\tilde{w} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{6.92}$$

and

$$\sigma_{\boldsymbol{q}}^{-,\mathrm{iso}} = \frac{1}{2}\begin{pmatrix} 0 & iq\left(w_1\Gamma - \tilde{w}\Gamma_\rho\right) & 0 \\ -iq\left(w_1\Gamma - \tilde{w}\Gamma_\rho\right) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{6.93}$$

where we have defined $\tilde{w} = w(1 - w_1 q^2/\gamma w)$ as well as damping coefficients $\Gamma = 1 - \rho_0 + q^2$ and $\Gamma_\rho = a_\rho + \nu_\rho q^2$. In fact, we may explicitly compute C^{iso} from the linearised dynamics and we refer to appendix C for the details. By substituting the result of this calculation back into eq. (6.88), we obtain

$$\dot{\mathcal{S}}_0^{+}(\Lambda) = \sum_{0<q\leq\Lambda} \frac{\gamma w^2}{\Gamma + \gamma^{-1} q^2 \Gamma_\rho}, \tag{6.94}$$

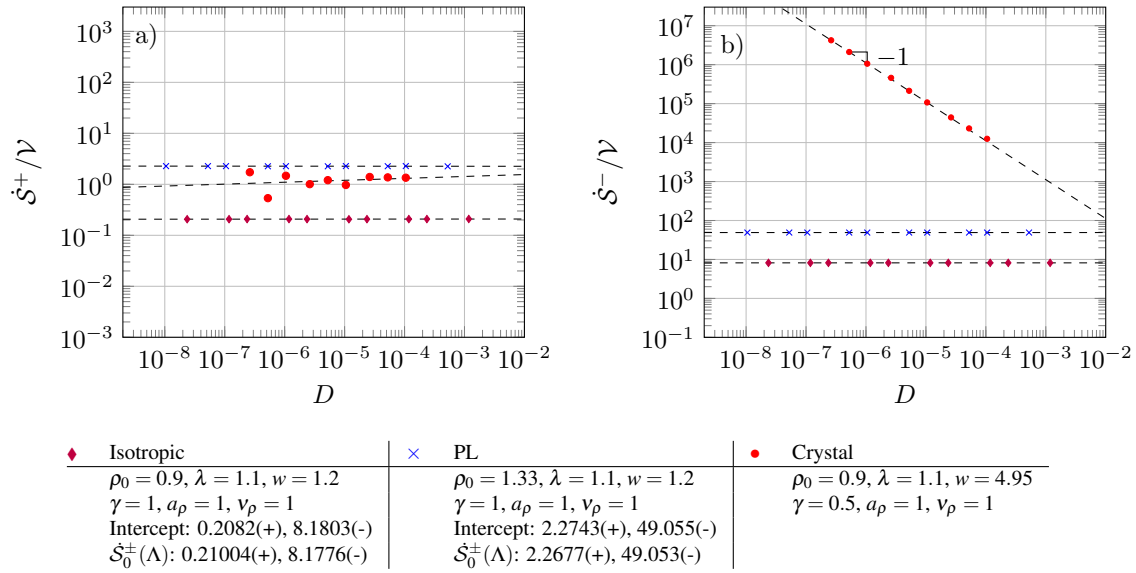


Fig. 6.4 Scaling of the EPRs $\dot{\mathcal{S}}^\pm$ (normalised by volume $\mathcal{V}$) with the noise coefficient D in the isotropic, PL, crystal phases. Dashed lines (- - - -) represent the best linear fit to the data from simulations (marked by ♦, ×, •), with the associated intercept (best estimate of $\lim_{D\to 0}\dot{\mathcal{S}}(D)$ from simulation data) reported in the legend for the isotropic and PL phases. The intercepts are compared with the numerically evaluated analytical results in eq. (6.88) for $\dot{\mathcal{S}}_0(\Lambda)/\mathcal{V}$, where $\Lambda = 2\pi N/L$, $L = 14\pi$ and $N = 96$.

and

$$\dot{\mathcal{S}}_0^-(\Lambda) = \sum_{q\leq\Lambda} \frac{\gamma^{-1}q^2(w_1\Gamma - \tilde{w}\Gamma_\rho)^2}{(\Gamma+\gamma^{-1}q^2\Gamma_\rho)(\tilde{w}w_1+\gamma^{-1}\Gamma_\rho\Gamma)}. \tag{6.95}$$

Interestingly, we see from direct power counting that $\dot{\mathcal{S}}_0^+(\Lambda)$ converges as $\Lambda\to\infty$, while $\dot{\mathcal{S}}_0^-(\Lambda)\sim\log\Lambda$. In fact, expressions eq. (6.94) and eq. (6.95) generalise trivially to dimensions $d\neq 2$, meaning that

$$\dot{\mathcal{S}}_0^+(\Lambda) \sim \begin{cases} 1, & d<4 \\ \Lambda^{d-4}, & d\geq 4 \end{cases} \tag{6.96}$$

and

$$\dot{\mathcal{S}}_0^-(\Lambda) \sim \begin{cases} 1, & d<2 \\ \Lambda^{d-2}, & d\geq 2 \end{cases}, \tag{6.97}$$

where we denote by Λ^0 a logarithmic divergence.

We may perform an identical procedure in the PL case, although we leave the details of this calculation in appendix D to simplify the presentation here. We note, however, that in the PL phase the scaling of $\dot{\mathcal{S}}_0^\pm(\Lambda)$ with the ultraviolet cutoff Λ changes. For our present case where $d=2$, we find that $\dot{\mathcal{S}}_0^+\sim\Lambda^2$ while $\dot{\mathcal{S}}_0^-\sim\Lambda^4$.

Similarly to our treatment in sections 6.2 and 6.3, we find good agreement between predictions and the results from simulations. In fig. 6.4 we demonstrate this comparison for both of the homogeneous phases. Furthermore, all considerations extend straightforwardly to the level where we explicitly track the current $\boldsymbol{J}$; the scaling exponent χ_J^+ of the EPR $\dot{\mathcal{S}}_J^+$ is at this level given by

$$\chi_J^+ = \begin{cases} 0, & \text{isotropic} \\ -1, & \text{PL} \end{cases} \tag{6.98}$$

while $\chi_J^- = \chi^-$. Note, however, that in the isotropic phase the coefficient of the leading order term of $\dot{\mathcal{S}}_J^+$ changes by virtue of eq. (6.80).

Finally, we consider inhomogeneous ground-states $(\rho_0, \boldsymbol{m}_0)$ (or $(\rho_0, \boldsymbol{J}_0, \boldsymbol{m}_0)$ at the level of $\boldsymbol{J}$), specifically the nonlinear PC, MPS and polar crystal states. Following the same reasoning as in section 6.3, we conclude that $\chi_J^\pm = \chi^\pm = -1$ for both the banded profiles and PCs. Indeed, these profiles still break both $\mathrm{T}^\pm$ and $\mathrm{PT}^\pm$ as before and are therefore strongly non-equilibrium at small noise. On the other hand, for the crystal state we find that $\chi^+ = 0$, while $\chi^- = -1$, as shown in fig. 6.4. We explain this by the observation that at $D = 0$ the ground-state solution is stationary, i.e. it has both $\partial_t \rho_0 = 0$ and $\partial_t m_0 = 0$. In particular, since $(\rho_0, \boldsymbol{m}_0)$ is independent of time, it is in fact also invariant under T^+. This may also be confirmed by inspection of e.g. eq. (D.14), where it is apparent that the stationary condition implies we must have $\chi^+ > -1$. Similarly, at the level of the current $\boldsymbol{J}$ we also find that $\chi_J^+ = 0$ for the crystal state. To see why this should be true, note that also $J_0 = 0$ for the polar crystal at ground-state level, meaning that there is no difference between $(\rho_0, \boldsymbol{J}_0, \boldsymbol{m}_0)$ and its time-reversal under T_J^+ for this phase. Coincidentally, this also shows that inhomogeneity is not necessarily sufficient alone to make the system strongly non-equilibrium. On the other hand, the polar crystal state is clearly not invariant under T^- or PT^- (nor T_J^-, PT_J^-), and so we conclude that $\chi_J^- = \chi^- = -1$.

6.6 Conclusion

In this chapter we have studied the entropy production rate in two related models of dry polar flocks, namely the HVM and DFM [1, 35, 99]. Our main results relate to the observation that the scaling of the EPR with the noise parameter D changes depending on the phase behaviour of the steady-state, and that the asymptotic scaling exponent takes integer values ≥ -1. This provides us with a handle to understand how the EPR reflects TRS violation at different orders due to small fluctuations away from the mean dynamics. In particular, strongly non-equilibrium behaviour is characterised by a divergent EPR in the limit $D \to 0$, and is caused by ground-state dynamics that violate detailed balance pathwise. On the other

	HVM	DFM			
	χ	χ^+	χ_J^+	χ^-	χ_J^-
Isotropic	1	0	0	0	0
PL	0	0	-1	0	0
MPS	-1	-1	-1	-1	-1
PC	-1	-1	-1	-1	-1
Crystal	N/A	0	0	-1	-1

Table 6.1 Scaling of the EPRs $\dot{\mathcal{S}} \sim D^{\chi}$, $\dot{\mathcal{S}}^{\pm} \sim D^{\chi^{\pm}}$ and $\dot{\mathcal{S}}_J^{\pm} \sim D^{\chi_J^{\pm}}$ with the noise parameter D when $D \ll 1$ for the isotropic, PL, MPS, PC and crystal regimes of the HVM and DFM.

hand, in the weak and effectively equilibrium cases where the scaling exponent is ≥ 0, the ground-state dynamics is pathwise equilibrium and entropy is produced only at the level of fluctuations. In particular, when the scaling exponent is strictly positive, the dynamics at small noise can be mapped onto equilibrium dynamics.

Both models studied display a transition from an isotropic gas to a PL, in addition to nonlinear PC and MPS phases [99, 100, 115]. The phase diagram of the DFM, which may be considered an extension of the HVM, contains additional structure at low densities where we find a novel crystal phase in which a stationary hexagonal lattice of high-density ridges surround low density valleys. Numerical integration of the DFM also shows that the same qualitative behaviour is retained at high densities even though the density dynamics are modified by the addition of a diffusive fluctuating current. This is, however, to be expected since the diffusive dynamics are only significant to the large scale behaviour when the advective transport is comparatively small [1].

Generally for systems with polar symmetry such as those considered here, the EPR may be constructed in two different ways depending on how we choose to implement time-reversal at the level of fluctuating trajectories [34, 35]. Specifically, we may choose whether the magnetisation density should transform as a velocity-like odd quantity or a head-to-tail-like even one under time-reversal, which changes the physics of the model. An exception to this is presented by the HVM, which is constructed in such a way that we only have one choice. Here, the continuity equation imposes a constraint on the space of observable trajectories, i.e. those that lie in the support of the transition probability density, which excludes the time-reversed trajectory of all observable trajectories when the magnetisation density does not flip sign. On the other hand, when the density advection is driven by independent fluctuations, such as in the DFM, we may consider both time-signatures. In addition, we may promote the current to an explicit dynamical variable and thus construct an additional EPR at this level [12]. Surprisingly, for this latter construction, we find that the additional knowledge

of the current changes the EPR only when the time-signature of the current differs from that of the magnetisation density. When it does, detailed balance at ground-state level for a homogeneously polarised system is broken by a mismatch between the density current and magnetisation density, analogously to the way in which TRS may be broken on the microscopic scale by ABPs or AOUPs [34].

For both time-signatures and models considered, as well as when explicitly tracking the density current, we find that the entropy production rate diverges in the limit $D \to 0$ in the MPS and PC regimes. We attribute this to the observation that both bands and PCs lead to traveling spatially asymmetrical profiles, which engenders a discrepancy between the time-forward and reversed movies that cannot be transformed away by parity. It is not sufficient that a profile is inhomogeneous alone, however, which is exemplified by the stationary crystal phase of the DFM. Indeed, in this case we find that when the magnetisation density does not change sign on time-reversal, the dynamics are only weakly non-equilibrium.

We also find that the PL phase is weakly non-equilibrium, except in the case where we explicitly track the density current and the magnetisation density is even under time-reversal, as noted above. When the magnetisation density transforms like a velocity, the EPR associated with ground-state flocking vanishes due to the rotational symmetry of the dynamics. Interestingly, we may conclude from this that if we were to break rotational symmetry a priori, for example by introducing an external driving field which is even under time-reversal, then flocking would in fact be strongly non-equilibrium when the magnetisation density is odd under time-reversal. This is not the case for the isotropic phase, however, which is at most weakly non-equilibrium in all cases.

We have also shown that for both the isotropic and PL phases, a linearisation of the dynamics at small noise allows us access the leading order coefficient of the EPR in the weakly non-equilibrium case by evaluating steady-state averages within the linear theory, extending similar work done in the context of Active Model B (and B+) by Nardini *et al.* [12] to polar models. In principle, this procedure can be adapted to access coefficients at arbitrary order in an expansion in small D, although the algebra involved becomes exceedingly complex at higher orders. Moreover, we find that our analytical predictions agree well with simulations, confirming that the procedure is well suited to analyse the EPR at small noise. In table 6.6 we summarise the scaling of the EPRs $\dot{\mathcal{S}}$ of the HVM in addition to $\dot{\mathcal{S}}^{\pm}$ and $\dot{\mathcal{S}}_J^{\pm}$ of the DFM with the noise parameter D for the various phases investigated.

Chapter 7

Conclusions and perspectives

Polar active matter provides a rich and interesting avenue for investigating how the irreversibility that is inherent to the microscopic motion imprints on larger scales in the form of currents. Particularly because a broad range of systems in this category, where particles have some mechanism of alignment, may organise into collectively moving steady-states that have no equivalent in scalar active systems. By the very nature of the collective motion, mass transport is always involved. However, as we have seen, whether these currents actually violate *time-reversal symmetry* (TRS) is usually a non-trivial question, and oftentimes a matter of definition. In the cases where the global dynamics do not violate TRS, one may have to study the fluctuations about the steady-state to reveal the non-equilibrium nature of the active phase. At high densities (low hydrodynamic noise), this implies that each active particle follow effectively equilibrium dynamics. On the other hand, when large scale currents do violate TRS, each active particle must contribute a constant non-vanishing entropy production on the microscopic scale at high densities, so that in consequence the total entropy production of the system diverges.

Our effort to investigate the properties of the *entropy production rate* (EPR) in polar active matter has involved a strategic choice of three model systems related to the much celebrated *Vicsek model* (VM) [17]. Firstly, these models can all be considered extensions of the VM to continuous time, and differ only as a result of the way in which they renormalise the diverging response time when coarse-graining from the time-discrete VM to the continuous time Langevin representation. This facilitates direct comparison of the results we have obtained for the three models. Perhaps more importantly, however, the *mean-field theory* (MFT) of these models can all be considered exact representations of the single particle dynamics at infinite density. In consequence, we are provided with an essential tool that allows us to analytically assess the EPR of the microscopic models when density fluctuations vanish.

The simplest choice of renormalisation prescribes a constant value to the response time when passing to the continuous representation. This leads to a model that we refer to in this thesis as the *continuous-time Vicsek model* (CTVM). Having first reviewed the phenomenology of the classical VM, we confirm in chapter 3 that the constant response time indeed reproduces the familiar phase diagram of the VM. Specifically, we observe from simulations that the phase diagram can be considered analogous to the canonical liquid-gas phase diagram in the ρ-T plane, with a critical point that is shifted to infinite density. In the miscibility gap, the CTVM exhibits *microphase-separation* (MPS) with a smectic arrangement of high-density collectively moving bands traveling against an isotropic gaseous background. At high temperatures (or low densities), above the upper binodal, there is no collective motion and the system behaves as an isotropic gas. Below the lower binodal, a collectively moving *polar liquid* (PL) phase fills the simulation domain.

Many features of the phase diagram of the CTVM can be explained by studying the MFT of the model. Being exact at infinite density, it is able to predict the precise location of the second order critical point where the two binodals converge. However, since the limit has already been taken when deriving the MFT, the density does not enter. This renders the theory unable to provide further details on the shape of the binodals. Still, we are able to confirm in section 3.4 that the PL phase is indeed linearly stable in the CTVM in the region where MFT is valid. This provides us with an important point of comparison when investigating the other choices of response times we have considered in this thesis. Finally, we also derive from the MFT a mapping to the 'Toner-Tu' theory space, employing a method initially developed by Farrell *et al.* [82].

In contrast to the CTVM and VM, the phase diagram that results from choosing a response time proportional to the local magnetisation does not contain a PL region. Instead, this choice leads to a model that we refer to as the Active XY model (AXYM), which, as we saw in chapter 4, instead displays a novel *polar cluster* (PC) phase at low temperatures where high density polarised clusters travel against an isotropic gaseous background. Still, the same procedure as that applied to derive the MFT of the CTVM can be utilised for the AXYM. In agreement with our numerical simulations, this allowed us to confirm in section 4.3 that the PL distribution obtained from the MFT of the AXYM is indeed always linearly unstable at infinite density, in contrast to the CTVM.

Importantly, the alignment interaction resulting from the AXYM choice of response time turns out to be locally extensive. From a modeling perspective, this may be a necessary requirement for certain physical systems. For example, one would expect that the interaction is proportional to the local density in a suspension of active particles with magnetic dipole

moments. In biological flocking systems, however, there is no a priori reason to assume that the alignment should be sensitive to the local signal strength.

Moreover, this local extensivity also turns out to be crucial to the resulting dynamics. In particular, when the effective interaction strength experienced by a particle grows proportionally to the local density, local freezing will energetically favor configurations in which cold high-density clusters are formed. It is this effect which causes the destabilisation of the PL phase in the AXYM, and leads instead to the PC phase. To back up this claim, we investigated further in section 4.4 the hydrodynamic theory in the Toner-Tu theory space representing the AXYM, identifying in particular a term in an effective free-energy functional responsible for the destabilisation. In addition, we showed in section 4.5 that by choosing instead a response time which is bounded and has the same MFT as the AXYM, leading to a model we refer to as the *non-extensive Active XY model* (NeAXYM), the PL phase is stabilised.

The differences apparent in the phenomenologies of the three models also has important consequences for their EPRs. In chapter 5 we presented numerical results that demonstrate that the *EPR per particle* (EPRP) reaches a constant non-zero value at small hydrodynamic noise (high density) in the PC phase of the AXYM. This stands in stark contrast to the CTVM, where the limiting spin dynamics in the PL phase described by MFT are in fact equilibrium dynamics, reflected by a vanishing EPRP. Although explained more thoroughly in chapter 6, this can in fact be seen as a consequence of the *global* breaking of TRS by the current engendered by the collectively moving clusters in the AXYM. More specifically, since the mean global PC dynamics violate TRS, the total EPR of the system diverges linearly in the density when density fluctuations are suppressed. In consequence, the contribution to the total EPR from each active particle must remain finite even when the limit of infinite density is taken.

Although the spin dynamics in the PL phase of the CTVM become equilibrium in the limit of infinite density, the same does not hold for the global dynamics. This is reflected by the *EPR per unit volume* (EPRV), which reaches a constant non-zero value in the PL phase at high density. In this case, hydrodynamic fluctuations are responsible for breaking TRS, rather than the mean global dynamics. Moreover, they break TRS in such a way that even for infinitesimally small density fluctuations, a finite amount EPR is produced by the system as a whole. Another way to phrase this is to say that at low noise, in the PL phase of the CTVM, density ripples diffuse irreversibly, while spin fluctuations decay reversibly.

A first important consequence of this ‘weak’ breaking of TRS displayed by the PL phase of the CTVM at high densities is that the MFT developed for the microscopic model cannot account for the finite limiting value of the EPRV. This is because, as mentioned above, the

MFT was derived by explicitly taking the infinite density limit to produce a single particle dynamics. Clearly, one cannot recover the EPRV of the system only from the knowledge that the EPRP is zero when sitting at infinite density. To remedy this, we included in section 5.3 finite-density fluctuations in a systematic expansion of the EPRV by utilising a Dean equation (DE) representation of the microscopic dynamics. This allowed us to analytically assess the EPRV at finite density first, before taking the infinite density limit. Moreover, this approach provided us explicit confirmation that the $\mathcal{O}(1)$ contribution to the EPRV at high densities could be attributed to the finite-density fluctuations (cf. eq. (5.53)).

Another insight is uncovered by realising that the global properties of the PL phase, similarly to the PC phase of the AXYM, again places significant constraints on the microscopic spin dynamics. Since the mean global dynamics does not break TRS, the total EPR of the system cannot diverge at low noise. Consequently, the EPRP must vanish in this limit, thus implying that the spin dynamics must be equilibrium at infinite density.

In chapter 6, this relationship between the global properties of the system and the total EPR was investigated more systematically. Specifically, by starting with the *hydrodynamic Vicsek model* (HVM) and the related *diffusive flocking model* (DFM) from the outset, we could investigate the scaling of the total EPR with the noise intensity in all phases reported for the microscopic models. Moreover, this allowed us to generalise our results to cases where different protocols for time-reversal could be chosen, as well as to include the density current as an independent fluctuating degree of freedom.

To extend the results in this thesis, we believe an interesting route to pursue for future investigations is to study the large deviations of the EPR in polar active matter. Particularly with the perspective of understanding the mechanisms that realise large fluctuations in the EPR. One motivation for this is to gain a better understanding of the irreversibility in these systems on a dynamical level. Another is to investigate how entropy production, or dissipation, may control the collective behaviour in them, thus connecting our research to other recent works on scalar active systems [28, 118, 119].

References

[1] Marchetti M C, Joanny J F, Ramaswamy S, Liverpool T B, Prost J, Rao M and Simha R A 2013 *Rev. Mod. Phys.* **85**(3) 1143–1189

[2] Ramaswamy S 2010 *Annu. Rev. Condens. Matter Phys.* **1**(1) 323–345

[3] Cates M E and Tailleur J 2015 *Annu. Rev. Condens. Matter Phys.* **6**(1) 219–244

[4] Kumar A, Maitra A, Sumit M, Ramaswamy S and Shivashankar G V 2014 *Sci. Rep.* **4**(1) 3781

[5] Katz Y, Tunstrøm K, Ioannou C C, Huepe C and Couzin I D 2011 *Proc. Natl. Acad. Sci. USA* **108**(46) 18720–18725

[6] Bialek W, Cavagna A, Giardina I, Mora T, Silvestri E, Viale M and Walczak A M 2012 *Proc. Natl. Acad. Sci. USA* **109**(13) 4786–4791

[7] Bain N and Bartolo D 2019 *Science* **363**(6422) 46–49

[8] Palacci J, Sacanna S, Steinberg A P, Pine D J and Chaikin P M 2013 *Science* **339**(6122) 936–940

[9] Deblais A, Barois T, Guerin T, Delville P H, Vaudaine R, Lintuvuori J S, Boudet J F, Baret J C and Kellay H 2018 *Phys. Rev. Lett.* **120**(18) 188002

[10] Deseigne J, Dauchot O and Chaté H 2010 *Phys. Rev. Lett.* **105**(9) 098001

[11] Kumar N, Soni H, Ramaswamy S and Sood A K 2014 *Nat. Commun.* **5**(1) 4688

[12] Nardini C, Fodor E, Tjhung E, van Wijland F, Tailleur J and Cates M E 2017 *Phys. Rev. X* **7**(2) 021007

[13] Caballero F, Nardini C and Cates M E 2018 *J. Stat. Mech.* **2018**(12) 123208

[14] Caballero F, Nardini C, van Wijland F and Cates M E 2018 *Phys. Rev. Lett.* **121**(2) 020601

[15] Caballero F and Cates M E 2020 *Phys. Rev. Lett.* **124**(24) 240604

[16] Chaté H 2020 *Annu. Rev. Condens. Matter Phys.* **11**(1) 189–212

[17] Vicsek T, Czirók A, Ben-Jacob E, Cohen I and Shochet O 1995 *Phys. Rev. Lett.* **75**(6) 1226–1229

[18] Toner J and Tu Y 1998 *Phys. Rev. E* **58**(4) 4828–4858

[19] Toner J, Tu Y and Ramaswamy S 2005 *Ann. Phys.* **318**(1) 170–244

[20] Toner J 2012 *Phys. Rev. Lett.* **108**(8) 088102

[21] Sekimoto K 2010 *Stochastic Energetics* 1st ed (*Lecture Notes in Physics* vol 799) (Berlin: Springer)

[22] Kubo R, Toda M and Hashitsume N 1991 *Statistical Physics II* 2nd ed (*Springer Series in Solid-State Sciences* vol 31) (Berlin: Springer)

[23] Zwanzig R 2001 *Nonequilibrium Statistical Mechanics* (Oxford: Oxford University Press)

[24] Lebowitz J L and Spohn H 1998 *J. Stat. Phys.* **95**(1) 333–365

[25] Seifert U 2012 *Rep. Prog. Phys.* **75**(12) 126001

[26] Wittkowski R, Tiribocchi A, Stenhammar J, Allen R J, Marenduzzo D and Cates M E 2014 *Nat. Commun.* **5**(1) 4351

[27] Tailleur J and Cates M E 2008 *Phys. Rev. Lett.* **100**(21) 218103

[28] Fodor E, Jack R L and Cates M E 2021 (*Preprint* arXiv:2104.06634)

[29] Ganguly C and Chaudhuri D 2013 *Phys. Rev. E* **88**(3) 032102

[30] Shim P S, Chun H M and Noh J D 2016 *Phys. Rev. E* **93**(1) 012113

[31] Fodor E, Nardini C, Cates M E, Tailleur J, Visco P and van Wijland F 2016 *Phys. Rev. Lett.* **117**(3) 038103

[32] Crosato E, Prokopenko M and Spinney R E 2019 *Phys. Rev. E* **100**(4) 042613

[33] Dabelow L, Bo S and Eichhorn R 2019 *Phys. Rev. X* **9**(2) 021009

[34] Shankar S and Marchetti M C 2018 *Phys. Rev. E* **98**(2) 020604

[35] Dadhichi L P, Maitra A and Ramaswamy S 2018 *J. Stat. Mech.* **2018**(12) 123201

[36] Borthne Ø L, Fodor É and Cates M E 2020 *New J. Phys.* **22**(12) 123012

[37] Evans D J, Cohen E G D and Morriss G P 1993 *Phys. Rev. Lett.* **71**(15) 2401–2404

[38] Seifert U 2008 *Eur. Phys. J. B* **64**(3) 423–431

[39] Kurchan J 1998 *J. Phys. A* **31**(16) 3719–3729

[40] Hummer G and Szabo A 2001 *Proc. Natl. Acad. Sci. USA* **98**(7) 3658–3661

[41] Jarzynski C 1997 *Phys. Rev. Lett.* **78**(14) 2690–2693

[42] Jarzynski C 1997 *Phys. Rev. E* **56**(5) 5018–5035

[43] Crooks G E 1999 *Phys. Rev. E* **60**(3) 2721–2726

[44] Crooks G E 2000 *Phys. Rev. E* **61**(3) 2361–2366

[45] Barato A C and Seifert U 2015 *Phys. Rev. Lett.* **114**(15) 158101

[46] Gingrich T R, Horowitz J M, Perunov N and England J L 2016 *Phys. Rev. Lett.* **116**(12) 120601

[47] Horowitz J M and Gingrich T R 2020 *Nat. Phys.* **16**(1) 15–20

[48] Hasegawa Y and Vu T V 2019 *Phys. Rev. E* **99**(6) 062126

[49] Vu T V and Hasegawa Y 2019 *Phys. Rev. E* **100**(3) 032130

[50] Dechant A and Sasa S i 2018 *J. Stat. Mech.* **2018**(6) 063209

[51] Dechant A 2018 *J. Phys. A: Math. Theor.* **52**(3) 035001

[52] Niggemann O and Seifert U 2020 *J. Stat. Phys.* **178**(5) 1142–1174

[53] Chun H M, Fischer L P and Seifert U 2019 *Phys. Rev. E* **99**(4) 042128

[54] Fischer L P, Pietzonka P and Seifert U 2018 *Phys. Rev. E* **97**(2) 022143

[55] Fischer L P, Chun H M and Seifert U 2020 *Phys. Rev. E* **102**(1) 012120

[56] Seifert U 2005 *Phys. Rev. Lett.* **95**(4) 040602

[57] Maes C 1999 *J. Stat. Phys.* **95**(1) 367–392

[58] Markovich T, Fodor E, Tjhung E and Cates M E 2020 (*Preprint* arXiv:2008.06735)

[59] Gardiner C 2009 *Stochastic Methods: A Handbook for the Natural and Social Sciences* 4th ed (*Springer Series in Synergetics*) (Berlin: Springer)

[60] Pavliotis G A 2014 *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations* (*Texts in Applied Mathematics* vol 60) (New York: Springer)

[61] Applebaum D 2009 *Lévy Processes and Stochastic Calculus* 2nd ed (*Cambridge Studies in Advanced Mathematics* vol 116) (Cambridge: Cambridge University Press)

[62] Rogers L C G and Williams D 2000 *Diffusions, Markov Processes and Martingales* 2nd ed (*Cambridge Mathematical Library* vol 2) (Cambridge: Cambridge University Press)

[63] Kaiser M, Jack R L and Zimmer J 2018 *J. Stat. Phys.* **170**(6) 1019–1050

[64] Song Y and Hyeon C 2021 *J. Chem. Phys.* **154**(13) 130901

[65] Pietzonka P and Seifert U 2018 *Phys. Rev. Lett.* **120**(19) 190602

[66] Pietzonka P, Barato A C and Seifert U 2016 *J. Stat. Mech.* **2016**(12) 124004

[67] Seifert U 2018 *Physica A* **504** 176–191

[68] Schmiedl T, Speck T and Seifert U 2007 *J. Stat. Phys.* **128**(1) 77–93

[69] Seifert U 2005 *EPL* **70**(1) 36–41

[70] Imparato A, Peliti L, Pesce G, Rusciano G and Sasso A 2007 *Phys. Rev. E* **76**(5) 050101

[71] Gallavotti G and Cohen E G D 1995 *J. Stat. Phys.* **80**(5) 931–970

[72] Demichev A and Chaichian M 2001 *Path Integrals in Physics I: Stochastic Processes and Quantum Mechanics* 1st ed (Boca Raton, FL: CRC Press)

[73] Tauber U C 2014 *Critical Dynamics: A Field Theory Approach to Equilibrium and Non-Equilibrium Scaling Behaviour* (Cambdrige: Cambridge University Press)

[74] Touchette H 2009 *Phys. Rep.* **478**(1-3) 1–69

[75] Bertini L, De Sole A, Gabrielli D, Jona-Lasinio G and Landim C 2015 *Rev. Mod. Phys.* **87**(2) 593–636

[76] Cates M E 2019 (*Preprint* arXiv:1904.01330)

[77] Romanczuk P, Bär M, Ebeling W, Lindner B and Schimansky-Geier L 2012 *Eur. Phys. J. Spec. Top.* **202**(1) 1–162

[78] Solon A P, Cates M E and Tailleur J 2015 *Eur. Phys. J. Spec. Top.* **224**(7) 1231–1262

[79] Bolley F, Cañizo J A and Carrillo J A 2012 *Appl. Math. Lett.* **25**(3) 339–343

[80] Brillinger D R 1997 *J. Theor. Probab.* **10**(2) 429–443

[81] Degond P and Motsch S 2008 *Math. Models Methods Appl. Sci.* **18**(supp01) 1193–1215

[82] Farrell F D C, Marchetti M C, Marenduzzo D and Tailleur J 2012 *Phys. Rev. Lett.* **108**(24) 248101

[83] Degond P, Frouvelle A and Liu J G 2013 *J. Nonlinear Sci.* **23**(3) 427–456

[84] Barré J, Chétrite R, Muratori M and Peruani F 2015 *J. Stat. Phys.* **158**(3) 589–600

[85] Mahault B, Patelli A and Chaté H 2018 *J. Stat. Mech.* **2018**(9) 093202

[86] Laighléis E O, Evans M R and Blythe R A 2018 *Phys. Rev. E* **98**(6) 062127

[87] Peshkov A, Ngo S, Bertin E, Chaté H and Ginelli F 2012 *Phys. Rev. Lett.* **109**(9) 098101

[88] Castellana M, Bialek W, Cavagna A and Giardina I 2016 *Phys. Rev. E* **93**(5) 052416

[89] Martin D, Chaté H, Nardini C, Solon A, Tailleur J and Van Wijland F 2021 *Phys. Rev. Lett.* **126**(14) 148001

[90] Ginelli F, Peruani F, Bär M and Chaté H 2010 *Phys. Rev. Lett.* **104**(18) 184502

[91] Peshkov A, Aranson I S, Bertin E, Chaté H and Ginelli F 2012 *Phys. Rev. Lett.* **109**(26) 268701

[92] Peshkov A, Bertin E, Ginelli F and Chaté H 2014 *Eur. Phys. J. Spec. Top.* **223**(7) 1315–1344

[93] Ngo S, Peshkov A, Aranson I S, Bertin E, Ginelli F and Chaté H 2014 *Phys. Rev. Lett.* **113**(3) 038302

[94] Patelli A, Djafer-Cherif I, Aranson I S, Bertin E and Chaté H 2019 *Phys. Rev. Lett.* **123**(25) 258001

[95] Bertin E, Droz M and Grégoire G 2006 *Phys. Rev. E* **74**(2) 022101

[96] Bertin E, Droz M and Grégoire G 2009 *J. Phys. A: Math. Theor.* **42**(44) 445001

[97] Grégoire G and Chaté H 2004 *Phys. Rev. Lett.* **92**(2) 025702

[98] Solon A P and Tailleur J 2013 *Phys. Rev. Lett.* **111**(7) 078101

[99] Solon A P, Chaté H and Tailleur J 2015 *Phys. Rev. Lett.* **114**(6) 068101

[100] Solon A P, Caussin J B, Bartolo D, Chaté H and Tailleur J 2015 *Phys. Rev. E* **92**(6) 062111

[101] Solon A P and Tailleur J 2015 *Phys. Rev. E* **92**(4) 042119

[102] Ginelli F 2016 *Eur. Phys. J. Spec. Top.* **225**(11-12) 2099–2117

[103] Toner J 2012 *Phys. Rev. E* **86**(3) 031918

[104] Ballerini M, Cabibbo N, Candelier R, Cavagna A, Cisbani E, Giardina I, Lecomte V, Orlandi A, Parisi G, Procaccini A, Viale M and Zdravkovic V 2008 *Proc. Natl. Acad. Sci. USA* **105**(4) 1232–1237

[105] Degond P and Yang T 2010 *Math. Models Methods Appl. Sci.* **20**(supp01) 1459–1490

[106] Watson G N 1995 *A Treatise on the Theory of Bessel Functions* (Cambridge: Cambridge University Press)

[107] Dean D S 1996 *J. Phys. A: Math. Gen.* **29**(24) L613–L617

[108] Greene R E and Krantz S G 2006 *Function Theory of One Complex Variable* (Providence, RI: American Mathematical Society)

[109] Krantz S G 2008 *A Guide to Complex Variables* (*Dolciani Mathematical Expositions Series* vol 32) (Washington, DC: Mathematical Association of America)

[110] Chepizhko O, Saintillan D and Peruani F 2021 *Soft Matter*

[111] Hohenberg P C and Halperin B I 1977 *Rev. Mod. Phys.* **49**(3) 435–479

[112] Bellac M L 1991 *Quantum and Statistical Field Theory* (*Oxford Science Publications*) (Oxford: Clarendon Press)

[113] Arovas D 2010 Physics 210a: Statistical physics *Lecture notes* URL https://courses.physics.ucsd.edu/2010/Spring/physics210a/LECTURES/CH07.pdf

[114] Canuto C, Hussaini M Y, Quarteroni A and Zang T A 2006 *Spectral Methods: Fundamentals in Single Domains* (*Scientific Computation*) (Berlin: Springer)

[115] Gopinath A, Hagan M F, Marchetti M C and Baskaran A 2012 *Phys. Rev. E* **85**(6) 061903

[116] Chen Q, Patelli A, Chaté H, Ma Y and Shi X 2017 *Phys. Rev. E* **96**(2) 020601

[117] Fodor E, Nemoto T and Vaikuntanathan S 2020 *New J. Phys.* **22**(1) 013052

[118] Nemoto T, Fodor E, Cates M E, Jack R L and Tailleur J 2019 *Phys. Rev. E* **99**(2) 022605

[119] Keta Y E, Fodor E, van Wijland F, Cates M E and Jack R L 2021 *Phys. Rev. E* **103**(2) 022603

[120] Cagnetta F, Corberi F, Gonnella G and Suma A 2017 *Phys. Rev. Lett.* **119**(15) 158002

Appendix A

Coarse-graining of microscopic Vicsek models

In chapters 3 and 4, we saw that all three microscopic models of collective motion considered in this thesis, namely the continuous-time Vicsek model (CTVM), Active XY model (AXYM) and non-extensive Active XY model (NeAXYM), could be mapped onto Toner-Tu type hydrodynamic theories. There, we referred to this space of theories as the hydrodynamic Vicsek model (HVM) class. In this appendix, we construct explicitly the mappings for each of the three models onto the HVM theory space, which is valid in the vicinity of their critical points at infinite density. The derivation we present here is a more detailed analysis building on that originally presented by Farrell *et al.* [82] for the AXYM, which also includes the CTVM and NeAXYM.

To start, we begin by introducing the general set of mean-field equations

$$\dot{\boldsymbol{x}}_i = \boldsymbol{e}_i, \tag{A.1}$$

$$\dot{\theta}_i = \beta \boldsymbol{e}_i' \cdot \boldsymbol{j}^{\mathcal{C}}(\boldsymbol{x}_i) + \sqrt{2}\xi_i, \tag{A.2}$$

where $\boldsymbol{e}_i = \boldsymbol{e}(\theta_i) \equiv (\cos\theta_i, \sin\theta_i)$ and the molecular field $\boldsymbol{j}(\boldsymbol{x})$ is depends on the specific model under consideration. Specifically, comparing with eqs. (3.25), (4.13) and (4.57) for, respectively, the CTVM, AXYM and NeAXYM, we see that we should take

$$\boldsymbol{j}^{\mathcal{C}}(\boldsymbol{x}) = \begin{cases} \hat{\boldsymbol{m}}^{\mathcal{C}}(\boldsymbol{x}) & \text{CTVM}, \\ L^2 \boldsymbol{m}^{\mathcal{C}}(\boldsymbol{x}) & \text{AXYM}, \\ \boldsymbol{m}^{\mathcal{C}}(\boldsymbol{x}) / \varrho^{\mathcal{C}}(\boldsymbol{x}) & \text{NeAXYM}, \end{cases} \tag{A.3}$$

where we recall that

$$\varrho^{\mathcal{C}}(\boldsymbol{x}) = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}' \int_0^{2\pi} \mathrm{d}\theta\, C_0(|\boldsymbol{x}-\boldsymbol{x}'|) p(\boldsymbol{x}',\theta), \tag{A.4}$$

$$\boldsymbol{m}^{\mathcal{C}}(\boldsymbol{x}) = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}' \int_0^{2\pi} \mathrm{d}\theta\, C_0(|\boldsymbol{x}-\boldsymbol{x}'|) \boldsymbol{e}(\theta) p(\boldsymbol{x}',\theta), \tag{A.5}$$

$p = \lim_{N\to\infty} N^{-1} \sum_{i=1}^{N} \delta(\boldsymbol{x}-\boldsymbol{x}_i)\delta(\theta-\theta_i)$ is the one-particle distribution function at infinite density, C_0 is an isotropic contact kernel with interaction length ℓ_0 (see section 3.2) and $\hat{\boldsymbol{m}}^{\mathcal{C}} = \boldsymbol{m}^{\mathcal{C}}/m^{\mathcal{C}}$. The one-particle distribution function p evolves, as in eqs. (3.26), (4.14) and (4.59), according to the Fokker-Planck equation (FPE)

$$\partial_t p + \boldsymbol{e}\cdot\nabla p = \partial_\theta \left(\mathrm{e}^{\beta \boldsymbol{e}\cdot\boldsymbol{j}^{\mathcal{C}}} \partial_\theta \left(\mathrm{e}^{-\beta \boldsymbol{e}\cdot\boldsymbol{j}^{\mathcal{C}}} p \right) \right). \tag{A.6}$$

Recall also that for the mean-field CTVM the critical point is located at $\beta_c = 0$, while for the AXYM and NeAXYM it is located at $\beta_c = 2$.

As argued in section 3.5, at infinite density the evolution of the system *on the hydrodynamic scale* is deterministic, meaning that we do not observe fluctuations due to noise in any hydrodynamic observable (such as the molecular field $\boldsymbol{j}^{\mathcal{C}}$). Our approach then was therefore to imagine introducing a small amplitude perturbation at finite $\beta \approx \beta_c$ of wavelength ℓ_p, that is large compared with the microscopic diffusion length scale $\ell_D = 1$ and interaction length ℓ_0, and study the deterministic relaxation of the hydrodynamic fields. In this limit, $\nabla \sim \ell_p^{-1}$, and so a gradient expansion is valid, implying in particular that we may expand

$$\boldsymbol{j}^{\mathcal{C}}(\boldsymbol{x}) = \boldsymbol{j}(\boldsymbol{x}) + \mathcal{O}\left((\ell_0/\ell_p)^2\right), \tag{A.7}$$

where

$$\boldsymbol{j}(\boldsymbol{x}) = \begin{cases} \hat{\boldsymbol{m}}(\boldsymbol{x}) & \text{CTVM}, \\ L^2 \boldsymbol{m}(\boldsymbol{x}) & \text{AXYM}, \\ \boldsymbol{m}(\boldsymbol{x})/\varrho(\boldsymbol{x}) & \text{NeAXYM}, \end{cases} \tag{A.8}$$

and the hydrodynamic fields $\varrho(\boldsymbol{x})$ and $\boldsymbol{m}(\boldsymbol{x})$ are given by

$$\varrho(\boldsymbol{x}) = \int_0^{2\pi} \mathrm{d}\theta\, p(\boldsymbol{x},\theta), \tag{A.9}$$

$$\boldsymbol{m}(\boldsymbol{x}) = \int_0^{2\pi} \mathrm{d}\theta\, \boldsymbol{e}(\theta) p(\boldsymbol{x},\theta). \tag{A.10}$$

For the remainder of this derivation, we ignore all higher order terms in ℓ_0/ℓ_p, which effectively corresponds to taking the limit of vanishing interaction length $\ell_0 \to 0$. Combining

eqs. (A.2) and (A.7) we then obtain for the angular dynamics

$$\dot{\theta}_i = \beta \boldsymbol{e}_i' \cdot \boldsymbol{j}(\boldsymbol{x}_i) + \sqrt{2}\xi_i. \tag{A.11}$$

Correspondingly, from eqs. (A.6) and (A.7) we also obtain for the one-particle distribution p the FPE

$$\partial_t p + \boldsymbol{e} \cdot \nabla p = \partial_\theta \left(\mathrm{e}^{\beta \boldsymbol{e}\cdot\boldsymbol{j}} \partial_\theta \left(\mathrm{e}^{-\beta \boldsymbol{e}\cdot\boldsymbol{j}} p \right) \right). \tag{A.12}$$

In order to obtain the HVM equations for the CTVM, AXYM and NeAXYM, we will seek to derive equations for the moments ϱ and $\boldsymbol{m}$ of the distribution p from the FPE in eq. (A.12). To do this, we follow [82] and compute equations for the Fourier coefficients p_k of p, defined here by

$$p_k(\boldsymbol{x}) = \int_0^{2\pi} \mathrm{d}\theta \, p(\boldsymbol{x}, \theta) \mathrm{e}^{-ik\theta}. \tag{A.13}$$

In particular, from eq. (A.12) we find that the p_k solve

$$\partial_t p_k + \frac{1}{2} \nabla \cdot \begin{pmatrix} p_{k-1} + p_{k+1} \\ -i(p_{k-1} - p_{k+1}) \end{pmatrix} = -k^2 p_k - i\frac{\beta k}{2} \boldsymbol{j} \cdot \begin{pmatrix} i(p_{k-1} - p_{k+1}) \\ p_{k-1} + p_{k+1} \end{pmatrix}. \tag{A.14}$$

Moreover, by comparing eq. (A.13) with the definitions in eqs. (A.9) and (A.10) we observe that the moments $\varrho = p_0$, $m_x = \mathrm{Re}\, p_1$ and $m_y = -\mathrm{Im}\, p_1$, where $\boldsymbol{m} = (m_x, m_y)^T$. In addition, we introduce $\boldsymbol{q} = (q_x, q_y)^T$ and $\boldsymbol{r} = (r_x, r_y)^T$ with

$$\boldsymbol{q} = \begin{pmatrix} \mathrm{Re}\, p_2 \\ -\mathrm{Im}\, p_2 \end{pmatrix}, \qquad \boldsymbol{r} = \begin{pmatrix} \mathrm{Re}\, p_3 \\ -\mathrm{Im}\, p_3 \end{pmatrix}. \tag{A.15}$$

By taking the real and imaginary parts of eq. (A.14) we then find for $k = 0$ the continuity equation

$$\partial_t \varrho + \nabla \cdot \boldsymbol{m} = 0. \tag{A.16}$$

Similarly, for $k = 1$ we obtain a pair of equations for the components m_x, m_y of $\boldsymbol{m}$, specifically

$$\partial_t m_x + \frac{1}{2} \nabla \cdot \begin{pmatrix} \varrho + q_x \\ q_y \end{pmatrix} = -m_x + \frac{\beta}{2} \boldsymbol{j} \cdot \begin{pmatrix} \varrho - q_x \\ -q_y \end{pmatrix}, \tag{A.17}$$

$$\partial_t m_y + \frac{1}{2} \nabla \cdot \begin{pmatrix} q_y \\ \varrho - q_x \end{pmatrix} = -m_y + \frac{\beta}{2} \boldsymbol{j} \cdot \begin{pmatrix} -q_y \\ \varrho + q_x \end{pmatrix}. \tag{A.18}$$

The final pair of equations we will need for now is obtained by setting $k = 2$, which gives

$$\partial_t q_x + \frac{1}{2}\nabla \cdot \begin{pmatrix} m_x + r_x \\ -m_y + r_y \end{pmatrix} = -4q_x + \beta \boldsymbol{j} \cdot \begin{pmatrix} m_x - r_x \\ -m_y - r_y \end{pmatrix}, \tag{A.19}$$

$$\partial_t q_y + \frac{1}{2}\nabla \cdot \begin{pmatrix} m_y + r_y \\ m_x - r_x \end{pmatrix} = -4q_y + \beta \boldsymbol{j} \cdot \begin{pmatrix} m_y - r_y \\ m_x + r_x \end{pmatrix}. \tag{A.20}$$

To get a closure relationship for $\boldsymbol{q}$, we first observe that for long times $t \gg 1$, the solution to eqs. (A.19) and (A.20) decays onto the slow manifold with

$$4q_x = \left[\beta \boldsymbol{j} - \frac{1}{2}\nabla\right] \cdot \begin{pmatrix} m_x \\ -m_y \end{pmatrix} - \left[\beta \boldsymbol{j} + \frac{1}{2}\nabla\right] \cdot \begin{pmatrix} r_x \\ r_y \end{pmatrix}, \tag{A.21}$$

$$4q_y = \left[\beta \boldsymbol{j} - \frac{1}{2}\nabla\right] \cdot \begin{pmatrix} m_y \\ m_x \end{pmatrix} - \left[\beta \boldsymbol{j} + \frac{1}{2}\nabla\right] \cdot \begin{pmatrix} r_y \\ -r_x \end{pmatrix}. \tag{A.22}$$

Notice also that since we assume we are near the critical point with $\beta \approx \beta_c$, $\boldsymbol{m}$ can be assumed to be small. Similarly, we in fact have that for all three models, i.e. the CTVM, AXYM and NeAXYM, the product $\beta \boldsymbol{j}$ is small. Indeed, for the CTVM $\boldsymbol{j} = \hat{\boldsymbol{m}} \sim \mathcal{O}(1)$ and so $\beta \boldsymbol{j} \sim \mathcal{O}(\beta)$, while for the AXYM and NeAXYM $\beta \boldsymbol{j} \sim \mathcal{O}(m)$. Moreover, by assuming that the Fourier coefficients $p_k \to 0$ exponentially fast as $k \to 0$, which is necessary for the Fourier sum to be well-behaved, verifying that the scaling hypothesis where $\boldsymbol{q}$ is of quadratic order in m, β and ∇, while $\boldsymbol{r}$ is cubic, is consistent can be done fairly routinely (although this requires deriving one more set of equations in addition to eqs. (A.19) and (A.20)). In particular, to quadratic order, we have

$$q_x = \frac{1}{4}\left[\beta \boldsymbol{j} - \frac{1}{2}\nabla\right] \cdot \begin{pmatrix} m_x \\ -m_y \end{pmatrix} \tag{A.23}$$

$$q_y = \frac{1}{4}\left[\beta \boldsymbol{j} - \frac{1}{2}\nabla\right] \cdot \begin{pmatrix} m_y \\ m_x \end{pmatrix} \tag{A.24}$$

At this stage, combining eqs. (A.17), (A.18), (A.23) and (A.24), we have everything we need to form a closed set of equations for the density ϱ and magnetisation $\boldsymbol{m}$. First, however, we note that to lowest order, we in fact do not need eqs. (A.23) and (A.24) for q_x and q_y to form a (non-trivial) closed set of equations for ϱ and $\boldsymbol{m}$. Indeed, to first order in β and ∇ for the CTVM, or m and ∇ for the AXYM and NeAXYM, we simply ignore the products

between ∇ and $\boldsymbol{q}$, and $\boldsymbol{j}$ and $\boldsymbol{q}$, in eqs. (A.17) and (A.18) to obtain

$$\partial_t \boldsymbol{m} = -\boldsymbol{m} + \frac{\beta}{2}\varrho\,\boldsymbol{j} - \frac{1}{2}\nabla\varrho\,. \tag{A.25}$$

For the CTVM, this approximation leads to the result stated in eq. (3.62). However, eq. (A.25) does not support a stable polar liquid (PL) phase for any of the three models, which we showed explicitly in section 3.5 for the case of the CTVM. To verify this also for the AXYM and NeAXYM, one simply substitutes the appropriate expression for $\boldsymbol{j}$ in eq. (A.25).

To obtain the next order approximation, we substitute eqs. (A.23) and (A.24) into eqs. (A.17) and (A.18) as announced above. Rather than performing three separate calculations for each model, we observe that we may write

$$\boldsymbol{j} = \chi(\varrho, m)\boldsymbol{m}, \tag{A.26}$$

where χ is a model dependent function, specifically

$$\chi(\varrho, m) = \begin{cases} 1/m & \text{CTVM}, \\ L^2 & \text{AXYM}, \\ 1/\varrho & \text{NeAXYM}. \end{cases} \tag{A.27}$$

After some tedious algebra, we obtain the HVM equation for $\boldsymbol{m}$ (cf. eq. (3.84))

$$\partial_t \boldsymbol{m} + \lambda\,(\boldsymbol{m}\cdot\nabla)\,\boldsymbol{m} = -a\,(\varrho_c - \varrho)\,\boldsymbol{m} - bm^2\boldsymbol{m} + \frac{1}{16}\nabla^2\boldsymbol{m} + \frac{\kappa_1}{2}\nabla m^2 - \kappa_2\boldsymbol{m}\nabla\cdot\boldsymbol{m} - w_1\nabla\varrho, \tag{A.28}$$

where the coefficients are given by

$$\varrho_c = \frac{2}{\beta\chi}, \tag{A.29}$$

$$\lambda_{\alpha\beta} = \frac{\beta}{4}\left(\frac{3\chi}{4}\delta_{\alpha\beta} + \frac{\partial\chi}{\partial m}\frac{m_\alpha m_\beta}{m}\right), \tag{A.30}$$

$$a = \frac{\beta\chi}{2}, \tag{A.31}$$

$$b = \frac{\beta^2\chi^2}{8}, \tag{A.32}$$

$$\kappa_1 = \frac{5\beta\chi}{16}\left(1 + \frac{m}{5\chi}\frac{\partial\chi}{\partial m}\right), \tag{A.33}$$

$$\kappa_2 = \frac{5\beta\chi}{16}, \tag{A.34}$$

$$w_{1,\alpha\beta} = \frac{1}{2}\left(\left(1 - \frac{\beta m^2}{4}\frac{\partial\chi}{\partial\varrho}\right)\delta_{\alpha\beta} + \frac{\beta}{2}\frac{\partial\chi}{\partial\varrho}m_\alpha m_\beta\right). \tag{A.35}$$

Notice also that we allow the coefficients of the HVM to take matrix values. Substituting in the appropriate expression for χ from eq. (A.27) then, we thus immediately obtain eqs. (3.76), (4.42) and (4.67) as reported in the main text in chapters 3 and 4.

Finally, we mention that the nematic tensor field $\mathcal{Q} = \left(\mathcal{Q}_{\alpha\beta}\right)$ in eq. (5.61) can be written in terms of $\boldsymbol{q}$ as

$$\mathcal{Q} = \frac{1}{2}\begin{pmatrix} q_x & q_y \\ q_y & -q_x \end{pmatrix}. \tag{A.36}$$

Thus, using the expressions for q_x and q_y in eqs. (A.23) and (A.24) we can compute the two-sided contraction of $\mathcal{Q}$ with $\hat{\boldsymbol{m}}$ in the CTVM. Collecting terms, we find after some algebra that

$$\hat{m}_\alpha \mathcal{Q}_{\alpha\beta}\hat{m}_\beta = \frac{\beta}{8}m + \frac{1}{16}\left(\nabla\cdot\boldsymbol{m} - \frac{1}{m}\hat{\boldsymbol{m}}\cdot\nabla m^2\right) \tag{A.37}$$

to quadratic order in β and ∇. In particular, from this the result in eq. (5.63) follows immediately.

Appendix B

Calculation of spin autocorrelations in the CTVM

To characterise the spin fluctuations, in addition to the fluctuations of the entropy production rate (EPR), in the continuous-time Vicsek model (CTVM), it is useful to calculate the spin autocorrelation function. Here we aim to delineate how this may be done in steady-state at infinite density near the second order critical point at inverse temperature $\beta \approx 0$ as a perturbative expansion in small β. In particular, we will derive eqs. (3.35) and (3.37) for the autocorrelation function $C(t)$ and effective diffusion $D_{\text{eff}}(\beta)$ respectively, as well as eq. (5.31) for fluctuations in the stochastic EPR of the reservoir $\Delta\mathfrak{s}_i^{\text{res}}$. Note that our treatment here will be restricted to $d = 2$ for simplicity, although a completely analogous derivation can be performed for arbitrary d.

First, we note that under the steady mean-field dynamics at infinite density in eq. (3.31), particles only align with the constant global effective magnetisation $\hat{\boldsymbol{M}}$, so that each particle is described by independent Langevin equations. In consequence, we may safely omit the dependence on the particle index i, and simply consider the equations of motion (in Itô form, see e.g. eq. (5.4))

$$\dot{\boldsymbol{x}} = \boldsymbol{e}, \tag{B.1}$$

$$\dot{e}_\alpha = -e_\alpha + \left(\delta_{\alpha\beta} - e_\alpha e_\beta\right)\left(\beta\hat{M}_\beta + \sqrt{2}\xi_\beta\right), \tag{B.2}$$

where $\boldsymbol{x}$ describes the position of the particle, and $\boldsymbol{e}$ is its spin.

We begin by defining the projection of $\boldsymbol{e}$ on $\hat{\boldsymbol{M}}$, i.e. we let $u(t) \equiv \hat{\boldsymbol{M}} \cdot \boldsymbol{e}_i(t)$. Our first aim will then be to calculate the autocorrelation function

$$U_n(t) \equiv \langle u(0)u(t)^n \rangle \tag{B.3}$$

perturbatively as a series in small β. From this, the variance $\left\langle (\Delta\mathfrak{s}_i^{\mathrm{res}})^2 \right\rangle$ in eq. (5.31) follows immediately. Subsequently, we will see that this result can be employed also to derive $C(t)$ and $D_{\mathrm{eff}}(\beta)$.

Our strategy in the following will be to construct a hierarchy of equations for the moments U_n. We will then see that his hierarchy admits a solution as an expansion in small β, from which we can extract the solution for U_1 and thus derive $\left\langle (\Delta\mathfrak{s}_i^{\mathrm{res}})^2 \right\rangle$. To do this, we first remark that using standard Itô calculus (see e.g. chapter 2) one may show that

$$\frac{\mathrm{d}}{\mathrm{d}t}U_n(t) = n\left\langle u(0)u(t)^{n-1}\dot{u}(t)\right\rangle + \frac{1}{2}n(n-1)\left\langle u(0)u(t)^{n-2}\left[\dot{u}(t),\dot{u}(t)\right]\right\rangle, \tag{B.4}$$

where $[\cdot,\cdot]$ is the quadratic covariation process. Now, by taking the scalar product of eq. (B.2) for the dynamics of $\boldsymbol{e}$ with $\hat{\boldsymbol{M}}$, one may compute that

$$\dot{u} = -u + \beta\left(1-u^2\right) + \sqrt{2}\hat{M}_\alpha\left(\delta_{\alpha\beta} - e_\alpha e_\beta\right)\xi_\beta. \tag{B.5}$$

Thus, it follows immediately that the quadratic covariation of u is given by

$$[\dot{u},\dot{u}] = 2\left(1-u^2\right), \tag{B.6}$$

and so combining eqs. (B.4) to (B.6) we conclude that

$$\frac{\mathrm{d}}{\mathrm{d}t}U_n = -n^2U_n + n\beta\left(U_{n-1} - U_{n+1}\right) + n(n-1)U_{n-2} \tag{B.7}$$

for $n \geq 0$.

To solve eq. (B.7), we look for a series solution in β near $\beta = 0$, i.e. we employ the ansatz

$$U_n = \sum_{k=0}^{\infty} U_n^k \beta^k \tag{B.8}$$

and aim to determine the coefficients $U_n^k(t)$. By substituting this expansion back into eq. (B.7) and collecting terms at equal order in β, we obtain a hierarchy of equations for the coefficients U_n^k given by

$$\frac{\mathrm{d}}{\mathrm{d}t}U_n^0 = -n^2U_n^0 + n(n-1)U_{n-2}^0 \tag{B.9}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}U_n^k = -n^2U_n^k + n\left(U_{n-1}^{k-1} - U_{n+1}^{k-1}\right) + n(n-1)U_{n-2}^k, \quad k \geq 1 \tag{B.10}$$

Clearly, eq. (B.9) can be solved for the coefficients U_n^0 with $n \geq 0$ in sequence, with the first two non-trivial solutions U_1^0, U_2^0 given by (note that U_0^k is simply a constant for all $k \geq 0$)

$$U_1^0 = U_1^0(0)\mathrm{e}^{-t}, \tag{B.11}$$

$$U_2^0 = U_2^0(0)\mathrm{e}^{-4t} + \frac{1}{2}U_0^0\left(1-\mathrm{e}^{-4t}\right). \tag{B.12}$$

From this, we may then proceed to solve eq. (B.10) for the higher order coefficients U_n^k with $k \geq 1$. For example, with $n = 1$ we find that

$$\frac{\mathrm{d}}{\mathrm{d}t}U_1^k = -U_1^k + U_0^{k-1} - U_2^{k-1}, \tag{B.13}$$

from which it follows that

$$U_1^1 = U_1^1(0)\mathrm{e}^{-t} + \frac{\mathrm{e}^{-t}}{3}U_2^0(0)\left(\mathrm{e}^{-3t}-1\right) - \frac{1}{6}U_0^0\left(\mathrm{e}^{-4t}+\mathrm{e}^{-t}-3\right). \tag{B.14}$$

Computing higher order coefficients quickly becomes algebraically messy, although it can be done straightforwardly using standard computer algebra packages.

Initial values for the system in eqs. (B.9) and (B.10) can be determined by expanding the moments of the distribution p_{ss} in eq. (3.27) in small β. Specifically, since the steady-state distribution for the process in eqs. (B.1) and (B.2) is given by

$$p_{\mathrm{ss}}(\boldsymbol{x},\boldsymbol{\theta}) = \frac{\mathrm{e}^{\beta \boldsymbol{e}\cdot\hat{\boldsymbol{M}}}}{2\pi L^2 I_0(\beta)}, \tag{B.15}$$

it follows that

$$U_n(0) = \frac{1}{I_0(\beta)}\frac{\mathrm{d}^n}{\mathrm{d}\beta^n}I_1(\beta). \tag{B.16}$$

Expanding the right-hand side in small β we thus find that

$$U_0 = \frac{\beta}{2} - \frac{\beta^3}{16} + \mathcal{O}(\beta^5), \tag{B.17}$$

$$U_1(0) = \frac{1}{2} + \frac{\beta^2}{16} + \mathcal{O}(\beta^4), \tag{B.18}$$

$$U_2(0) = \frac{3}{8}\beta - \frac{\beta^3}{24} + \mathcal{O}(\beta^5), \tag{B.19}$$

$$U_3(0) = \frac{3}{8} + \frac{\beta^2}{16} + \mathcal{O}(\beta^4), \tag{B.20}$$

$$\vdots$$

Finally then, we may read off from eqs. (B.17) to (B.20) the initial values of the coefficients U_n^k, from which one may deduce using the procedure outlined above that

$$U_1 = \frac{1}{2}\mathrm{e}^{-t} + \frac{\beta^2}{144}\left(36 + 2\mathrm{e}^{-4t} - (29+12t)\mathrm{e}^{-t}\right) + \mathcal{O}(\beta^4). \tag{B.21}$$

In particular, comparing with eq. (5.30), we see that

$$\left\langle (\Delta\mathfrak{s}_i^{\mathrm{res}})^2 \right\rangle = \beta^2 \left\langle (u(\tau) - u(-\tau))^2 \right\rangle \tag{B.22}$$

$$= \beta^2\left(1 - \mathrm{e}^{-2\tau}\right) + \frac{\beta^4}{72}\left((29+24\tau)\,\mathrm{e}^{-2\tau} - 2\mathrm{e}^{-8\tau} - 27\right) + \mathcal{O}\left(\beta^6\right) \tag{B.23}$$

as promised. Note also that by employing the functions U_n for $n \geq 1$, we can in principle obtain all moments of $\Delta\mathfrak{s}_i^{\mathrm{res}}$, so that this approach extends beyond the variance computed here.

Next, we proceed to calculate the correlation function $C(t)$ and effective diffusion coefficient $D_{\mathrm{eff}}(\beta)$. To this end, we introduce a second set of autocorrelation moments, namely

$$V_n(t) = \langle \boldsymbol{e}(0)\cdot\boldsymbol{e}(t)u(t)^n \rangle\,. \tag{B.24}$$

Again, one may use standard Itô calculus to show that by differentiating V_n with respect to time, we may express

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}V_n(t) = {}& \langle e_\alpha(0)u(t)^n\dot{e}_\alpha(t)\rangle + n\left\langle e_\alpha(0)u(t)^{n-1}\left[\dot{u}(t),\dot{e}_\alpha(t)\right]\right\rangle \\ & + n\left\langle \boldsymbol{e}(0)\cdot\boldsymbol{e}(t)u(t)^{n-1}\dot{u}(t)\right\rangle + \frac{1}{2}n(n-1)\left\langle \boldsymbol{e}(0)\cdot\boldsymbol{e}(t)u(t)^{n-2}\left[\dot{u}(t),\dot{u}(t)\right]\right\rangle.\end{aligned} \tag{B.25}$$

Moreover, from eqs. (B.2) and (B.5) we may compute the quadratic covariation

$$[\dot{u},\dot{e}_\alpha] = 2\left(\hat{M}_\alpha - ue_\alpha\right), \tag{B.26}$$

from which we conclude that

$$\frac{\mathrm{d}}{\mathrm{d}t}V_n = -(n+1)^2V_n + \beta\left(U_n - V_{n+1}\right) + n\beta\left(V_{n-1} - V_{n+1}\right) + 2nU_{n-1} + n(n-1)V_{n-2}. \tag{B.27}$$

Assuming knowledge of the functions U_n, we will now show that the solution to the recursive equation for V_n can also be obtained as a series expansion in small β. Indeed, as

before, we look for a solution of the form

$$V_n = \sum_{k=0}^{\infty} V_n^k \beta^k \tag{B.28}$$

and aim to determine the coefficients $V_n^k(t)$. Combining then the recursive equation in eq. (B.27) with the expansions in eqs. (B.8) and (B.28) and collecting terms at equal order in β we thus obtain the hierarchy

$$\frac{\mathrm{d}}{\mathrm{d}t}V_n^0 = -(n+1)^2 V_n^0 + 2nU_{n-1}^0 + n(n-1)V_{n-2}^0, \tag{B.29}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}V_n^k = -(n+1)^2 V_n^k + U_n^{k-1} - V_{n+1}^{k-1} + n\left(V_{n-1}^{k-1} - V_{n+1}^{k-1}\right) + 2nU_{n-1}^k + n(n-1)V_{n-2}^k \quad k \geq 1 \tag{B.30}$$

Again, a solution can be found by first solving eq. (B.29) for V_n^0 to the required order in n, and subsequently eq. (B.30) to obtain V_n to the required order in β. The initial values for the system in eqs. (B.29) and (B.30) are straightforwardly found by observing that $V_0(0) = 1$ while

$$V_n(0) = U_{n-1}(0) \quad n \geq 1, \tag{B.31}$$

which again can be read off from eqs. (B.17) to (B.20). Thus, one may show that

$$V_0(t) = \mathrm{e}^{-t} + \frac{\beta^2}{36}\left(9 - 2\left(5+3t\right)\mathrm{e}^{-t} + \mathrm{e}^{-4t}\right) + \mathcal{O}(\beta^4). \tag{B.32}$$

Finally, combining our results obtained thus far, we may deduce the autocorrelation function $C(t)$ in eq. (3.35) and effective diffusion coefficient $D_{\mathrm{eff}}(\beta)$ in eq. (3.37). First, using the expansion for $\langle \boldsymbol{e}(0)\rangle = (I_1(\beta)/I_0(\beta))\hat{\boldsymbol{M}}$ in eq. (3.34), in addition to eq. (B.32), we find that

$$C(t) = V_0(t) - \langle \boldsymbol{e}(0)\rangle^2 \tag{B.33}$$

$$= \mathrm{e}^{-t}\left(1 + \frac{\beta^2}{36}\left(\mathrm{e}^{-3t} - 10 - 6t\right) + \mathcal{O}\left(\beta^4\right)\right), \tag{B.34}$$

as stated in eq. (3.35). Similarly, the effective diffusion in eq. (3.37) can be computed from

$$D_{\mathrm{eff}}(\beta) = \lim_{\tau\to\infty}\frac{1}{2\tau}\left\langle\left(\Delta\boldsymbol{x} - \langle\Delta\boldsymbol{x}\rangle\right)^2\right\rangle \tag{B.35}$$

$$= \lim_{\tau\to\infty}\frac{1}{2\tau}\left(\int_0^\tau \mathrm{d}t\int_0^\tau \mathrm{d}s\, C(|t-s|) - \left(\frac{I_1(\beta)}{I_0(\beta)}\tau\right)^2\right), \tag{B.36}$$

where $\Delta \boldsymbol{x} = \int_0^\tau \mathrm{d}t\, \boldsymbol{e}$. Indeed, using eq. (B.34), the integral can be computed straightforwardly, and one finds that

$$D_{\mathrm{eff}}(\beta) = 1 - \frac{7}{16}\beta^2 + \mathcal{O}(\beta^4) \tag{B.37}$$

as required.

Appendix C

Linear analysis of the HVM and DFM

In this appendix we derive and summarize the results employed in chapter 6 from the linear theory of the HVM as well as the DFM. Specifically, we derive expressions for the correlators used when calculating the coefficient at $\mathcal{O}(D^0)$ of the entropy production rates $\dot{\mathcal{S}}$ and $\dot{\mathcal{S}}^{\pm}$ in chapter 6.

We begin by assuming that an expansion of the fields ρ and $\boldsymbol{m}$ in small D as in eqs. (6.21) and (6.22) is valid, and that the ground-state trajectory $(\rho_0, \boldsymbol{m}_0)$ is constant and homogeneous. Substituting this into the dynamics in eqs. (6.1) and (6.3) we obtain a hierarchy of equations by equating terms at $\mathcal{O}(D^\alpha)$, $\alpha = 0, \frac{1}{2}, 1, \ldots$. Since the continuity equation is linear, we obtain the trivial hierarchy

$$\partial_t \rho_n = -w\nabla \cdot \boldsymbol{m}_n, \qquad n \geq 0 \tag{C.1}$$

for the density coefficients ρ_n. The equation for $\boldsymbol{m}$ requires more work, however. At $\mathcal{O}(D^0)$ we find that

$$a_0 \boldsymbol{m}_0 = 0 \tag{C.2}$$

where $a_0 = 1 - \rho_0 + m^2$. Solving this equation for the magnetisation density gives the isotropic and polar liquid (PL) solutions $\boldsymbol{m}_0 = 0$ and $m^2 = \rho_0 - 1$ respectively. At higher orders, we find that

$$\partial_t \boldsymbol{m}_n + \lambda \boldsymbol{m}_0 \cdot \nabla \boldsymbol{m}_n = -\frac{\delta F_L}{\delta \boldsymbol{m}}[\rho_n, \boldsymbol{m}_n] + \boldsymbol{\Delta}_n^m\left(\{\rho_k, \boldsymbol{m}_k, \nabla\rho_k, \nabla\boldsymbol{m}_k, \ldots\}_{k<n}\right), \quad n \geq 1. \tag{C.3}$$

Here, F_L is the quadratic functional defined in eq. (6.25). The driving term $\boldsymbol{\Delta}_n^m$ at each order $n \geq 2$ must be derived explicitly for each case, although it depends only on the fields ρ_k, $\boldsymbol{m}_k$ (and their spatial derivatives) for $k < n$. For $n = 1$, $\boldsymbol{\Delta}_1^m = \boldsymbol{\eta}_1$ is a mean zero Gaussian white

noise process with covariance

$$\langle \eta_{1\alpha}(\boldsymbol{x},t)\eta_{1\beta}(\boldsymbol{x}',t')\rangle = 2\delta_{\alpha\beta}\delta(\boldsymbol{x}-\boldsymbol{x}')\delta(t-t'), \tag{C.4}$$

and in particular does not depend on D. At first nontrivial order, i.e. $n=2$, the driving term is given explicitly by

$$\boldsymbol{\Delta}_2^m = -\lambda \boldsymbol{m}_1\cdot\nabla\boldsymbol{m}_1 + (\rho_1 - 2\boldsymbol{m}_0\cdot\boldsymbol{m}_1)\boldsymbol{m}_1 - |\boldsymbol{m}_1|^2\boldsymbol{m}_0 + \frac{\kappa}{2}\nabla|\boldsymbol{m}_1|^2 - \kappa\boldsymbol{m}_1\nabla\cdot\boldsymbol{m}_1. \tag{C.5}$$

In particular, when $\boldsymbol{m}_0 = 0$, we know from chapter 6 that $(\rho_1,\boldsymbol{m}_1)$ is in fact an equilibrium dynamics, albeit with Fourier modes driven by heat baths at different temperatures. The coupling of these via eq. (C.5) consequently drives the next order process $(\rho_2,\boldsymbol{m}_2)$ out of equilibrium. Moreover, since eq. (C.3) is inhomogeneous and linear in ρ_n, $\boldsymbol{m}_n$, the system in eqs. (C.1) to (C.3) may in principle be solved recursively to arbitrary order. Thus, we may think of the higher order driving terms $\boldsymbol{\Delta}_n^m$ in a similar vein. Despite this, our analysis here will be restricted to $n \leq 1$.

The situation is quite similar for the DFM, although the continuity equation is no longer linear and the hierarchy in eq. (C.1) changes accordingly. Specifically, we find that for the DFM

$$\partial_t\rho_n = -\nabla\cdot\left(w\boldsymbol{m}_n - \gamma^{-1}\nabla\frac{\delta F_L}{\delta\rho}[\rho_n,\boldsymbol{m}_n]\right) + \Delta_n^\rho, \qquad n\geq 1, \tag{C.6}$$

where F_L is now given in eq. (6.83). Note that even though F_L is modified slightly for the DFM, the hierarchy of equations in eq. (C.3) is in fact unchanged since $\delta F_L/\delta\boldsymbol{m}$ remains the same. Again, the driving term $\Delta_n^\rho \equiv \Delta_n^\rho(\{\rho_k,\boldsymbol{m}_k,\ldots\}_{k<n})$ is in general a nonlinear function of ρ_k, $\boldsymbol{m}_k$ and their gradients for $k<n$. For $n=1$, $\Delta_1^\rho = -\nabla\cdot\boldsymbol{\xi}_1$, where $\boldsymbol{\xi}_1$ is a mean zero Gaussian white noise process with covariance

$$\langle \xi_{1\alpha}(\boldsymbol{x},t)\xi_{1\beta}(\boldsymbol{x}',t')\rangle = 2\gamma^{-1}\delta_{\alpha\beta}\delta(\boldsymbol{x}-\boldsymbol{x}')\delta(t-t'). \tag{C.7}$$

In the following, we wish to examine both the linear stability of the isotropic state of the HVM and DFM. Moreover, we will also deduce expressions for the correlators in the linearised theory of the HVM and DFM in both the isotropic and polar liquid (PL) phases. In particular, since we performed the linear stability analysis of the PL state of the HVM in section 4.6, it will not feature in this appendix. For the DFM, however, the linear stability analysis of the PL state becomes too algebraically involved to obtain exact analytical results, and so we will not attempt to pursue this here.

Further, the calculations we perform will be presented in two parts by treating first both the linear stability of and correlators in the isotropic state with $\boldsymbol{m}_0 = 0$, and subsequently the

correlators in the PL phase with $m_0^2 = \rho_0 - 1$. In both cases, we consider the more general DFM and observe that predictions for the HVM may be made by considering the limit $\gamma \to \infty$.

Beginning with the isotropic state, we transform the linearised equations of motion in eqs. (C.3) and (C.6) for the DFM to Fourier space when $n = 1$. The result is most conveniently expressed in matrix form as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} \rho_{1,\boldsymbol{q}} \\ m_{L,\boldsymbol{q}} \\ m_{T,\boldsymbol{q}} \end{pmatrix} = -\underbrace{\begin{pmatrix} q^2\Gamma_\rho/\gamma & i\tilde{w}q & 0 \\ iw_1 q & \Gamma & 0 \\ 0 & 0 & \Gamma \end{pmatrix}}_{\mathcal{L}_{\boldsymbol{q}}^{\mathrm{iso}}} \begin{pmatrix} \rho_{1,\boldsymbol{q}} \\ m_{L,\boldsymbol{q}} \\ m_{T,\boldsymbol{q}} \end{pmatrix} + \begin{pmatrix} -iq\xi_{L,\boldsymbol{q}} \\ \eta_{L,\boldsymbol{q}} \\ \eta_{T,\boldsymbol{q}} \end{pmatrix}. \tag{C.8}$$

Here, $\xi_{L,\boldsymbol{q}} = \hat{\boldsymbol{q}} \cdot \boldsymbol{\xi}_{1,\boldsymbol{q}}$ is the longitudinal component of the Fourier coefficient $\boldsymbol{\xi}_{1,\boldsymbol{q}}$, while $\eta_{L,\boldsymbol{q}} = \hat{\boldsymbol{q}} \cdot \boldsymbol{\eta}_{1,\boldsymbol{q}}$, $\eta_{T,\boldsymbol{q}} = \hat{\boldsymbol{q}}_\perp \cdot \boldsymbol{\eta}_{1,\boldsymbol{q}}$ are the longitudinal and transverse components of $\boldsymbol{\eta}_{1,\boldsymbol{q}}$ respectively, where $\hat{\boldsymbol{q}}_\perp$ is perpendicular to $\hat{\boldsymbol{q}}$. Because of the convention in eq. (6.36) we have chosen for Fourier transforms, noise correlations in Fourier space contain an extra factor of $\mathcal{V}^{-1}$, i.e.

$$\langle \xi_{L,\boldsymbol{q}}(t)\xi^*_{L,\boldsymbol{q}'}(t')\rangle = 2(\gamma\mathcal{V})^{-1}\delta_{\boldsymbol{q},\boldsymbol{q}'}\delta(t-t'), \tag{C.9}$$

$$\langle \eta_{L,\boldsymbol{q}}(t)\eta^*_{L,\boldsymbol{q}'}(t')\rangle = \langle \eta_{T,\boldsymbol{q}}(t)\eta^*_{T,\boldsymbol{q}'}(t')\rangle = 2\mathcal{V}^{-1}\delta_{\boldsymbol{q},\boldsymbol{q}'}\delta(t-t'), \tag{C.10}$$

while $\langle \eta_{L,\boldsymbol{q}}(t)\eta^*_{T,\boldsymbol{q}'}(t')\rangle = \langle \eta_{L,\boldsymbol{q}}(t)\xi^*_{L,\boldsymbol{q}'}(t')\rangle = \langle \eta_{T,\boldsymbol{q}}(t)\xi^*_{L,\boldsymbol{q}'}(t')\rangle = 0$. Note also that the damping coefficient $\Gamma(q) = 1 - \rho_0 + q^2 > 0$ ensures that transverse fluctuations $\boldsymbol{m}_{T,\boldsymbol{q}}$ decay on non-hydrodynamic timescales when $\rho_0 < 1$.

Observe that solutions to eq. (C.8) are stable and decay at an exponential rate when the eigenvalues $\sigma_\pm$, σ_T of the linear operator $\mathcal{L}_{\boldsymbol{q}}^{\mathrm{iso}}$ have positive real parts. These are straightforwardly found from the characteristic equation of $\mathcal{L}_{\boldsymbol{q}}^{\mathrm{iso}}$, and are given by

$$\sigma_\pm = \frac{1}{2}\left(\Gamma + \gamma^{-1}q^2\Gamma_\rho \pm \sqrt{(\Gamma + \gamma^{-1}q^2\Gamma_\rho)^2 - 4q^2(\tilde{w}w_1 + \gamma^{-1}\Gamma_\rho\Gamma)}\right), \tag{C.11}$$

$$\sigma_T = \Gamma. \tag{C.12}$$

Equation (C.11) for $\sigma_\pm$ may be unraveled by first understanding its behaviour at $\gamma = \infty$. In this limit, we find that

$$\sigma_\pm(\gamma \to \infty) = \frac{1}{2}\left(\Gamma \pm \sqrt{\Gamma^2 - 4q^2 w w_1}\right). \tag{C.13}$$

Since $w, w_1 > 0$ it follows that Re $\sigma_\pm(\gamma \to \infty) > 0$ if and only if $\sigma_T = \Gamma > 0$. Thus, at $\gamma = \infty$, the isotropic state is linearly stable when $\rho_0 < 1$, while it is unstable otherwise.

From this, it is fairly straightforward to see that a similar condition holds for finite γ. Specifically, Re $\sigma_\pm > 0$ only when $\rho_0 < 1$. However, we see from eq. (C.11) that in this case stability also requires that

$$\tilde{w} w_1 + \gamma^{-1} \Gamma_\rho \Gamma > 0. \tag{C.14}$$

After some algebra, one finds that this latter condition holds for all q if and only if

$$w_1^2 < a_\rho + \nu_\rho(1-\rho_0) + 4\gamma\nu_\rho + 2\sqrt{\nu_\rho(1-\rho_0+2\gamma)(a_\rho + 2\gamma\nu_\rho)}. \tag{C.15}$$

Thus, the phase diagram of the DFM in the region where $\rho_0 < 1$ as predicted by the linear theory is no longer trivial. Indeed, when condition eq. (C.15) is broken, there is a finite range of wave numbers $q \in [q_-, q_+]$ for which the corresponding modes $\rho_{1,\boldsymbol{q}}$, $m_{L,\boldsymbol{q}}$ grow in time, where

$$2\nu_\rho q_\pm^2 = w_1^2 - \alpha_\rho - \nu_\rho(1-\rho_0) \pm \sqrt{\left(w_1^2 - \alpha_\rho - \nu_\rho(1-\rho_0)\right)^2 - 4\nu_\rho\left(\gamma w w_1 + \alpha_\rho(1-\rho_0)\right)}. \tag{C.16}$$

From simulations, we find that this instability leads to the polar crystal phase reported in section 6.5.

We are also interested in calculating the equal-time correlation functions of the linear theory in order to make analytical predictions about the entropy production rate (EPR). Since the dynamics is linear, the correlation matrix C^{iso} in eq. (6.91) solves the algebraic Riccati equation

$$\mathcal{L}_{\boldsymbol{q}}^{\text{iso}} C^{\text{iso}} + C^{\text{iso}} \left(\mathcal{L}_{\boldsymbol{q}}^{\text{iso}}\right)^\dagger = 2\mathcal{D}, \tag{C.17}$$

where we have defined the diffusion matrix $\mathcal{D}$ by

$$\mathcal{D} = \frac{1}{\mathcal{V}} \begin{pmatrix} q^2/\gamma & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{C.18}$$

To derive this, one simply needs to apply the chain rule to the left-hand side of

$$\left\langle \frac{\mathrm{d}}{\mathrm{d}t}\left(u_{i,\boldsymbol{q}}^{\text{iso}} \left(u_{j,\boldsymbol{q}'}^{\text{iso}}\right)^*\right)\right\rangle = 0, \tag{C.19}$$

where $\boldsymbol{u}^{\text{iso}}$ is defined in eq. (6.49). It is well known that the solution to the Riccati equation eq. (C.17) may be expressed in integral form. However, we find that for our present purposes

it is less cumbersome to tackle it straight on. First we observe that the linear system in eq. (C.8) reduces to the two-dimensional coupled dynamics of $(\rho_{1,\boldsymbol{q}}, m_{L,\boldsymbol{q}})$, in addition to the one-dimensional dynamics of $m_{T,\boldsymbol{q}}$. Thus, clearly, we may treat these separately. Beginning with the former, the Riccati equation eq. (C.17) may be re-expressed as a four-by-four linear system, specifically

$$\underbrace{\begin{pmatrix} 2q^2\Gamma_\rho/\gamma & -i\tilde{w}q & i\tilde{w}q & 0 \\ -iw_1 q & \gamma^{-1}q^2\Gamma_\rho+\Gamma & 0 & i\tilde{w}q \\ iw_1 q & 0 & \gamma^{-1}q^2\Gamma_\rho+\Gamma & -i\tilde{w}q \\ 0 & iw_1 q & -iw_1 q & 2\Gamma \end{pmatrix}}_{R^{\mathrm{iso}}_{\boldsymbol{q}}} \begin{pmatrix} C^{\mathrm{iso}}_{11} \\ C^{\mathrm{iso}}_{12} \\ C^{\mathrm{iso}}_{21} \\ C^{\mathrm{iso}}_{22} \end{pmatrix} = \frac{2}{\mathcal{V}} \begin{pmatrix} q^2/\gamma \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{C.20}$$

To find the correlators of the linear theory we therefore simply invert the matrix $R^{\mathrm{iso}}_{\boldsymbol{q}}$, and one may check that the solution is given by

$$C^{\mathrm{iso}}_{11} = \frac{1}{\mathcal{V}} \frac{\tilde{w}w + \gamma^{-1}\Gamma(\Gamma+\gamma^{-1}q^2\Gamma_\rho)}{(\Gamma+\gamma^{-1}q^2\Gamma_\rho)(\tilde{w}w_1+\gamma^{-1}\Gamma_\rho\Gamma)}, \tag{C.21}$$

$$C^{\mathrm{iso}}_{12} = (C^{\mathrm{iso}}_{21})^* = \frac{1}{\mathcal{V}} \frac{i\gamma^{-1}q(w_1\Gamma - \tilde{w}\Gamma_\rho)}{(\Gamma+\gamma^{-1}q^2\Gamma_\rho)(\tilde{w}w_1+\gamma^{-1}\Gamma_\rho\Gamma)}, \tag{C.22}$$

$$C^{\mathrm{iso}}_{22} = \frac{1}{\mathcal{V}} \frac{ww_1 + \gamma^{-1}\Gamma_\rho(\Gamma+\gamma^{-1}q^2\Gamma_\rho)}{(\Gamma+\gamma^{-1}q^2\Gamma_\rho)(\tilde{w}w_1+\gamma^{-1}\Gamma_\rho\Gamma)}. \tag{C.23}$$

The final non-trivial component of C^{iso} is found straightforwardly from the dynamics of $m_{T,\boldsymbol{q}}$, and is given by

$$C^{\mathrm{iso}}_{33} = \frac{1}{\mathcal{V}\Gamma}. \tag{C.24}$$

To recover the correlators in the linearised HVM, we simply take the limit $\gamma \to \infty$ in eqs. (C.21) to (C.24). We obtain that

$$C^{\mathrm{iso}}(q, \gamma\to\infty) = \frac{1}{\mathcal{V}\Gamma} \begin{pmatrix} w/w_1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{C.25}$$

In particular, this result verifies that $\langle \rho_{1,\boldsymbol{q}} m^*_{L,\boldsymbol{q}} \rangle = 0$ as advertised in eq. (6.33).

The linearised dynamics about the homogeneous PL phase may be treated similarly to the isotropic case considered above. Transforming eqs. (C.3) and (C.6) for $n = 1$ and $m^2 = \rho_0 - 1$

Table C.1 Leading order asymptotic expressions for the components C_{ij}^{pl} of the equal-time correlation matrix C^{pl} in the limit where $q \to \infty$, as well as when $\gamma \to \infty$ is taken first.

C_{ij}^{pl}	$\gamma \to \infty$, $q \to \infty$	$q \to \infty$
C_{11}^{pl}	$w/(\mathcal{V} w_1 q^2)$	$1/(\mathcal{V} \nu_\rho q^2)$
C_{12}^{pl}	$m_0 w(\cos^2\theta + (1+\kappa w_1)\sin^2\theta)/(\mathcal{V} w_1 q^4)$	$i w_1 \cos\theta/(\mathcal{V}\nu_\rho q^3)$
C_{13}^{pl}	$-m_0 \kappa w \cos\theta \sin\theta/(\mathcal{V} q^4)$	$i w_1 \sin\theta/(\mathcal{V}\nu_\rho q^3)$
C_{22}^{pl}	$1/(\mathcal{V} q^2)$	$1/(\mathcal{V} q^2)$
C_{23}^{pl}	$-i m_0 \kappa \sin\theta/(\mathcal{V} q^3)$	$-i m_0 \kappa \sin\theta/(\mathcal{V} q^3)$
C_{33}^{pl}	$1/(\mathcal{V} q^2)$	$1/(\mathcal{V} q^2)$

to Fourier space we find that

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} \rho_{1,\boldsymbol{q}} \\ m_{\parallel,\boldsymbol{q}} \\ m_{\perp,\boldsymbol{q}} \end{pmatrix} = -\mathcal{L}_{\boldsymbol{q}}^{\mathrm{pl}}\begin{pmatrix} \rho_{1,\boldsymbol{q}} \\ m_{\parallel,\boldsymbol{q}} \\ m_{\perp,\boldsymbol{q}} \end{pmatrix} + \begin{pmatrix} -iq\xi_{L,\boldsymbol{q}} \\ \hat{\eta}_{\parallel} \\ \hat{\eta}_{\perp} \end{pmatrix}, \tag{C.26}$$

where we have defined

$$\mathcal{L}_{\boldsymbol{q}}^{\mathrm{pl}} = \begin{pmatrix} q^2\Gamma_\rho/\gamma & i\tilde{w}q_{\parallel} - q^2 m_0/\gamma & i\tilde{w}q_{\perp} \\ i w_1 q_{\parallel} - m_0 & \Gamma_{\parallel} + i\lambda m_0 q_{\parallel} & i\kappa m_0 q_{\perp} \\ i w_1 q_{\perp} & -i\kappa m_0 q_{\perp} & \Gamma_{\perp} + i\lambda m_0 q_{\parallel} \end{pmatrix}, \tag{C.27}$$

in addition to new damping coefficients

$$\Gamma_{\parallel} = 2(\rho_0 - 1) + q^2, \tag{C.28}$$

$$\Gamma_{\perp} = q^2. \tag{C.29}$$

As before, $\xi_{L,\boldsymbol{q}} = \hat{\boldsymbol{q}} \cdot \boldsymbol{\xi}_{\boldsymbol{q}}$ is the longitudinal component of the noise $\boldsymbol{\xi}_{\boldsymbol{q}}$, while $m_{\parallel,\boldsymbol{q}}$ ($\eta_{\parallel,\boldsymbol{q}}$) and $m_{\perp,\boldsymbol{q}}$ ($\eta_{\perp,\boldsymbol{q}}$) are respectively the parallel and perpendicular components of $\boldsymbol{m}_{\boldsymbol{q}}$ ($\boldsymbol{\eta}_1$) with respect to $\boldsymbol{m}_0$. Furthermore, all noise terms $\xi_{L,\boldsymbol{q}}$, $\hat{\eta}_{\parallel,\boldsymbol{q}}$ and $\hat{\eta}_{\perp,\boldsymbol{q}}$ are independent and of the same covariance as in the isotropic case.

From this, we may investigate the structure of the correlators in the linearised theory of the HVM and DFM about the PL state. In analogy with the isotropic calculation, the matrix C^{pl} in eq. (6.50) solves an algebraic Riccati equation as in eq. (C.17), although in this case all three modes $(\rho_{1,\boldsymbol{q}}, m_{\parallel,\boldsymbol{q}}, m_{\perp,\boldsymbol{q}})$ remain coupled for general $\boldsymbol{q}$. Finding its solution can be

achieved by solving the linear system

$$R_{\boldsymbol{q}}^{\mathrm{pl}} \boldsymbol{C}^{\mathrm{pl}} = 2\boldsymbol{D}, \tag{C.30}$$

where $\boldsymbol{C}^{\mathrm{pl}} = (C_{11}^{\mathrm{pl}}, C_{12}^{\mathrm{pl}}, \ldots, C_{33}^{\mathrm{pl}})^T$, $\boldsymbol{D} = (\mathcal{D}_{11}, \mathcal{D}_{12}, \ldots, \mathcal{D}_{33})^T$. We avoid explicitly writing out the nine-by-nine matrix $R_{\boldsymbol{q}}^{\mathrm{pl}}$ here for sake of clarity of presentation, although it may be found fairly straightforwardly from the Riccati equation. Analytical inversion of eq. (C.30) may be done using standard computer algebra systems that perform symbolic computations. Due to the algebraic complexity of the resulting expressions, we choose to only state the result in certain limiting cases. More specifically, we look at the asymptotic behaviour of the components C_{ij}^{pl} in the limit $q \to \infty$ at fixed $q_{\parallel}/q_{\perp}$, as well as when $\gamma \to \infty$ is taken first. From this, we deduce in appendix D the dependence of the EPR on the ultraviolet cutoff Λ quoted in the main text. For the six independent components of C^{pl} we summarize the results in table C.

Appendix D

Calculation of the EPR in the HVM and DFM

Here we derive explicitly the expression in eq. (6.14) for the entropy production rate $\dot{\mathcal{S}}$ of the HVM, as well as eq. (6.71) and eq. (6.72) for $\dot{\mathcal{S}}^{\pm}$ of the DFM. In addition, we deduce the small noise expansion of the EPRs about the constant homogeneous isotropic and polar liquid (PL) ground-states quoted in the main text. For clarity, we choose to consider the HVM and the DFM separately. In particular, in contrast with our treatment above in appendix C, the results obtained in the former model cannot in general be computed from the latter by sending $\gamma \to \infty$.

Hydrodynamic Vicsek model

Starting from the definitions eq. (6.13) for the EPR $\dot{\mathcal{S}}$ and eq. (6.9) for the path transition probability density $\mathcal{P}$ of the HVM, one finds that we may equivalently express $\dot{\mathcal{S}}$ in terms of the T-antisymmetric part of the Freidlin-Wentzell action $\mathcal{A}$. Specifically,

$$\dot{\mathcal{S}} = \lim_{\tau\to\infty} \frac{\overleftarrow{\mathcal{A}} - \mathcal{A}}{2D\tau}, \tag{D.1}$$

where $\overleftarrow{\mathcal{A}}[\rho, \boldsymbol{m}] = \mathcal{A}[\overleftarrow{\rho}, \overleftarrow{\boldsymbol{m}}]$, $\overleftarrow{\rho} = \mathrm{T}\rho$ and $\overleftarrow{\boldsymbol{m}} = \mathrm{T}\boldsymbol{m}$ as before and the time-reversal operation T is defined as in eq. (6.10). By applying T to $\mathcal{A}$ as given in eq. (6.8), we find that the action $\overleftarrow{\mathcal{A}}$ for the time-reversed ensemble may be expressed explicitly by

$$\overleftarrow{\mathcal{A}} = \frac{1}{4}\int_{-\tau}^{\tau} \mathrm{d}t \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \left| \partial_t \boldsymbol{m} + \lambda \boldsymbol{m}\cdot\nabla\boldsymbol{m} - \frac{\delta F^S}{\delta \boldsymbol{m}} + \frac{\delta F^A}{\delta \boldsymbol{m}} \right|^2 \quad \text{if} \quad \partial_t \rho + w\nabla\cdot\boldsymbol{m} = 0, \tag{D.2}$$

and $\overleftarrow{\mathcal{A}} = \infty$ otherwise, where $F = F^S + F^A$ and F^A is defined in eq. (6.12). In particular, from eq. (D.1) one deduces that

$$\dot{\mathcal{S}} = \lim_{\tau\to\infty} \frac{-1}{2D\tau} \int_{-\tau}^{\tau} \mathrm{d}t \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \left(\partial_t \boldsymbol{m} + \lambda \boldsymbol{m} \cdot \nabla \boldsymbol{m} + \frac{\delta F^A}{\delta \boldsymbol{m}} \right) \circ \frac{\delta F^S}{\delta \boldsymbol{m}}. \tag{D.3}$$

Note that in eq. (D.3) we indicate as usual by $\circ$ that the product with $\partial_t \boldsymbol{m}$ should be interpreted in the Stratonovich sense (see chapter 2).

To transform eq. (D.3) into the expression given in eq. (6.14), we simply have to observe that some terms in eq. (D.3) are integrable. Specifically, we find that we may write

$$\dot{\mathcal{S}} = \lim_{\tau\to\infty} \left[\frac{-\Delta\mathcal{I}}{2D\tau} + \frac{1}{2D\tau} \int_{-\tau}^{\tau} \mathrm{d}t \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \left(\rho \boldsymbol{m} \circ \partial_t \boldsymbol{m} - \left(\lambda \boldsymbol{m} \cdot \nabla \boldsymbol{m} + \frac{\delta F^A}{\delta \boldsymbol{m}} \right) \cdot \frac{\delta F^S}{\delta \boldsymbol{m}} \right) \right], \tag{D.4}$$

where we have defined $\mathcal{I}$ to be the functional

$$\mathcal{I}[\boldsymbol{m}] = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \left(\frac{1}{2} m^2 + \frac{1}{2} (\nabla_\alpha m_\beta)^2 + \frac{1}{4} m^4 \right), \tag{D.5}$$

i.e. as the part of F^S which does not explicitly depend on the density ρ. Since we assume that the moments of $\boldsymbol{m}$ and its higher order spatial derivatives are finite in steady-state, it follows that $\Delta\mathcal{I}/\tau \to 0$ as $\tau \to \infty$ so that the first term in eq. (D.4) may safely be ignored. In a similar vein, an integration by parts allows us to write the integral over the first term appearing in the integrand in eq. (D.4) as

$$\int_{-\tau}^{\tau} \mathrm{d}t \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \rho \circ \partial_t m^2 = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \rho m^2 \Big|_{-\tau}^{\tau} + w \int_{-\tau}^{\tau} \mathrm{d}t \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, (\nabla \cdot \boldsymbol{m}) m^2, \tag{D.6}$$

where we have used the continuity equation $\partial_t \rho = -w \nabla \cdot \boldsymbol{m}$. Again, after dividing by $4D\tau$, the first term on the right-hand side of eq. (D.6) goes away in the limit $\tau \to \infty$ since it grows sublinearly in τ. Thus, assuming that we may replace temporal averages by averages over noise realizations, we finally arrive at the expression

$$\dot{\mathcal{S}} = D^{-1} \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \left\langle \frac{w}{2} m^2 (\nabla \cdot \boldsymbol{m}) - \left(\lambda \boldsymbol{m} \cdot \nabla \boldsymbol{m} + \frac{\delta F^A}{\delta \boldsymbol{m}} \right) \cdot \frac{\delta F^S}{\delta \boldsymbol{m}} \right\rangle \tag{D.7}$$

reported in eq. (6.14).

Next, we investigate the expansion of eq. (D.7) about the constant homogeneous ground-states. To do this, we substitute the expansions eq. (6.21) and eq. (6.22) into eq. (D.7) and collect terms at equal order in D. It is fairly straightforward to check that there are no contributions to $\dot{\mathcal{S}}$ at orders D^{-1} or $D^{-1/2}$. At order D^0 we find after some fairly tedious

algebra that

$$\dot{\mathcal{S}}_0 = \boldsymbol{m}_0 \cdot \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \Big\langle w\,(\nabla\cdot\boldsymbol{m}_1)\,\boldsymbol{m}_1 + \rho_1\left((2w_1-\kappa)\nabla\,(\boldsymbol{m}_0\cdot\boldsymbol{m}_1) + \lambda\,(\boldsymbol{m}_0\cdot\nabla)\,\boldsymbol{m}_1\right) + 2\kappa\,(\nabla\cdot\boldsymbol{m}_1)\left(\nabla^2\boldsymbol{m}_1 - m_0^2\boldsymbol{m}_1\right)\Big\rangle \tag{D.8}$$

In order to arrive at this expression we have only assumed that $\boldsymbol{m}_0$ satisfies the zeroth order equation eq. (C.2), and used eq. (C.1) repeatedly. If we further assume that $\boldsymbol{m}_0$ is a polarised solution with $m_0 > 0$, we may write $\boldsymbol{m}_0 = (m_0, 0)$ without loss of generality, from which eq. (6.31) follows immediately after integrating out total derivatives.

After transforming eq. (D.8) to Fourier space, we obtain eq. (6.47) as stated in the main text. Using our results from appendix C, we may now investigate the scaling of this expression with $\Lambda \to \infty$. To determine this, we note that the scaling of the sum in this limit is determined by the corresponding integral in q-space, i.e.

$$\dot{\mathcal{S}}_0/\mathcal{V} = \sum_{q\leq\Lambda} \mathrm{Tr}(\dot{\sigma}^{\mathrm{pl}} C^{\mathrm{pl}}) \sim \frac{\mathcal{V}}{(2\pi)^2}\int_0^{2\pi}\mathrm{d}\theta\int_0^{\Lambda}\mathrm{d}q\,q\,\mathrm{Tr}(\dot{\sigma}^{\mathrm{pl}} C^{\mathrm{pl}}), \tag{D.9}$$

where $\dot{\sigma}^{\mathrm{pl}}$ is given by eq. (6.48). In particular, direct substitution from table C and subsequently performing the integrals over q and θ yields

$$\dot{\mathcal{S}}_0/\mathcal{V} \sim \frac{m_0^2\kappa^2}{4\pi}\Lambda^2. \tag{D.10}$$

Diffusive flocking model

In analogy with the above calculation, we may compute the EPRs $\dot{\mathcal{S}}^{\pm}$ of the DFM from the $\mathrm{T}^{\pm}$-antisymmetric part of the Freidlin-Wentzell action $\mathcal{A}_{\mathrm{DF}}$, specifically

$$\dot{\mathcal{S}}^{\pm} = \lim_{\tau\to\infty}\frac{\overleftarrow{\mathcal{A}}^{\pm}_{\mathrm{DF}} - \mathcal{A}_{\mathrm{DF}}}{2D\tau}, \tag{D.11}$$

where we have defined $\overleftarrow{\mathcal{A}}^{\pm}_{\mathrm{DF}} = \mathcal{A}_{\mathrm{DF}}[\mathrm{T}^{\pm}\rho, \mathrm{T}^{\pm}\boldsymbol{m}]$, and $\mathcal{A}_{\mathrm{DF}}$ and $\mathrm{T}^{\pm}$ are defined in eqs. (6.65) and (6.66) respectively. Writing out the actions for the time-reversed ensembles explicitly, we find that

$$\overleftarrow{\mathcal{A}}^{+}_{\mathrm{DF}} = \frac{1}{4}\int_{-\tau}^{\tau}\mathrm{d}t\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\left[\gamma\left|\nabla^{-1}\left(\partial_t\rho - \nabla\cdot\boldsymbol{J}_d\right)\right|^2 + \left|\partial_t\boldsymbol{m} - \lambda\boldsymbol{m}\cdot\nabla\boldsymbol{m} - \frac{\delta F}{\delta\boldsymbol{m}}\right|^2\right], \tag{D.12}$$

and

$$\overleftarrow{\mathcal{A}}^{-}_{\mathrm{DF}} = \frac{1}{4}\int_{-\tau}^{\tau}\mathrm{d}t\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\left[\gamma\left|\nabla^{-1}\left(\partial_t\rho-\nabla\cdot\boldsymbol{J}^S_d+\nabla\cdot\boldsymbol{J}^A_d\right)\right|^2 + \left|\partial_t\boldsymbol{m}+\lambda\boldsymbol{m}\cdot\nabla\boldsymbol{m}-\frac{\delta F^S}{\delta\boldsymbol{m}}+\frac{\delta F^A}{\delta\boldsymbol{m}}\right|^2\right]. \tag{D.13}$$

Thus, proceeding as above we straightforwardly find that

$$\dot{\mathcal{S}}^{+} = \lim_{\tau\to\infty}\left[\frac{-\Delta F}{2D\tau} + \frac{1}{2D\tau}\int_{-\tau}^{\tau}\mathrm{d}t\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\left(\gamma w(\nabla\cdot\boldsymbol{m})\nabla^{-2}\circ\partial_t\rho-\lambda\left[(\boldsymbol{m}\cdot\nabla)\boldsymbol{m}\right]\circ\partial_t\boldsymbol{m}\right)\right] \tag{D.14}$$

and

$$\dot{\mathcal{S}}^{-} = \lim_{\tau\to\infty}\left[\frac{-\Delta F^S}{2D\tau} + \frac{1}{2D\tau}\int_{-\tau}^{\tau}\mathrm{d}t\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\left(\boldsymbol{J}^A_d\cdot\nabla\mu^S-\left(\lambda\boldsymbol{m}\cdot\nabla\boldsymbol{m}+\frac{\delta F^A}{\delta\boldsymbol{m}}\right)\cdot\frac{\delta F^S}{\delta\boldsymbol{m}}\right)\right]. \tag{D.15}$$

The expression for $\dot{\mathcal{S}}^{-}$ immediately reduces to eq. (6.72) provided in the main text upon replacing the temporal average by an average over noise realizations and noting again that $\Delta F^S/\tau\to 0$ as $\tau\to\infty$. On the other hand, eq. (D.14) requires a bit more massaging. Firstly, we observe that the equal-time expectation

$$\left\langle(\nabla\cdot\boldsymbol{m})\nabla^{-2}\nabla\circ\boldsymbol{\xi}\right\rangle = 0, \tag{D.16}$$

since here the Stratonovich product coincides with the corresponding Itô product (i.e. there is no spurious drift). It therefore follows that

$$\left\langle(\nabla\cdot\boldsymbol{m})\nabla^{-2}\circ\partial_t\rho\right\rangle = -w\left\langle(\nabla\cdot\boldsymbol{m})\nabla^{-2}\nabla\cdot\boldsymbol{m}\right\rangle+\gamma^{-1}\left\langle(\nabla\cdot\boldsymbol{m})\mu\right\rangle. \tag{D.17}$$

To treat the product with $\partial_t\boldsymbol{m}$ in eq. (D.14) we must explicitly compute the spurious drift. We will show that, in fact, the spurious drift is a total derivative and therefore does not contribute to the EPR $\dot{\mathcal{S}}^{+}$. To do this, we consider a finite discretisation of the process in Fourier space with $q\leq\Lambda$, which is consistent with our numerical scheme. By applying standard stochastic

calculus, we then obtain

$$\begin{aligned}\int_{\mathcal{V}} \mathrm{d}\boldsymbol{x}\, \langle [(\boldsymbol{m}\cdot\nabla)\boldsymbol{m}]\circ\boldsymbol{\eta}\rangle &= \mathcal{V}\sum_{|\boldsymbol{q}|\leq\Lambda}\sum_{|\boldsymbol{k}|\leq\Lambda} ik_\beta \left\langle m_\beta(\boldsymbol{q}-\boldsymbol{k})m_\alpha(\boldsymbol{k})\circ\eta_\alpha(-\boldsymbol{q})\right\rangle \\ &= D\sum_{|\boldsymbol{q}|\leq\Lambda}\sum_{|\boldsymbol{k}|\leq\Lambda} ik_\beta \left\langle \left(\delta_{\boldsymbol{k},\boldsymbol{0}} m_\beta(\boldsymbol{k}) + 2\delta_{\boldsymbol{q},\boldsymbol{k}} m_\beta(\boldsymbol{q}-\boldsymbol{k})\right)\right\rangle \\ &= 0 \end{aligned} \tag{D.18}$$

Here, the second equality follows from the transformation rule between Ito and Stratonovich processes (see [59] or chapter 2), i.e.

$$\langle h(\boldsymbol{m}(\boldsymbol{q}_1),\ldots,\boldsymbol{m}(\boldsymbol{q}_n))\circ\eta_\alpha(\boldsymbol{k})\rangle = D\sum_{m=1}^{n}\left\langle \frac{\partial h}{\partial m_\alpha(\boldsymbol{q}_m)}\right\rangle \delta_{\boldsymbol{q}_m,-\boldsymbol{k}}, \tag{D.19}$$

and the fact that $\delta_{\alpha\alpha} = 2$. To see why the final equality holds, note also that

$$\sum_{k\leq\Lambda} k_\alpha = 0 \tag{D.20}$$

since the sum is finite. From this, and taking $\tau\to\infty$ in eq. (D.14), the result in eq. (6.71) reported in the main text follows immediately.

For the DFM, we may additionally calculate the EPRs $\dot{\mathcal{S}}_J^{\pm}$ at the level of the fluctuating density current $\boldsymbol{J}$. At this level, the actions $\overleftarrow{\mathcal{A}}_{\mathrm{DF}}^{J,\pm}$ for the time-reversed ensembles may be expressed as

$$\overleftarrow{\mathcal{A}}_{\mathrm{DF}}^{J,+} = \frac{1}{4}\int_{-\tau}^{\tau}\mathrm{d}t\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\left[\gamma|\boldsymbol{J}+\boldsymbol{J}_d|^2 + \left|\partial_t\boldsymbol{m} - \lambda\boldsymbol{m}\cdot\nabla\boldsymbol{m} - \frac{\delta F}{\delta\boldsymbol{m}}\right|^2\right] \tag{D.21}$$

and

$$\overleftarrow{\mathcal{A}}_{\mathrm{DF}}^{J,-} = \frac{1}{4}\int_{-\tau}^{\tau}\mathrm{d}t\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\left[\gamma\left|\boldsymbol{J}+\boldsymbol{J}_d^S-\boldsymbol{J}_d^A\right|^2 + \left|\partial_t\boldsymbol{m} + \lambda\boldsymbol{m}\cdot\nabla\boldsymbol{m} - \frac{\delta F^S}{\delta\boldsymbol{m}} + \frac{\delta F^A}{\delta\boldsymbol{m}}\right|^2\right] \tag{D.22}$$

if $\partial_t\rho + \nabla\cdot\boldsymbol{J} = 0$ and $\overleftarrow{\mathcal{A}}_{\mathrm{DF}}^{J,\pm} = \infty$ otherwise. By direct substitution we then find that

$$\dot{\mathcal{S}}_J^{+} = \frac{1}{2D\tau}\int_{-\tau}^{\tau}\mathrm{d}t\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\left(\gamma\boldsymbol{J}_d\circ\boldsymbol{J} - \left(\lambda\boldsymbol{m}\cdot\nabla\boldsymbol{m} + \frac{\delta F}{\delta\boldsymbol{m}}\right)\circ\partial_t\boldsymbol{m}\right), \tag{D.23}$$

Table D.1 Independent components of the Hermitian bilinear EPR coupling matrices $\dot{\sigma}^{\pm,\mathrm{pl}}$.

(i,j)	$\dot{\sigma}_{ij}^{+,\mathrm{pl}}$	$\dot{\sigma}_{ij}^{-,\mathrm{pl}}$
$(1,1)$	0	0
$(1,2)$	$\frac{i}{2}\left(w\Gamma_\rho - m_0^2\lambda - im_0\lambda w_1 q_\parallel\right) q_\parallel$	$\frac{i}{2}\left(w_1\Gamma_\parallel - \tilde{w}\Gamma_\rho + m_0^2\lambda\right) q_\parallel$
$(1,3)$	$\frac{i}{2}\left(w\Gamma_\rho - im_0\lambda w_1 q_\parallel\right) q_\perp$	$\frac{i}{2}\left(w_1\Gamma_\perp - \tilde{w}\Gamma_\rho + m_0^2\kappa\right) q_\perp$
$(2,2)$	$\left(\gamma w\tilde{w} + m_0^2\lambda^2 q^2\right) q_\parallel^2/q^2$	0
$(2,3)$	$-\frac{i}{2}w\left(m_0 q^2 + 2i\gamma\tilde{w}q_\parallel\right) q_\perp/q^2$	$\frac{i}{2}m_0\left(\tilde{w} - \kappa\left(\Gamma_\parallel + \Gamma_\perp\right)\right) q_\perp$
$(3,3)$	$\left(\gamma w\tilde{w} + m_0^2\lambda^2 q^2\left(\frac{q_\parallel}{q_\perp}\right)^2\right) q_\perp^2/q^2$	0

in addition to

$$\dot{\mathcal{S}}_J^- = \frac{1}{2D\tau}\int_{-\tau}^{\tau}\mathrm{d}t\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\left(\gamma\left(\boldsymbol{J}-\boldsymbol{J}_d^A\right)\circ\boldsymbol{J}_d^S - \left(\partial_t\boldsymbol{m} + \lambda\boldsymbol{m}\cdot\nabla\boldsymbol{m} + \frac{\delta F^A}{\delta\boldsymbol{m}}\right)\circ\frac{\delta F^S}{\delta\boldsymbol{m}}\right). \quad \text{(D.24)}$$

Now, it is fairly easy to see that

$$\begin{aligned}\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\,\langle\boldsymbol{J}_d\circ\boldsymbol{J}\rangle &= \int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\left\langle w\boldsymbol{m}\circ\boldsymbol{J} - \gamma^{-1}\boldsymbol{J}\cdot\nabla\mu\right\rangle \\ &= \int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\left\langle w^2m^2 - \gamma^{-1}w\boldsymbol{m}\cdot\nabla\mu - \gamma^{-1}\frac{\delta F}{\delta\rho}\circ\partial_t\rho\right\rangle, \end{aligned} \quad \text{(D.25)}$$

where the second equality follows from an integration by parts and the fact that $\langle\boldsymbol{m}\circ\boldsymbol{J}\rangle = \langle\boldsymbol{m}\cdot\boldsymbol{J}_d\rangle$. Substituting this back into eq. (D.23) gives the desired result for $\dot{\mathcal{S}}_J^+$, stated in eq. (6.76). Similarly, we have that

$$\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\left\langle\left(\boldsymbol{J}-\boldsymbol{J}_d^A\right)\circ\boldsymbol{J}_d^S\right\rangle = -\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\left\langle\gamma^{-1}\frac{\delta F^S}{\delta\rho}\circ\partial_t\rho + \boldsymbol{J}_d^A\cdot\boldsymbol{J}_d^S\right\rangle, \quad \text{(D.26)}$$

from which the fact that $\dot{\mathcal{S}}_J^- = \dot{\mathcal{S}}^-$ follows upon substitution back into eq. (D.24).

Again we may linearise the expressions in eqs. (6.71) and (6.72) about the homogeneous isotropic and PL states by substituting in an expansion of the form of eqs. (6.21) and (6.22).

Treating this as above for the HVM, we find that

$$\begin{aligned}\dot{\mathcal{S}}_0^+ = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \Big\langle & \gamma w^2 |K\boldsymbol{m}_1|^2 + w\left(\nabla\cdot\boldsymbol{m}_1\right)\left(a_\rho\rho_1 - \nu_\rho\nabla^2\rho_1\right) - m_0 w m_\parallel\left(\partial_\perp m_\perp\right) \\ & - w w_1\left(\nabla\cdot\boldsymbol{m}_1\right)^2 + m_0^2\lambda^2|\partial_\parallel\boldsymbol{m}_1|^2 - m_0^2\lambda\rho_1\left(\partial_\parallel m_\parallel\right) \\ & + m_0\lambda w_1\left(\partial_\parallel\rho_1\right)\left(\nabla\cdot\boldsymbol{m}_1\right)\Big\rangle, \end{aligned} \tag{D.27}$$

$$\begin{aligned}\dot{\mathcal{S}}_0^- = \int_{\mathcal{V}} \mathrm{d}\boldsymbol{x} \Big\langle & m_0 m_\parallel\partial_\perp\left(w m_\perp + \gamma^{-1}w_1\nabla^2 m_\perp\right) \\ & - \left(a_\rho\rho_1 - \nu_\rho\nabla^2\rho_1\right)\nabla\cdot\left(w\boldsymbol{m}_1 + \gamma^{-1}w_1\nabla^2\boldsymbol{m}_1\right) \\ & + m_0^2\left(2w_1 - \kappa + \lambda\right)\rho_1\partial_\parallel m_\parallel - 2m_0\kappa\left(\partial_\perp m_\perp\right)\left(m_0^2 m_\parallel - \nabla^2 m_\parallel\right)\Big\rangle, \end{aligned} \tag{D.28}$$

where both expressions hold for general $m_0 \geq 0$. We may equivalently express eqs. (D.27) and (D.28) in Fourier space, and for the constant homogeneous ground-states we obtain expressions analogous to those which we encountered for the HVM in eq. (6.47). Specifically, we find that for the isotropic and PL states,

$$\dot{\mathcal{S}}_0^\pm/\mathcal{V} = \sum_{|\boldsymbol{q}|\leq\Lambda} \mathrm{Tr}\left(\dot{\sigma}^{\pm,\mathrm{iso}} C^{\mathrm{iso}}\right), \tag{D.29}$$

and

$$\dot{\mathcal{S}}_0^\pm/\mathcal{V} = \sum_{|\boldsymbol{q}|\leq\Lambda} \mathrm{Tr}\left(\dot{\sigma}^{\pm,\mathrm{pl}} C^{\mathrm{pl}}\right), \tag{D.30}$$

respectively. Furthermore, we may choose to write the sum such that $\dot{\sigma}^{\pm,\mathrm{iso}}$ and $\dot{\sigma}^{\pm,\mathrm{pl}}$ are Hermitian. Taking $m_0 = 0$ in eqs. (D.27) and (D.28) and transforming to Fourier space we find that $\dot{\sigma}^{\pm,\mathrm{iso}}$ are given by eqs. (6.92) and (6.93) as advertised. For $\sigma^{\pm,\mathrm{pl}}$ we list the six independent components of each matrix in table D. Finally, from eq. (D.30) in addition to tables C and D, we deduce that in the PL phase, $\dot{\mathcal{S}}_0^+/\mathcal{V} \sim m_0^2\lambda^2\Lambda^2/(4\pi)$ and $\dot{\mathcal{S}}_0^-/\mathcal{V} \sim w_1^2\Lambda^4/(8\pi\gamma)$, while the exact results in the isotropic phase are presented in the main text.

Appendix E

Derivation of the Dean equation for the CTVM

In this appendix we include, for completeness, the derivation of the Dean equation (DE) for the continuous-time Vicsek model (CTVM) first mentioned in section 3.4 (see eq. (3.52)), and subsequently employed to study the asymptotic expansion of the entropy production rate (EPR) in high density of the CTVM in section 5.3 (see eq. (5.40)). Our derivation here closely follows that originally presented by Dean [107]. By now, this is considered a standard tool in the non-equilibrium statistical mechanics toolbox, and has been applied also to the NeAXYM previously [84]. Obtaining a DE from a Langevin dynamics can be done using an application of Itô's lemma, and importantly, is an equivalent representation of the dynamics.

We begin by recalling the CTVM Langevin equation from section 3.2 for the positions $\{\boldsymbol{x}_i\}$ and spins $\{\boldsymbol{e}_i\}$ in $d = 2$, specifically

$$\dot{\boldsymbol{x}}_i = \boldsymbol{e}_i, \tag{E.1}$$

$$\dot{\theta}_i = \beta \boldsymbol{e}_i' \cdot \hat{\boldsymbol{m}}_i^{\mathcal{C}} + \sqrt{2}\xi_i, \tag{E.2}$$

where i is the particle index, $\boldsymbol{e}_i \equiv \boldsymbol{e}(\theta_i) = (\cos\theta_i, \sin\theta_i)$, $\{\xi_i\}$ are unit Gaussian white noises and $\hat{\boldsymbol{m}}_i^{\mathcal{C}} = \boldsymbol{m}_i^{\mathcal{C}} / m_i^{\mathcal{C}}$ is the direction of the local magnetisation

$$\boldsymbol{m}_i^{\mathcal{C}} = \sum_{j=1}^{N} \mathcal{C}_{ij} \boldsymbol{e}_j. \tag{E.3}$$

As before, we consider here only isotropic contact kernels $\mathcal{C}_{ij} = \mathcal{C}_0(|\boldsymbol{x}_j - \boldsymbol{x}_i|)$.

To construct the DE, we consider some arbitrary smooth function $f(\boldsymbol{x},\theta)$. Applying Itô's lemma (see eq. (2.10)), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}f(\boldsymbol{x}_i,\theta_i) = \dot{\boldsymbol{x}}_i\cdot\nabla f(\boldsymbol{x}_i,\theta_i) + \dot{\theta}_i\partial_\theta f(\boldsymbol{x}_i,\theta_i) + \partial_\theta^2 f(\boldsymbol{x}_i,\theta_i). \tag{E.4}$$

Now, we substitute the dynamics in eqs. (E.1) and (E.2) for $\dot{\boldsymbol{x}}_i$ and $\dot{\theta}_i$, from which it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}f_i = \boldsymbol{e}_i\cdot\nabla f_i + \left(\beta\boldsymbol{e}_i'\cdot\hat{\boldsymbol{m}}_i^{\mathcal{C}} + \sqrt{2}\xi_i\right)\partial_\theta f_i + \partial_\theta^2 f_i, \tag{E.5}$$

where we have defined $f_i = f(\boldsymbol{x}_i,\theta_i)$ and ∇ acts on the spatial coordinate $\boldsymbol{x}$. Observe further that the left-hand side of eq. (E.5) can be expressed as

$$\frac{\mathrm{d}}{\mathrm{d}t}f_i = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\int_0^{2\pi}\mathrm{d}\theta\, f(\boldsymbol{x},\theta)p_i(\boldsymbol{x},\theta) \tag{E.6}$$

$$= \int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\int_0^{2\pi}\mathrm{d}\theta\, f(\boldsymbol{x},\theta)\partial_t p_i(\boldsymbol{x},\theta) \tag{E.7}$$

where $p_i(\boldsymbol{x},\theta)\equiv\delta(\boldsymbol{x}-\boldsymbol{x}_i)\delta(\theta-\theta_i)$ and

$$p(\boldsymbol{x},\theta)\equiv\frac{1}{N}\sum_{i=1}^{N}p_i(\boldsymbol{x},\theta) \tag{E.8}$$

is the one-particle distribution function at finite density. In the same vein, notice that the local magnetisation can be expressed in terms of p, i.e.

$$\boldsymbol{m}_i^{\mathcal{C}} = N\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\int_0^{2\pi}\mathrm{d}\theta\,\mathcal{C}_0(|\boldsymbol{x}-\boldsymbol{x}_i|)\boldsymbol{e}(\theta)p(\boldsymbol{x},\theta), \tag{E.9}$$

and so we can consider $\hat{\boldsymbol{m}}_i^{\mathcal{C}}\equiv\hat{\boldsymbol{m}}^{\mathcal{C}}(\boldsymbol{x}_i)$ a function of $\boldsymbol{x}_i$. Thus, performing a similar rewriting of the right-hand side of eq. (E.5), we then find that

$$\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\int_0^{2\pi}\mathrm{d}\theta\left[f\partial_t p_i - p_i\left(\boldsymbol{e}\cdot\nabla f + \left(\beta\boldsymbol{e}'\cdot\hat{\boldsymbol{m}}^{\mathcal{C}} + \sqrt{2}\xi_i\right)\partial_\theta f + \partial_\theta^2 f\right)\right] = 0. \tag{E.10}$$

At this stage, all that is left to do is to sum eq. (E.10) over i and integrate by parts, which yields

$$\int_{\mathcal{V}}\mathrm{d}\boldsymbol{x}\int_0^{2\pi}\mathrm{d}\theta\, f\left(\partial_t p + \boldsymbol{e}\cdot\nabla p + \partial_\theta\left(\beta\boldsymbol{e}'\cdot\hat{\boldsymbol{m}}^{\mathcal{C}}p - \partial_\theta p + \sqrt{2/N}\Xi\right)\right), \tag{E.11}$$

and acknowledge that since f was assumed to be arbitrary, the integrand must vanish identically. Thus, we obtain the DE

$$\partial_t p + \boldsymbol{e}\cdot\nabla p = -\partial_\theta\left(\beta \boldsymbol{e}'\cdot\hat{\boldsymbol{m}}^{\mathcal{C}} p - \partial_\theta p + \sqrt{2p/N}\Xi\right) \tag{E.12}$$

as reported in eqs. (3.52) and (5.40). Notice that the noise term Ξ in eqs. (E.11) and (E.12) can be written in terms of p_i and the noises $\{\xi_i\}$ as

$$\Xi = \frac{1}{\sqrt{pN}}\sum_{i=1}^{N} p_i\xi_i. \tag{E.13}$$

Clearly, Ξ is Gaussian, and moreover it is mean zero and has autocovariance

$$\left\langle \Xi(\boldsymbol{x},\theta,t)\Xi(\boldsymbol{x}',\theta',t')\right\rangle = \frac{\delta(t-t')}{N\sqrt{p(\boldsymbol{x},\theta)p(\boldsymbol{x}',\theta')}}\sum_{i=1}^{N} p_i(\boldsymbol{x},\theta)p_i(\boldsymbol{x}',\theta'). \tag{E.14}$$

Since $p_i(\boldsymbol{x},\theta)p_i(\boldsymbol{x}',\theta') = \delta(\boldsymbol{x}-\boldsymbol{x}')\delta(\theta-\theta')p_i(\boldsymbol{x},\theta)$, it therefore follows that Ξ is a unit Gaussian white noise, i.e. it has autocovariance

$$\left\langle \Xi(\boldsymbol{x},\theta,t)\Xi(\boldsymbol{x}',\theta',t')\right\rangle = \delta(\boldsymbol{x}-\boldsymbol{x}')\delta(\theta-\theta')\delta(t-t'). \tag{E.15}$$