PAPER • OPEN ACCESS

Thermodynamically consistent flocking: from discontinuous to continuous transitions

View the article online for updates and enhancements.

New Journal of Physics
The open access journal at the forefront of physics

PAPER

CrossMark

Thermodynamically consistent flocking: from discontinuous to continuous transitions

Tal Agranov[1,*], Robert L Jack[1,2] ⓘ, Michael E Cates[1] and Étienne Fodor[3] ⓘ

[1] DAMTP, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom
[2] Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom
[3] Department of Physics and Materials Science, University of Luxembourg, L-1511 Luxembourg, Luxembourg
[*] Author to whom any correspondence should be addressed.

E-mail: tal.agranov@mail.huji.ac.il

Abstract

We introduce a family of lattice-gas models of flocking, whose thermodynamically consistent dynamics admits a proper equilibrium limit at vanishing self-propulsion. These models are amenable to an exact coarse-graining which allows us to study their hydrodynamic behavior analytically. We show that the equilibrium limit here belongs to the universality class of Model C, and that it generically exhibits tricritical behavior. Self-propulsion has a non-perturbative effect on the phase diagram, yielding novel phase behaviors depending on the type of aligning interactions. For aligning interaction that increase monotonically with the density, the tricritical point diverges to infinite density reproducing the standard scenario of a discontinuous flocking transition accompanied by traveling bands. In contrast, for models where the aligning interaction is non-monotonic in density, the system can exhibit either (the nonequilibrium counterpart of) an azeotropic point, associated with a continuous flocking transition, or a state with counterpropagating bands.

1. Introduction

Flocking is a prominent phenomenon in nonequilibrium statistical mechanics [1, 2], comprising the collective motion of a large group of aligning self-propelled agents. It appears in systems ranging from bird flocks [3] to human crowds [4] and synthetic self-propelled colloids [5]. Various theoretical models lead to a qualitatively similar flocking behavior, which suggests possible universality. A pioneering theoretical account of flocking is due to Vicsek *et al* [6], who proposed an agent-based model almost three decades ago, followed by Toner and Tu [7] who offered a hydrodynamic description. Since then, these models and their variants have played a leading role in the study of active matter [8].

In the Vicsek model, locally aligning spins move with fixed speed and individually fluctuating orientation, which drives the dynamics out of equilibrium. Strikingly, under local alignment, this model undergoes an ordering transition even in two dimensions, which would be precluded in equilibrium by the Mermin–Wagner theorem [9, 10]. When increasing either density or alignment, the system transitions from a homogeneous disordered gas to a state of homogeneous collective motion with long-ranged order. Originally thought of as a continuous transition [6], it took almost another decade to establish flocking as a discontinuous transition [11]: in between the two homogeneous phases, there is a region in parameter space where dense ordered bands propagate against a dilute disordered background. Unlike equilibrium liquid-gas transitions, the critical point is at infinite density. This feature has been attributed to the fact that the dense and dilute phases have different symmetries, so that they cannot be connected via a continuous transformation [12].

A recent lattice-based model, known as the active Ising model (AIM), has led to a series of analytical predictions for the phase diagram of the flocking transition [13, 14]. In this simple extension of the Ising model, motile spins with discrete symmetry can only self-propel along a single axis. Importantly, this model

is amenable to coarse-graining, yielding an exact hydrodynamic description [12, 15, 16]. Although the phenomenology of AIM is clearly distinct from any equilibrium analogue, it is tempting to try and rationalize how AIM departs from an equilibrium reference model, with proper reversible dynamics, as was done for repulsive active particles [17, 18]. Interestingly, AIM does not reduce to equilibrium at vanishing self-propulsion, as one might expect. This inconsistency reveals that AIM is actually not compatible with the theory of stochastic thermodynamics [19], which assumes that the system is in contact with equilibrium reservoirs and driven by external forces (so that it naturally relaxes to equilibrium in the absence of drive). The same holds for other flocking models with continuous spin symmetry [6], or with topological interactions that involve nonreciprocal alignment [20, 21], which also lead to irreversibility even at vanishing self-propulsion.

It is then natural to examine flocking models where the microscopic dynamics is amended to entail an equilibrium limit in the absence of self-propulsion, which we refer to as *thermodynamic consistency*. Several questions remain open: Would novel qualitative behavior emerge from thermodynamic consistency? To which universality class does the equilibrium limit of such flocking models belong? Inspired by AIM, we explore these issues in the context of flocking systems with a discrete symmetry. To make analytical progress, we choose lattice models which, like AIM, admit an analytically exact and tractable hydrodynamic description.

This work is part of a broader current effort in recasting various active systems in a thermodynamically consistent framework such as in phase field crystal models [22], active fluids [23, 24] , and reaction-diffusion systems [25]. Although our models are permanently out of thermodynamic equilibrium, we find that their behavior can be usefully analyzed in terms of the underlying free energy that controls their equilibrium limit at vanishing self-propulsion. Indeed, the free energy plays an informative role even outside the regime of linear response to the nonequilibrium drive. As noted previously [12, 13, 26], the standard scenario of a discontinuous flocking transition stems from a density modulated coupling of the alignment interaction. In our models, we reveal a strikingly different phenomenology between the cases of monotonic and non-monotonic density modulations. For models where ordering increases monotonically with density, the equilibrium limit exhibits a single tricritical point [16], and introducing activity is a non-perturbative effect: the tricritical point diverges to infinite density, and the flocking transition is discontinuous all across phase space. In contrast, for models where ordering is non-monotonic with density, the equilibrium limit exhibits a pair of tricritical points. At finite self-propulsion, such points can either collide into a single azeotropic point, yielding a continuous flocking transition, or instead a new collective state emerges with counterpropagating bands.

This paper is organized as follows. In section 2, we set up the general definition of our thermodynamically consistent flocking models. We propose a microscopic dynamics with a generic type of aligning interactions, and derive the corresponding hydrodynamic description through systematic coarse-graining. Then, we examine in sections 3 and 4 the phase diagrams for monotonic and non-monotonic dependence of the aligning strength on density, respectively. Overall, our results illustrate how various choices of density-dependent alignment affect the flocking transition, yielding unprecedented phenomenology in the flocking context. Yet, some of this phenomenology echoes that of equilibrium systems, so that our thermodynamically consistent description allows one to anticipate and classify the topologies of nonequilibrium phase diagrams in terms of their equilibrium counterparts.
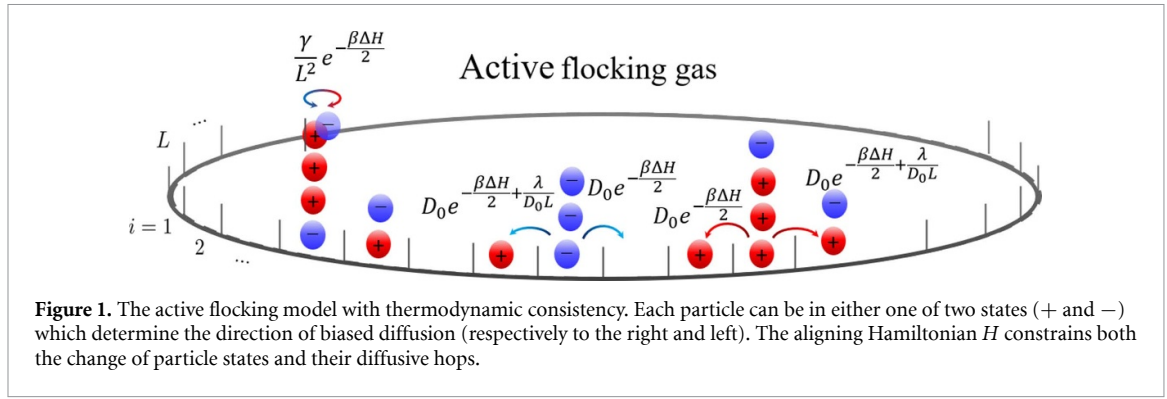
2. Thermodynamic consistency: from microscopics to hydrodynamics

In this section, we introduce a class of lattice dynamics for active spins that obey thermodynamic consistency. We take locally aligning interactions with mesoscopic range to capture a transition to collective motion. Such a form of interactions is amenable to exact coarse-graining, which allows us to derive the corresponding hydrodynamic equations for some relevant fields. Importantly, our derivation is given in a functional form with an arbitrary dependence of the alignment strength on density.

2.1. Microscopic dynamics of aligning active spins

Flocks with discrete symmetry, like in AIM [12], are defined by two species of particles ($+$ and $-$) which hop with a bias to the right or left (respectively) on a periodic lattice of $L \gg 1$ sites, see figure 1. We consider for simplicity a one dimensional lattice, yet our fluctuating hydrodynamics is trivially extended to higher dimensions. The dynamics does not exclude multiple occupancy, so that each site i can have any number of particles. The total number of particles in the system is $N = \rho_0 L$, where ρ_0 is the mean density.

The dynamical update rules consist of diffusive hops biased by self-propulsion, and spin flips $+ \leftrightarrow -$. Differently from previous models [12, 16], at zero self-propulsion we make both these rules derive from the same Hamiltonian H via a detailed balance constraint [19, 27]. We leave H unspecified at this stage.

Figure 1. The active flocking model with thermodynamic consistency. Each particle can be in either one of two states ($+$ and $-$) which determine the direction of biased diffusion (respectively to the right and left). The aligning Hamiltonian H constrains both the change of particle states and their diffusive hops.

Self-propulsion is then added as a weak bias of hops in the direction of the particle's spin, yielding the following update rules (see figure 1):

(i) Site hoping: Any particle jumps to a neighboring site with rate $D_0 e^{-\frac{\beta \Delta H}{2} + \frac{\lambda}{D_0 L}}$ if the jump is in the direction of its spin, and $D_0 e^{-\frac{\beta \Delta H}{2}}$ otherwise. Here, ΔH is the energy difference between the configurations of the system before and after the jump.

(ii) Tumbling: A $+ (-)$ particle converts into a $- (+)$ particle with rate $(\gamma/L^2) e^{-\frac{\beta \Delta H}{2}}$.

Here, D_0 is the bare diffusion constant, λ is the self-propulsion strength, and γ sets the tumble rate. The scaling of the rates with L ensures that in the hydrodynamic limit ($L \to \infty$ at fixed ρ_0) all processes occur on diffusive time scales [15, 28]. Indeed, the time it takes for particles to traverse a macroscopic system of size L, either via diffusive motion or on account of self-propulsion, scales as L^2 which is also the time scale for tumbling events. This L-dependent scaling of the rates is standard for other lattice-based diffusive models [15, 28]. The main difference with AIM [12, 13, 26] is that our site hopping now accounts for the energy difference ΔH. This feature renders our models reversible at vanishing self-propulsion, thus setting a proper equilibrium limit, as a consequence of thermodynamic consistency [19, 27].

We now restrict H such that the interaction range spans a mesoscopic scale $\Delta x = L^\delta$, with $0 < \delta < 1$. This scale is sub-extensive in system size, yet contains a large number of particles $1 \ll \Delta x \ll L$, and is the scale appearing in the coarse-graining procedure detailed below. Within this choice of scaling each particle interacts with a vanishing fraction of the system of the order of $\sim \Delta x/L \ll 1$, yet it still enables a phase transition in one dimension even in the equilibrium limit. Indeed, in defiance of the usual Landau argument (which states that having only a finite energetic cost of domain walls in one dimension destroys long range order [29]) let us consider the free energy cost of the interface between phases. The configurational entropy of a single domain wall scale as $\log L$ (as there are L sites to center it upon them). Now since each particle across the domain wall interacts with about $\sim \Delta x$ other particles, then the energetic cost of a domain wall would scale at least as $\Delta x \sim L^\delta$, which is much larger then the entropic $\log L$ contribution (compare this with a local interaction range where the energetic cost is $\mathcal{O}(1)$). Thus, the large $\mathcal{O}(L^\delta)$ free energy cost will make domain walls thermodynamically unfavorable and keep the ordered state stable.

Note that it was recently found [30] that the flocking transition does not survive in the one-dimensional AIM. Instead it is replaced by either direction-reversing ordered aggregates or static asters composed of opposing clusters of left and right movers that block each other. These arise by virtue of the purely local interaction range of the AIM and do not appear in our model as a result of our choice of a mesoscopic interaction range.

In practice, interactions are given in terms of some local average occupancies for the density and 'magnetization' variables, respectively $\eta_i^\rho = \eta_i^+ + \eta_i^-$ and $\eta_i^m = \eta_i^+ - \eta_i^-$, where η_i^+ and η_i^- are the number of $+$ and $-$ particles at site i. The local averages around site i are given by

$$\hat{\rho}_i = \frac{\sum_{|i-j| < \Delta x} \eta_j^\rho}{2 \Delta x}, \quad \hat{m}_i = \frac{\sum_{|i-j| < \Delta x} \eta_j^m}{2 \Delta x}. \tag{1}$$

We then choose the aligning Hamiltonian to take the form

$$H = -\sum_{i=1}^{L} \frac{\hat{m}_i^2}{2 \hat{\rho}_i} f(\hat{\rho}_i). \tag{2}$$

For $f(\hat{\rho}_i) = 1$, this is the sum of L fully connected Ising Hamiltonians within some restricted range Δx, where the coupling is scaled with the total number of interacting spins, see also [12]. Indeed, denote by $s_i^k \in \{+1, -1\}$ with $k = 1, \ldots, \eta_i^\rho$ the spin variable of the k'th particle on site i. Then the Ising Hamiltonian with all-to-all interactions within the range Δx of site i reads

$$-\frac{1}{2} \frac{\sum_{|i-j|<\Delta x} \sum_{|i-k|<\Delta x} \sum_{l=1}^{\eta_j} \sum_{m=1}^{\eta_k} s_j^l s_k^m}{2\Delta x \hat{\rho}_i} = -2\Delta x \frac{\hat{m}_i^2}{2\hat{\rho}_i}. \tag{3}$$

The factor $1/2$ avoids double counting. The normalization $2\Delta x \hat{\rho}_i$ ensures the interactions remain extensive with the total number of interacting spins, as is common for the fully-connected Ising Hamiltonian. Then summing the right hand side (3) over all disjoint compartments of size $2\Delta x$ is equivalent to the sum (2) with $f(\hat{\rho}_i) = 1$ which runs over all lattice points (up to sub-leading corrections in L). Then the density-dependent modulation $f(\hat{\rho}_i)$, which plays a key role in the following, tunes the alignment strength. For AIM, a similar modulation appears at a coarse-grained level, as an effective renormalization due to the hydrodynamic fluctuations [26]. In our model, since the microscopic dynamics already contains a mesoscale interaction range, we choose to directly introduce such a modulation at this level. Various forms of f can be motivated on account of different types of close neighbor interactions [31–34].

In the following, we consider the hydrodynamic limit of this model. As a preliminary step, note that for a single isolated particle, there are two important length scales: $\xi_p = \lambda L/\gamma$ is the average displacement of particles between tumbles due to self-propulsion, and $\xi_d = L\sqrt{D_0/\gamma}$ is the typical diffusive displacement, over the same time. These distances are measured in units of the lattice spacing and they are both $\mathcal{O}(L)$, because of the L-dependent hopping rates of the model.

2.2. Hydrodynamic equations for density and magnetization

We follow a standard coarse-graining procedure, inspired by similar lattice models with diffusive scaling [35], which has already been deployed in some active lattice gas models [36, 37]. We start with a spatial coarse-graining over the diffusive displacement ξ_d. I.e., we define the macroscopic spatial coordinate $x = i/\xi_d \in [0, \ell_s]$ where $\ell_s = L/\xi_d = \sqrt{\gamma/D_0}$ is the system size in the hydrodynamic representation (measured in units of ξ_d, as in [15]). The reason for this choice is that we will later analyse phase-separated profiles: their interfacial widths are of order 1 in this hydrodynamic representation and it will be convenient to take $\ell_s \gg 1$ so that the system is large compared to this interfacial width. For consistency with these choices we make a diffusive scaling of time, setting $t = \gamma \hat{t}/L^2$ with $\hat{t}$ the time variable of the microscopic system.

The local density in the vicinity of point x is defined by the mesoscopic coarse-graining (1) as:

$$\rho(x, t) = \hat{\rho}_{\frac{xL}{\ell_s}}(t), \quad m(x, t) = \hat{m}_{\frac{xL}{\ell_s}}(t). \tag{4}$$

Since ρ is conserved and m is not, the dynamics of these fields must take the general form

$$\begin{bmatrix} \partial_t \rho \\ \partial_t m \end{bmatrix} = -\partial_x \begin{bmatrix} J_\rho \\ J_m \end{bmatrix} - \begin{bmatrix} 0 \\ 2K \end{bmatrix} \tag{5}$$

where J_ρ, J_m are the conservative fluxes and K is the net rate of change of the ρ_- density due to spin flipping. As detailed in appendix A, to leading order in system size we find the following expressions

$$\begin{bmatrix} J_\rho \\ J_m \end{bmatrix} = -\frac{1}{2} \mathbb{C}(\rho, m) \partial_x \begin{bmatrix} \frac{\delta F}{\delta \rho} \\ \frac{\delta F}{\delta m} \end{bmatrix} + \text{Pe} \begin{bmatrix} m \\ \rho \end{bmatrix}, \quad K = M(\rho, m) \sinh\left(\frac{\delta F}{\delta m}\right), \quad \text{Pe} = \frac{\lambda}{\sqrt{D_0 \gamma}}. \tag{6}$$

The Péclet number $\text{Pe} = \xi_p/\xi_d$ is the standard measure of activity in similar active systems with diffusive scaling [38]. Here, $F[\rho, m] = \frac{L}{\ell_s} \int_0^{\ell_s} dx \mathcal{F}[\rho(x, t), m(x, t)]$ is the free-energy functional of the equilibrium ($\lambda = 0$) system with the corresponding free energy density given by the function

$$\mathcal{F}(\rho, m) = \frac{\rho}{2} \ln \frac{\rho^2 - m^2}{4} + \frac{m}{2} \ln \frac{\rho + m}{\rho - m} - \rho - \beta \frac{m^2}{2\rho} f(\rho), \tag{7}$$

while $\mathbb{C}$ and M are mobility coefficients given by

$$\mathbb{C} = 2 \begin{bmatrix} \rho & m \\ m & \rho \end{bmatrix}, \quad M = \sqrt{\rho^2 - m^2}. \tag{8}$$

The expressions (6) and (7) are exact as long as $\text{Pe} \neq 0$. Yet in order to describe correctly the equilibrium ($\text{Pe} = 0$) hydrodynamics for non convex $\mathcal{F}$, one has to retain surface tension terms that are sub-leading

$\mathcal{O}\left(L^{-2}|\nabla\rho|^2, L^{-2}|\nabla m|^2\right)$. Still, as is usual in the thermodynamics of phase separation, the bulk term $\mathcal{F}$ suffices for determining the binodal (coexisting) densities [4].

The dynamics (5) and (6) has a thermodynamically consistent structure. Indeed, for Pe = 0, all terms derive from the free energy $\mathcal{F}$, which then serves as a Lyapunov function, as discussed in appendix A. As shown in what follows, this equilibrium limit belongs to the universality class of Model C [16, 39]. Note that, as written, the dynamics (5) and (6) neglects some sub-leading noise terms of order $L^{-1/2}$; these can be readily deduced from the detailed-balance condition (and are included by definition in Model C).

The general structure of (5) and (6) should hold rather broadly for all thermodynamically consistent flocking models with discrete symmetry and under diffusive scaling, including models with purely local interaction range (albeit with a renormalized free energy depending on dimensionality). The conservative fluxes in (6) are linear in the free-energy derivatives, as is common when expanding close to equilibrium. That these expressions hold also far away from equilibrium is a consequence of the L-dependent scaling of the microscopic rates, and the diffusive scaling of space and time; any higher order spatial gradients will be sub-leading in system size L^{-1}. This reasoning does not apply for the flipping term K, since it does not involve spatial variations of the fields. Indeed, this term goes beyond linear order in the free energy derivatives even at the hydrodynamic level [40]. Activity (Pe > 0) leads to the extra terms Pe m and Pe ρ in the conservative fluxes (6) but otherwise does not interfere with the remaining bare equilibrium terms. Again, this is a consequence of the diffusive scaling, where self-propulsion enters as a weak bias in the microscopic dynamics.

Our thermodynamically consistent dynamics in (5) and (6) is easily related to the non-consistent one of AIM [12, 13, 26]. Our flipping term K is the same as in the hydrodynamic equations of AIM, under appropriate matching of f. The conservative fluxes J_ρ and J_m are different from those of AIM, since they stem from a different choice of hopping rules. Nevertheless, one can reproduce the corresponding terms in AIM by sending the temperature that enters these terms to infinity (namely, taking $\beta = 0$ in the free energy that enters the conservative fluxes J_ρ and J_m), as discussed in appendix A. Indeed, denote by $\mathcal{F}_0$ the free energy density evaluated at $\beta = 0$ and write F_0 for the corresponding free energy functional. Then the hydrodynamics of the AIM are also given by (5) and (6), except that one replaces $F \to F_0$ in the expressions for J

$$\begin{bmatrix} J_\rho \\ J_m \end{bmatrix}_{\text{AIM}} = -D_0 \partial_x \begin{bmatrix} \rho \\ m \end{bmatrix} + \text{Pe} \begin{bmatrix} m \\ \rho \end{bmatrix} = -\frac{1}{2} \mathbb{C}(\rho, m) \partial_x \begin{bmatrix} \frac{\delta F_0}{\delta \rho} \\ \frac{\delta F_0}{\delta m} \end{bmatrix} + \text{Pe} \begin{bmatrix} m \\ \rho \end{bmatrix}, \qquad (9)$$

while retaining the full F in K. To this extent, the source of residual irreversibly in the zero self-propulsion limit of AIM can be regarded as a *mismatch between two temperatures* controlling respectively the conservative fluxes (J_ρ, J_m) and the flipping term K.

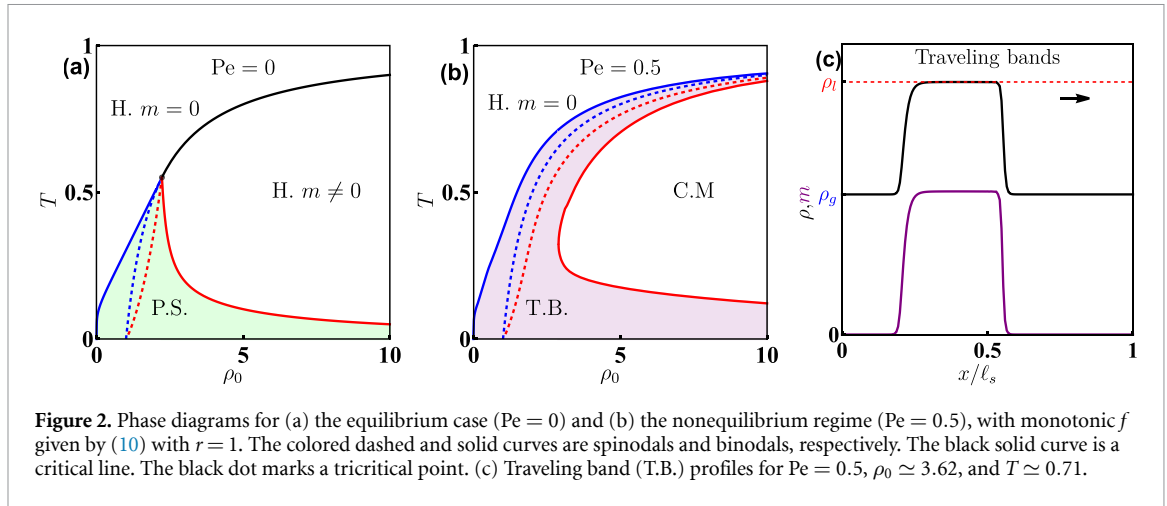## 3. Phase diagram for monotonic alignment strength

Various forms of the density-dependent alignment strength $f(\rho)$, entering in the microscopic alignment (2) and in the hydrodynamic free energy (7), can be motivated to reflect different types of close neighbor interactions [31–34]. We first explore the phase behavior for a monotonic $f(\rho)$ by considering the form appearing previously in AIM [12, 13, 26]:

$$f(\rho) = 1 - \frac{r}{\rho}. \qquad (10)$$

This choice is convenient mathematically, although on physical grounds one might instead favor choices that are positive everywhere, and do not diverge at $\rho \to 0$. Yet this does not change the phenomenology reported below. Other monotonic forms have been recently proposed in extensions of AIM yielding a similar phase behavior [16]. In some of these models, the alignment term is not purely quadratic in the magnetization as assumed here (see (2)). Yet, its expansion close to criticality reduces to this form, and the topology of the phase diagram actually follows from the behavior close to criticality.

The resulting phase diagram, given in terms of $T = 1/\beta$ and ρ_0, is composed of three phases shown in figure 2: (i) a homogeneous disordered phase (H; $m = 0$), (ii) a homogeneous ordered phase (H; $m \neq 0$), which becomes a collective motion (C.M.) for non-vanishing self-propulsion, and (iii) a phase-separated state (P.S.) between different discrete symmetries ($m = 0$ and $m \neq 0$), which becomes a traveling band (T.B.) for non-vanishing self-propulsion. In this section, we study the topology of this phase diagram and derive its phase boundaries. We address first the equilibrium limit (Pe = 0), and then the nonequilibrium regime (Pe > 0).

---

[4] The profiles presented in the following correspond to finite Pe $\neq 0$ and were computed without the need for the sub-leading interfacial tension terms.

Figure 2. Phase diagrams for (a) the equilibrium case (Pe = 0) and (b) the nonequilibrium regime (Pe = 0.5), with monotonic f given by (10) with $r = 1$. The colored dashed and solid curves are spinodals and binodals, respectively. The black solid curve is a critical line. The black dot marks a tricritical point. (c) Traveling band (T.B.) profiles for Pe = 0.5, $\rho_0 \simeq 3.62$, and $T \simeq 0.71$.

3.1. The equilibrium limit (Pe = 0): Model C dynamics

As pointed out in [16], the equilibrium, zero-propulsion limit of thermodynamically consistent flocks with discrete symmetry must generically lie within the universality class of Model C [39]. In our model, this should remain true for any non-trivial choice of $f(\rho)$, while allowing different global phase behaviors even at equilibrium. Model C describes an equilibrium dynamics which couples a conserved and a non-conserved scalar order parameter, and the canonical Landau form close to criticality is given by [41, 42]

$$\mathcal{F}_{\text{cri}} = \frac{a}{2}(\delta\rho)^2 + \frac{b}{4}(\delta\rho)^4 + \frac{c}{2}m^2 + \frac{d}{4}m^4 + em^2\delta\rho, \tag{11}$$

where higher order terms and terms purely linear in the density are omitted. Expanding our free energy in (7) around a homogeneous neutral state ($\rho = \rho_0, m = 0$), one can derive the expressions of the different coefficients in (11), which are determined by the choice of $f(\rho)$ [5]. Importantly, the coupling coefficient e in (11) reads

$$e = \beta f'(\rho_0), \tag{12}$$

which vanishes for a constant $f(\rho)$; in this special case, our model falls outside the Model C universality class. (As discussed in section 4, this also makes the nonequilibrium Pe > 0 regime different from the usual flocking scenario.) As detailed in [42], the topology of the phase diagram is set by the values of the coefficients in (11). As we have the exact form of the free energy at our disposal (7), we follow the procedure presented in [42] to map the phase diagram across the whole phase space, even beyond criticality.
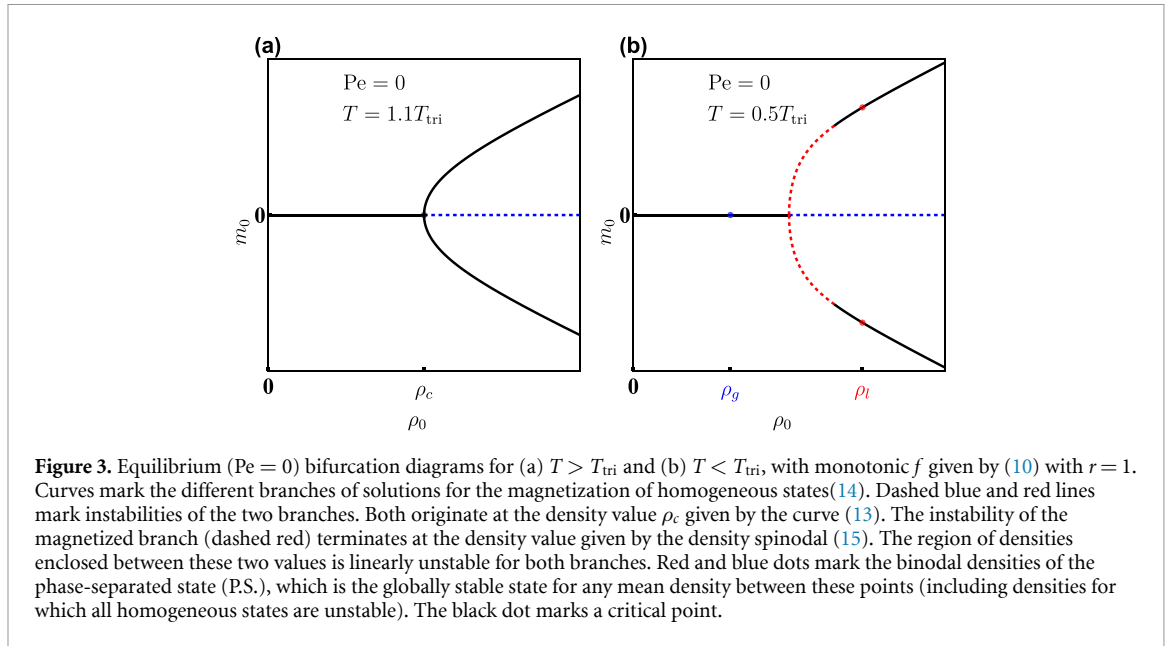
We find the stable phases by minimizing the free energy (7) under the constraint of conserved density $(1/\ell_s)\int_0^{\ell_s} \rho \, dx = \rho_0$. These belong to one of three possibilities: the homogeneous neutral state ($\rho = \rho_0, m = 0$) (marked in figure 2 by H.$m = 0$), the homogeneous magnetized state ($\rho = \rho_0, m = m_0 \neq 0$) (marked by H.$m \neq 0$) or a phase separated state (P.S.) with spatially varying magnetization and density. In all of the stable states, since magnetization is not conserved, its value is locked to the local value of the density via the minimization of F with respect to the magnetization. A non-trivial minimizer $m_0 \neq 0$ then emerges at $\partial_{mm}\mathcal{F}|_{m=0} = 0$ which defines the curve

$$T_c = f(\rho). \tag{13}$$

Generally, the relation (13) only establishes linear instability of the neutral state ($m = 0$). In regions that also admit phase separation, as discussed below, this curve will then mark a spinodal instability which we will term the *magnetization spinodal*, see dashed blue curve in figure 2(a). Only in the absence of phase separation, this curve marks a second-order transition between disordered and ordered states, see solid black curve in figure 2(a). In the ordered phase, the value $m_0(\rho, T)$ is found from $\partial_m\mathcal{F} = 0$, which gives the magnetization in an implicit form:

$$m_0 = \rho z(\rho, T), \qquad \frac{\text{arctanh} z}{z} = \beta f(\rho), \tag{14}$$

---

[5] The expansion also produces constant terms, a linear term $\delta\rho$, and a cubic term $\delta\rho^3$. The linear term integrates to zero due to mass conservation. The cubic term in general does not, yet it shall not affect the phase diagram, see [41].

Figure 3. Equilibrium (Pe = 0) bifurcation diagrams for (a) $T > T_{\text{tri}}$ and (b) $T < T_{\text{tri}}$, with monotonic f given by (10) with $r = 1$. Curves mark the different branches of solutions for the magnetization of homogeneous states(14). Dashed blue and red lines mark instabilities of the two branches. Both originate at the density value ρ_c given by the curve (13). The instability of the magnetized branch (dashed red) terminates at the density value given by the density spinodal (15). The region of densities enclosed between these two values is linearly unstable for both branches. Red and blue dots mark the binodal densities of the phase-separated state (P.S.), which is the globally stable state for any mean density between these points (including densities for which all homogeneous states are unstable). The black dot marks a critical point.

where the second equality defines the dependence of z on ρ and T. A state with magnetization (14) minimizes the free energy only for homogeneous profiles. To find phase separation, one has to follow the usual procedure of a common-tangent construction over the single-variable function $\mathcal{F}[\rho, m_0(\rho, T)]$ [42]. Non-convexity of this function marks linear instability against spatial density variations, and is equivalent to non-convexity of the two-variable free energy $\mathcal{F}$:

$$|\text{Hess}(\mathcal{F})| < 0, \tag{15}$$

where $|\text{Hess}(\mathcal{F})|$ is the Hessian determinant of $\mathcal{F}(\rho, m)$. An explicit expression for (15) is presented in appendix B.

The relation (15) defines a spinodal curve that is complementary to the magnetization spinodal (13), see dashed red line figure 2(a), which we will term the *density spinodal* (since instability here also involves density modulations). The two spinodal curves meet at a tricritical point, as detailed in appendix B, whose location obeys for the density ρ_{tri}

$$\left[\rho_{\text{tri}} \frac{f'(\rho_{\text{tri}})}{f(\rho_{\text{tri}})}\right]^2 = \frac{2}{3}. \tag{16}$$

The region enclosed between the two spinodals (13) and (15) defines the region of spinodal decomposition. This is shown by the bifurcation diagram figure 3 which marks instabilities along a constant temperature cross section of figure 2(a). At $T > T_{\text{tri}}$ (figure 3(a)) the neutral branch $m = 0$ loses stability at the density given by (13) while the entire branch of magnetized states $m_0 \neq 0$, prescribed by the nontrivial solution to (14), is stable. In contrast, at $T < T_{\text{tri}}$ (figure 3(b)), the magnetized branch has an unstable segment which terminates at the spinodal density (15). Thus, any homogeneous state with density within the region enclosed between the two spinodals (13) and (15) is linearly unstable.

Any monotonic $f(\rho)$ that vanishes at a finite density while saturating at infinity will have at least one root of (16), yielding a phase diagram like that in figure 2(a). At $T > T_c(\rho_{\text{tri}})$, the curve (13) marks a line of critical points separating the homogeneous disordered ($m = 0$) and ordered ($m \neq 0$) phases. At $T < T_c(\rho_{\text{tri}})$ the transition to the ordered phase becomes discontinuous with a region of coexistence between an ordered liquid and a disordered gas. Within the miscibility gap, the relation (13) marks the magnetization spinodal, and the complementary density spinodal given by (15) lies within the symmetry-broken phase. The location of the binodal curves follows from a common tangent construction, as detailed in appendix B. Note that, although coexisting bulk phases are found within this procedure, properly resolving the interfaces would necessitate retaining sub-leading gradient terms, which quantify surface tension in the free energy. Note also that Model C can exhibit additional phase diagram topologies, such as critical points and double critical end points [42], under appropriate tuning of $f(\rho)$. These are absent for the curves (10), but could arise for other sufficiently elaborate choices, including strictly monotonic curves.

3.2. The nonequilibrium regime (Pe > 0): Discontinuous flocking transition

The phase diagram for the monotonic $f(\rho)$ obeying (10) at finite Pe = 0.5 is shown in figure 2(b). The tricritical point shifts to infinite density, turning the second-order transition line into a first order coexistence region (miscibility gap) along its entire length. The same phase behavior is observed for AIM and its recent variants [13, 16]. Within the miscibility gap, the phase separated state becomes a traveling band (T.B.), as shown in figure 2(c). Indeed, at non-vanishing self-propulsion, any magnetized bulk phase attains a non-vanishing density flux given by Pe m, according to (6). Then, simple flux balance across the interface implies that it must propagate with velocity given by

$$V = \mathrm{Pe}\frac{\Delta m}{\Delta \rho}, \tag{17}$$

where $\Delta\rho$ and Δm are the density and magnetization difference between phases. Moreover, we have found numerically that the propagation velocity is bounded below as $|V| \geqslant \mathrm{Pe}$, for all models that we have examined.

In the absence of a free energy minimization principle, there is no simple recipe to establish the binodals in the nonequilibrium case. Throughout this paper, the nonequilibrium phase diagrams were found by numerically integrating the hydrodynamics (5) and (6) using a finite difference scheme for space and time. Then binodals are identified as the plateaus of the traveling density profiles as shown in figure 2(c). For more details see appendix D.

To establish more broadly the generic topology of the phase diagram in figure 2(b), we analyze the linear instability of the dynamics. First, the magnetization spinodal line (13) remains unchanged even at finite Pe > 0, since it describes instabilities against *homogeneous* variations. As the active terms in the dynamics (5) and (6) only couple to spatial flux derivatives, they do not affect this instability. In contrast, the complementary density spinodal (15) is changed. As we show in appendix C the instability in (15) is now modified into the following criterion

$$|\mathrm{Hess}\,(\mathcal{F})| < \frac{\mathrm{Pe}^2}{2M(\rho_0, m)\,\rho_0}\left[\left(\frac{\partial_{\rho m}\mathcal{F}}{\partial_{mm}\mathcal{F}}\right)^2 - 1\right], \tag{18}$$

where the mobility M is given in (8), and the magnetization m is to be evaluated at $m = m_0(\rho, T)$ in (14).

Based on other flocking models, it was argued that the flocking transition should generically be discontinuous everywhere [2]. In our model, this holds true if the conditions in (13) and (18) predict a finite spinodal gap at any Pe > 0. It is useful to notice that, at the magnetized state $m = m_0(\rho, T)$ in (14), we get $\partial_{\rho m}\mathcal{F}/\partial_{mm}\mathcal{F} = -\partial m_0/\partial\rho$, so that the density spinodal (18) becomes

$$|\mathrm{Hess}\,(\mathcal{F})| < \frac{\mathrm{Pe}^2}{2M(\rho_0, m)\,\rho_0}\left[\left(\frac{\partial m_0}{\partial\rho}\right)^2 - 1\right]. \tag{19}$$

Moreover, close to the magnetization spinodal (13), we have $m_0 \sim \sqrt{T_c(\rho) - T}$ with $T_c = f(\rho)$, yielding
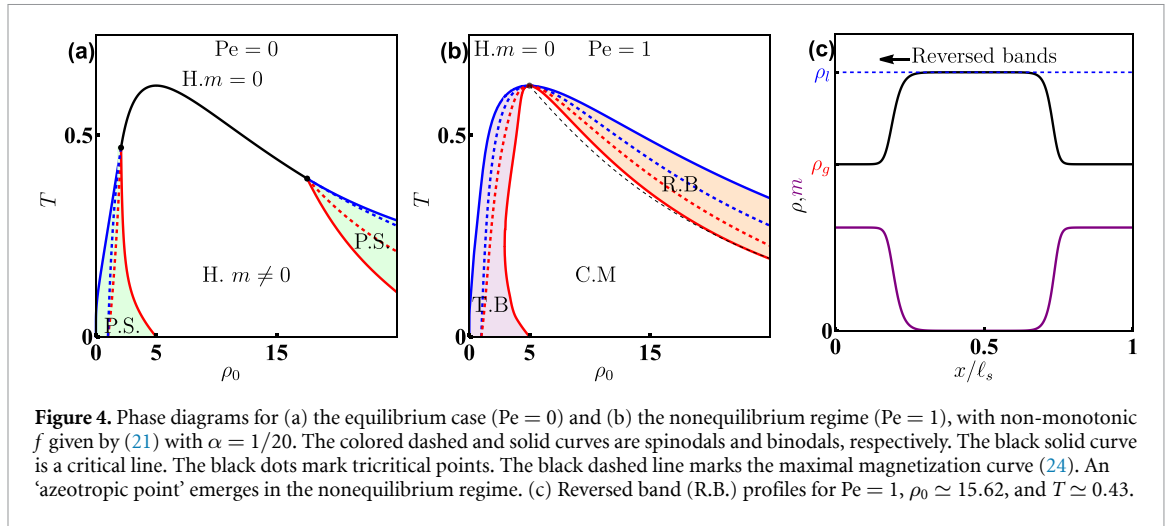
$$\left.\frac{\partial m_0}{\partial\rho}\right|_{\rho_0} \sim \frac{T_c'(\rho_0)}{\sqrt{T_c(\rho_0) - T}}. \tag{20}$$

Given that (20) diverges close to the magnetization spinodal (13), a finite neighborhood of the magnetization spinodal remains unstable at arbitrary small Pe no matter where one is on the phase diagram. Therefore, the non-perturbative effect of activity on the phase diagram can be attributed to the singularity of the order disorder transition of the equilibrium limit, as encoded in m_0. Remarkably, however, this argument fails for any non-monotonic $f(\rho)$ at its maximum $f'(\rho_0) = 0$. As discussed next in section 4, the case of non-monotonic f actually leads to novel phase behaviors and a diagram very different from figure 2(b).

4. Phase diagram for non-monotonic alignment strength

Non-monotonic alignment strength functions $f(\rho)$ are likely to arise physically in systems where alignment is restricted by visual obstruction, as was recently proposed to model natural bird flocking [43]. Non-monotonicity opens the door to novel phase behaviors, not encountered for the monotonic case in section 3 above. In equilibrium (Pe = 0), when $f(\rho)$ features a single maximum, one typically finds two roots for (16) marking a pair of tricritical points. This scenario leads to two distinct miscibility gaps, as shown in figure 4(a). The miscibility gap to the left of the maximum has a disordered low-density phase coexisting

Figure 4. Phase diagrams for (a) the equilibrium case (Pe = 0) and (b) the nonequilibrium regime (Pe = 1), with non-monotonic f given by (21) with $\alpha = 1/20$. The colored dashed and solid curves are spinodals and binodals, respectively. The black solid curve is a critical line. The black dots mark tricritical points. The black dashed line marks the maximal magnetization curve (24). An 'azeotropic point' emerges in the nonequilibrium regime. (c) Reversed band (R.B.) profiles for Pe = 1, $\rho_0 \simeq 15.62$, and $T \simeq 0.43$.

with an ordered higher-density phase. The miscibility gap to the right of the maximum has a disordered high-density phase coexisting with an ordered lower-density phase.

In the rest of this section, we consider how self-propulsion (Pe > 0) affects this equilibrium picture. The regime of non-monotonic f is more elaborate compared with the monotonic case in section 3. In particular, we distinguish two qualitatively different phase behaviors, depending on the curvature of $f(\rho)$ close to its maximum. For illustration purposes, we use particular examples in a specific family of non monotonic-curves:

$$f(\rho) = \left(1 - \frac{1}{\rho}\right) e^{-\alpha\rho}, \tag{21}$$

with $\alpha > 0$. In what follows, we show that the topology of the phase diagram is controlled by local properties of f close to the maximum, so that the exponential cutoff at large densities is immaterial for our purposes.

4.1. Azeotropic point in the nonequilibrium case (Pe > 0)

For mildly curved functions $f(\rho)$, we find numerically that the tricritical points collide at the maximum of f at any finite Pe, see figure 4(b). As for the case of monotonic $f(\rho)$ in section 3, the effect of self-propulsion on the phase diagram is non-perturbative: the equilibrium critical line expands almost everywhere into a finite miscibility gap, except at a single point which remains critical. Thus, in contrast with the case of monotonic $f(\rho)$, there does remain a critical point on the phase diagram at finite Pe. Therefore, a non-monotonic $f(\rho)$ can allow for a continuous transition from homogeneous disorder (H., $m = 0$) to collective motion (C.M.).

For equilibrium systems showing a phase-diagram topology equivalent to figure 4(b), the critical point is called an azeotropic point [44], so that we also adopt the same term here. This point marks the pairwise meeting of two branches of first order lines, with miscibility gaps that display distinct behavior. The left gap features the usual traveling bands (T.B.) as in figure 2(c). The right gap features reversed bands (R.B.), shown in figure 4(c). These bands propagate in the direction opposite to the bulk magnetization. Indeed, in the corresponding miscibility gap, the coexistence is between a magnetized lower-density phase and a paramagnetic high-density phase. Then, according to the flux balance condition (17), it follows that the phase boundaries propagate in the reverse direction from usual.

Our scenario shows that the flocking transition need not *necessarily* be discontinuous, at variance with a widely held view among those studying flocking [2], although it remains an exceptional case. In order to rule out some continuous transition scenarios, note that the disordered (H., $m = 0$) and the ordered phases (H., $m \neq 0$ at Pe = 0 or C.M. at Pe $\neq$ 0) entail different symmetries. Thus, it is not possible to pass between these pure phases without crossing a (symmetry-breaking) phase transition of some kind. This argument precludes the standard scenario of liquid–liquid phase separation, where a single pair of binodals terminate at a critical point. Indeed, one sees from the equilibrium phase diagrams in figures 2(a) and 4(a) that the binodals terminate at tricritical points, and any path between the pure phases must cross a phase transition line. In the nonequilibrium phase diagrams of figures 2(b) and 4(b), one observes discontinuous transitions as well as an azeotropic point, all of which are consistent with the relevant symmetry principles.
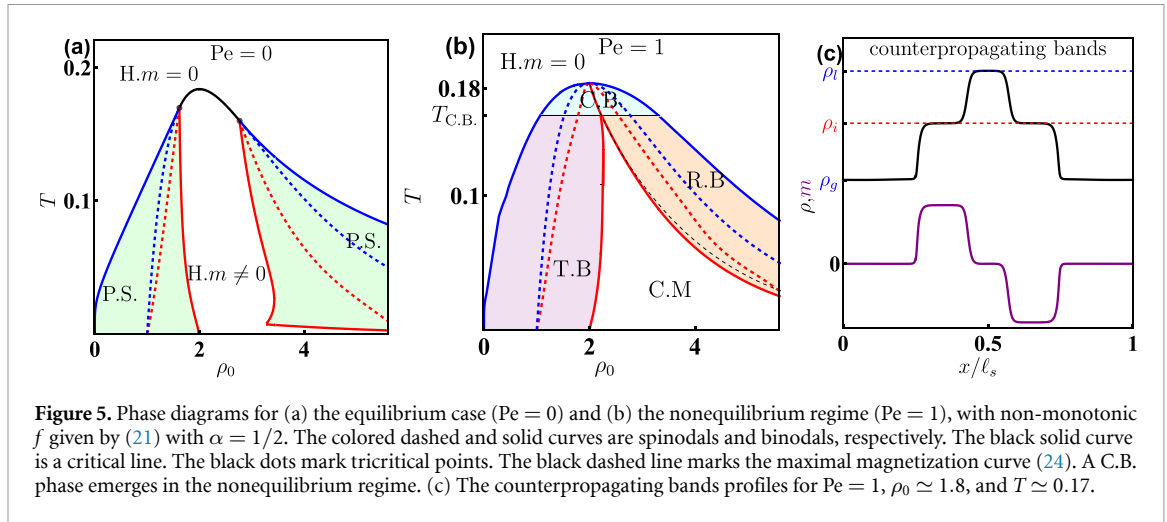
Figure 5. Phase diagrams for (a) the equilibrium case (Pe = 0) and (b) the nonequilibrium regime (Pe = 1), with non-monotonic f given by (21) with $\alpha = 1/2$. The colored dashed and solid curves are spinodals and binodals, respectively. The black solid curve is a critical line. The black dots mark tricritical points. The black dashed line marks the maximal magnetization curve (24). A C.B. phase emerges in the nonequilibrium regime. (c) The counterpropagating bands profiles for Pe = 1, $\rho_0 \simeq 1.8$, and $T \simeq 0.17$.

4.2. Counterpropagating bands in the nonequilibrium case (Pe > 0)

When $f(\rho)$ is sufficiently curved, two tricritical points in the equilibrium limit are closer to the maximum of $f(\rho)$, see figure 5(a). Here, we observe a different behavior at finite Pe compared with section 4.1. Instead of the azeotropic point reported in figure 4(b), there now appears a region with counterpropagating bands (C.B.), as shown in figure 5(b). This happens when the two magnetised binodals (red solid lines) merge, at a finite temperature $T_{\text{C.B.}}$. For temperatures between $T_{\text{C.B.}}$ and $T_{\max} = \max[f(\rho)]$, the system enters the C.B. phase (see figure 5(c)), a strongly fluctuating dynamical steady state whose properties can be understood in terms of the partial restoration of a broken symmetry, as we discuss next.

Recall that within the symmetry-broken T.B. and R.B. phases, the steady state consists of a single magnetized band, which travels through the system. The magnetization may be either positive or negative: this spontaneous breaking of spin-reversal symmetry is familiar from the Ising model. In turn, this symmetry breaking drives a particle current, which causes the band to move either left or right. This additionally breaks the spatial reflection symmetry of the model. Due to the periodic boundary conditions of the system, the resulting steady state is also time-periodic, with period ℓ_s/V.

In the C.B. phase, both left- and right-moving bands are present at the same time, as shown in figure 6. Since they propagate in opposite directions, the two bands regularly meet each other. It turns out that they interact weakly (we recall that the number of particles on each site is unbounded, in contrast with exclusion processes). The resulting situation is most simply represented in terms of the density of $+$ and $-$ particles

$$\rho_\pm (x,t) = \frac{\rho(x,t) \pm m(x,t)}{2}. \tag{22}$$

As shown in figures 7(a) and (b), the space-time dependence of these profiles is well-approximated by $\rho_+ \simeq \bar{\rho}(x - Vt)$ and $\rho_- \simeq \bar{\rho}(-x - Vt)$. In this case, the density $\rho(x,t)$ is still time-periodic, with behaviour invariant under spatial reflection, so that it is analogous to a standing-wave solution, while the T.B. and R.B. are analogous to travelling-wave solutions. These analogues of standing waves and travelling waves are strongly anharmonic, as expected since the governing equations are non-linear. In fact, $\bar{\rho}$ is close to a top-hat profile, as it was in the T.B./R.B. phases: there are extended bulk regions at densities (ρ_1, ρ_2), separated by interfaces whose widths are $\mathcal{O}(1) \ll \ell_s$. Plotting the density profile at generic times, as shown in figure 6, one typically finds up to four distinct regions whose densities are the following. Outside of either band the density is that of the dilute binodal $\rho_g = 2\rho_1$. Inside both bands, the density is that of the dense binodal $\rho_l = 2\rho_2$. Finally, inside one band and outside the other, the density is that of the intermediate binodal $\rho_i = \rho_1 + \rho_2$. Hence the intermediate binodal is always halfway between the extreme binodals, while their difference equals twice the magnetization m_0:

$$\rho_g(T) = \rho_i(T) - m_0(T), \quad \rho_l(T) = \rho_i(T) + m_0(T). \tag{23}$$

An important consequence of (23) is that the propagation velocity (17) attains its limiting minimal value $|V| = \text{Pe}$. The magnetization in (23) marks the bulk magnetization of the intermediate binodal $m_0 = m_0(\rho_i, T)$, as given explicitly by (14). Therefore, both binodals $\rho_{g,l}$ in (23) are fully determined by the value of the intermediate binodal ρ_i. In practice, we find empirically that the binodal ρ_i coincides with the maximal magnetization curve $\rho_{\max}$, defined by $\partial_\rho m_0(\rho_{\max}, T) = 0$, as shown in dashed black line in figure 5(b). We get an implicit representation of $\rho_{\max}(T)$ by differentiating (14) with respect to ρ, yielding
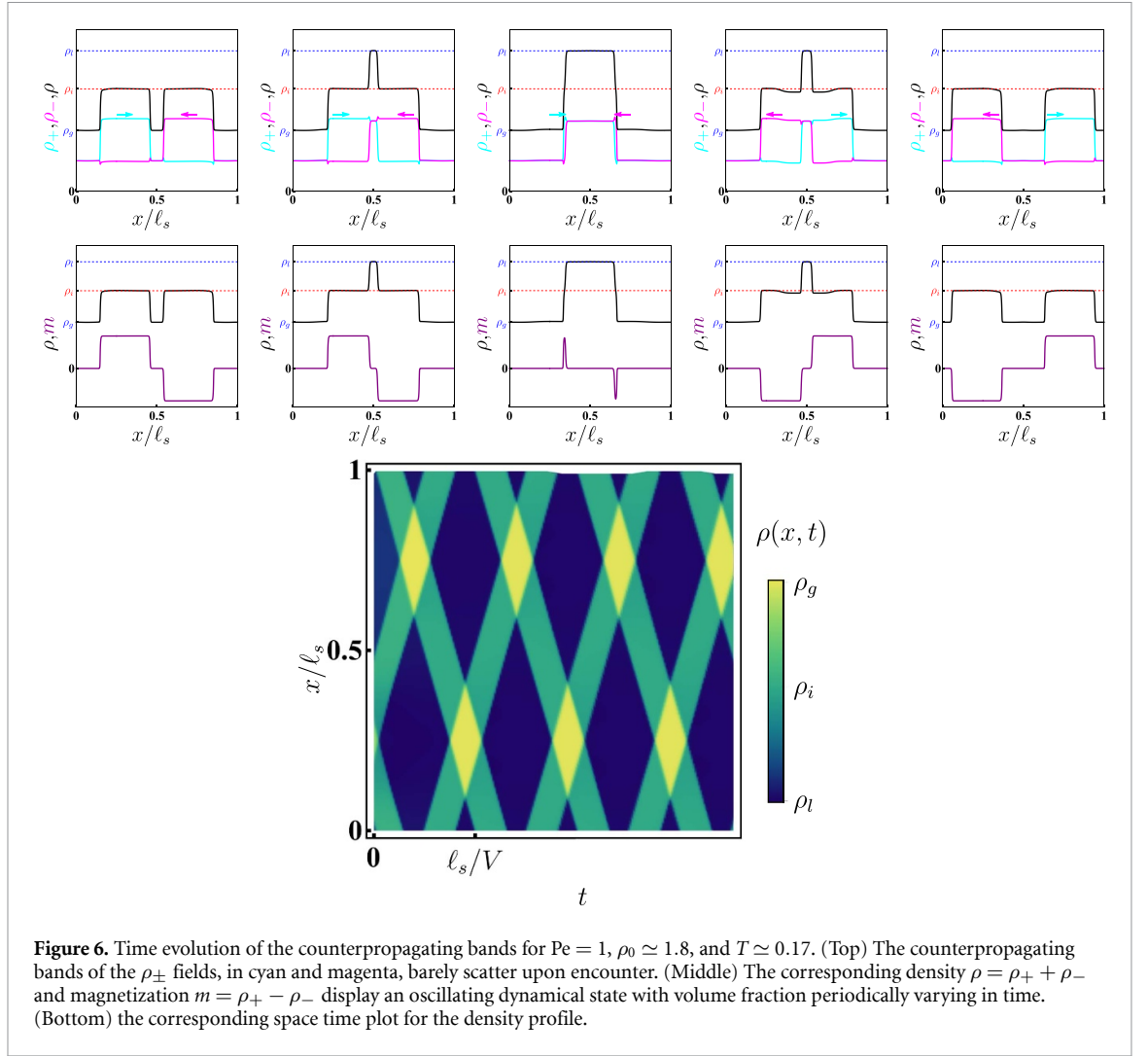
Figure 6. Time evolution of the counterpropagating bands for Pe = 1, $\rho_0 \simeq 1.8$, and $T \simeq 0.17$. (Top) The counterpropagating bands of the $\rho_\pm$ fields, in cyan and magenta, barely scatter upon encounter. (Middle) The corresponding density $\rho = \rho_+ + \rho_-$ and magnetization $m = \rho_+ - \rho_-$ display an oscillating dynamical state with volume fraction periodically varying in time. (Bottom) the corresponding space time plot for the density profile.

$$\frac{z}{(1-z^2)\operatorname{arctanh}z} = 1 - \frac{\rho_{\max}f'(\rho_{\max})}{f(\rho_{\max})}, \quad T = \frac{z}{\operatorname{arctanh}z}f(\rho_{\max}). \tag{24}$$

Overall, equations (23) and (24) specify all three binodals of the C.B. phase in closed analytic form.
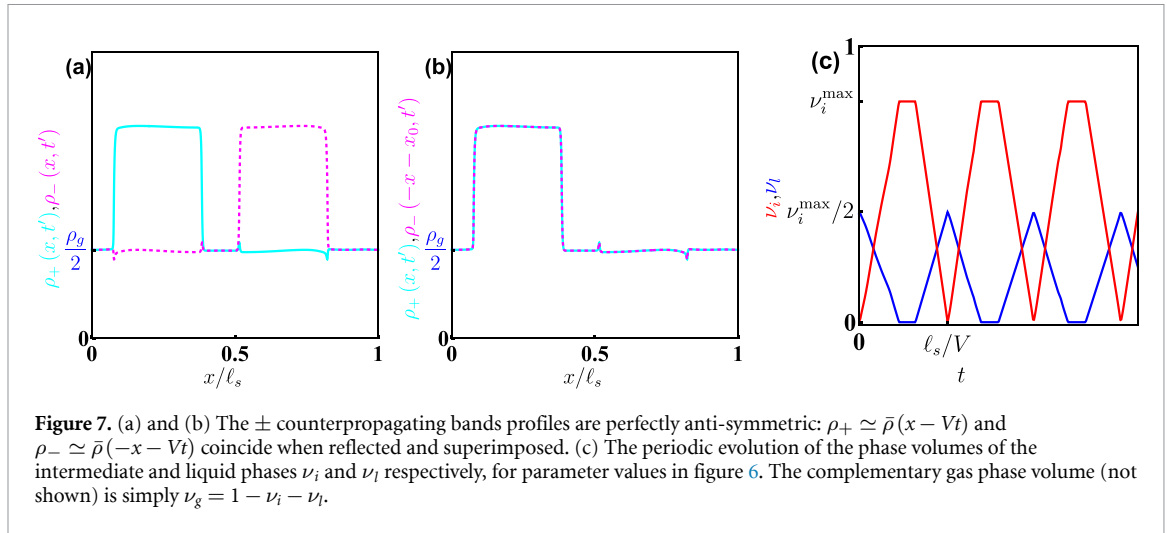
We now address the phase volumes of the bulk phases, defined as the fraction of space that each phase occupies, and the corresponding highly dynamical nature of the C.B. phase state. Since the profiles of ρ_+ and ρ_- are symmetric, each one of them carries exactly half of the total mass $L\rho_0/2$. Besides, given that $\pm$ bands counter-propagate at constant speed, their overlap region varies periodically with time. We report in figure 7(c) the corresponding dynamics of the phase volumes for a mean density below the intermediate binodal ($\rho_0 < \rho_i$). The maximal phase volume of the intermediate phase $\nu_i^{\max}$ is set by the usual lever rule between the intermediate and gas binodals: $\nu_i^{\max} = (\rho_i - \rho_0)/(\rho_i - \rho_g)$. Then, the maximal phase volume for the liquid phase occurs at perfect overlap of the $\pm$ profiles and is exactly $\nu_i^{\max}/2$, as shown in figure 7(c).

Interestingly, one can formulate a necessary condition for the emergence of the C.B. phase state. Indeed, the liquid and gas binodals must lie outside the spinodals ρ_g^s and ρ_l^s:

$$\rho_g(T) < \rho_g^s(T) < \rho_l^s(T) < \rho_l(T). \tag{25}$$

The condition (25) must hold across the C.B. phase. In particular, it should hold in the vicinity of the maximum of f, where the C.B. phase first emerges. Approximating $f(\rho)$ by a parabola in this region, the spinodal densities $\rho_{g,l}^s$ defined in (13) are symmetrically positioned around the maximum ρ^*:

$$\rho_g^s \simeq \rho^* - \Delta_s, \quad \rho_l^s \simeq \rho^* + \Delta_s, \quad \Delta_s = \sqrt{\frac{2\epsilon f(\rho^*)}{|f''(\rho^*)|}}, \quad \epsilon = \frac{f(\rho^*) - T}{f(\rho^*)} \ll 1. \tag{26}$$

Figure 7. (a) and (b) The $\pm$ counterpropagating bands profiles are perfectly anti-symmetric: $\rho_+ \simeq \bar{\rho}(x - Vt)$ and $\rho_- \simeq \bar{\rho}(-x - Vt)$ coincide when reflected and superimposed. (c) The periodic evolution of the phase volumes of the intermediate and liquid phases ν_i and ν_l respectively, for parameter values in figure 6. The complementary gas phase volume (not shown) is simply $\nu_g = 1 - \nu_i - \nu_l$.

Performing a similar expansion for the maximal magnetization curve in (24) yields

$$\rho_{\max}(T) \simeq \rho^*, \quad m_0(\rho_{\max}, T) \simeq \rho^* \sqrt{3\epsilon}. \tag{27}$$

Substituting the expression (27) in (23), and recalling that $\rho_i = \rho_{\max}$ in the C.B. phase, we find that the liquid and gas binodals are also symmetrically positioned around the maximum:

$$\rho_g \simeq \rho^* - \Delta_b, \quad \rho_l \simeq \rho^* + \Delta_b, \quad \Delta_b = \sqrt{3\epsilon}\rho^*. \tag{28}$$

Finally, the condition (25) for existence of the C.B. phase is equivalent to the ordering $\Delta_b > \Delta_s$ of the binodal and spinodal gaps. Using the expressions for these gaps, respectively in (26) and (28), we arrive at the following condition for the curvature near the maximum:

$$\frac{\rho^{*2} |f''(\rho^*)|}{f(\rho^*)} > \frac{2}{3}. \tag{29}$$

Remarkably, the condition (29) is independent of the Péclet number, which corroborates that the C.B. phase appears as a singular perturbation of the underlying equilibrium phase diagram at any finite activity, no matter how small.

5. Discussion

Our results show that the density-dependence of the aligning interactions can play an unexpectedly major role in flocking models, yielding a richer flocking phenomenology than any previously reported. These results are found by choosing a thermodynamically consistent approach which entails a proper equilibrium limit, pertaining to the universality class of Model C [16, 39], at vanishing self-propulsion. Such rich phenomenology should not be restricted to thermodynamically consistent models provided an appropriate choice of density dependent interactions. However, we find that thermodynamically consistent models are best suited for anticipating and understanding these phenomena. In particular, the spinodal condition (13) which is given in terms of the equilibrium free energy, helps rationalize why a tricritical point in the equilbirium limit (figure 2(a)) 'runs away' to infinite density in the presence of self propulsion (figure 2(b)), recovering the standard scenario of discontinuous flocking [2]. Moreover, for non-monotonic alignment *f*, one has that the pair of tricritical points that exist in the equilibrium limit (figure 4(a)) must collide at the maximum of *f* in the presence of self propulsion, giving way to more exotic phase topologies which could involve either an azeotropic point (figure 4(b)) or a region of counterpropagating bands phase (figure 5(b)). Remarkably, in contrast with standard flocking models [2], the flocking transition is continuous in the presence of an azeotropic point, while remaining discontinuous elsewhere on the phase diagram.

Our models employ a diffusive scaling of the microscopic rates to allow for an analytically exact derivation of a hydrodynamic description. However, the phenomenology is expected to persist for models without such scaling. Indeed, although the hydrodynamic description of the AIM employs a similar diffusive scaling, microscopic simulations show that the predicted flocking transition phenomenology holds without such scaling, see [12]. The mesoscopic interaction range Δx used here enables these phenomena to appear in

our simple one-dimensional setting. We expect that in higher dimensions much of the phenomenology would survive also for locally interacting models.

Given the rich topology of these phase diagrams, it would be interesting to examine how the entropy production rate (EPR) varies across the various phase transitions that we have unveiled. In thermal systems [19, 27], EPR measures both the breakdown of time-reversal symmetry and the amount of energy dissipated by nonequilibrium currents. For flocking systems, irreversibility measures have been provided in both microscopic models [45–48] and field theories [49, 50]. Yet, these measures should coincide with a quantification of dissipation only for thermodynamically consistent models [51–53]. In that respect, our approach opens the unprecedented opportunity to properly study how the dissipation varies across the flocking transition. For instance, it could lead to a spatial resolution of dissipation, thus uncovering the profile of energy dissipation associated with travelling bands.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgments

Appendix A. Deriving hydrodynamic equations

The microscopic dynamics described in section 2 consists of biased diffusive hops along with reaction dynamics. To derive the corresponding hydrodynamics, we follow the recipe outlined in [28], which has already been implemented in other models involving active lattice gases [15, 36–38]. A necessary property for the derivation is that the underlying microscopic dynamics is dominated by symmetric diffusive hops, which have the fastest transition rate.

First, we express the Hamiltonian (2) in terms of the hydrodynamic fields (4). Assuming that these fields are slowly varying over the macroscopic scale (which will be verified *a posteriori*), then one has, to leading order at large L,

$$H = \frac{L}{\ell_s} \int_0^{\ell_s} \mathrm{d}x \mathcal{H}\left[\rho\left(x,t\right), m\left(x,t\right)\right], \tag{A.1}$$

with the Hamiltonian density

$$\mathcal{H}\left(\rho, m\right) = -\frac{m^2}{2\rho} f(\rho). \tag{A.2}$$

When a $+$ particle at position $i = Lx_i/\ell_s$ hops to the right, the field variations follow as

$$\rho \to \rho + \frac{\ell_s}{L}\delta\left(x - x_{i+1}\right) - \frac{\ell_s}{L}\delta\left(x - x_i\right) = \rho - \frac{\ell_s^2}{L^2}\partial_x\delta\left(x - x_i\right) + o\left(L^{-2}\right). \tag{A.3}$$

Similarly, we have

$$m \to m - \frac{\ell_s^2}{L^2}\partial_x\delta\left(x - x_i\right) + o\left(L^{-2}\right), \tag{A.4}$$

and analogous expressions hold for the other jumps. The expression for the field variations under a flip of $+ \to -$ is simply given by

$$\rho \to \rho, \quad m \to m - 2\frac{\ell_s}{L}\delta\left(x - x_i\right) + o\left(L^{-1}\right). \tag{A.5}$$

Substituting the expressions (A.3) and (A.4) into (A.1), we find that the jump of a $+$ particle to the right leads to the Hamiltonian increment

$$\Delta H = \frac{\ell_s}{L}\partial_x\left(\frac{\delta H}{\delta\rho} + \frac{\delta H}{\delta m}\right), \tag{A.6}$$

with similar expressions hold for the other jumps. The Hamiltonian increment under $+ \to -$ flipping is simply

$$\Delta H = -2\frac{\delta H}{\delta m}. \tag{A.7}$$

Substituting (A.6) and (A.7) in the definition of the rates of section 2, we find that these rates are functions of the local fields to leading order:

$$
\begin{aligned}
\text{rate}_+ (i \to i+1) &= D_0 \left[1 + \frac{1}{L} \left(\frac{\lambda}{D_0} - \frac{\beta}{2} \ell_s \partial_x \left(\frac{\delta H}{\delta \rho} + \frac{\delta H}{\delta m} \right) \Big|_{\rho(x,t),m(x,t)} \right) \right], \\
\text{rate}_+ (i \to i-1) &= D_0 \left[1 + \frac{\beta \ell_s}{2L} \left(\frac{\delta H}{\delta \rho} + \frac{\delta H}{\delta m} \right) \Big|_{\rho(x,t),m(x,t)} \right], \\
\text{rate}_- (i \to i+1) &= D_0 \left[1 - \frac{\beta \ell_s}{2L} \left(\frac{\delta H}{\delta \rho} - \frac{\delta H}{\delta m} \right) \Big|_{\rho(x,t),m(x,t)} \right], \\
\text{rate}_- (i \to i-1) &= D_0 \left[1 + \frac{1}{L} \left(\frac{\lambda}{D_0} + \frac{\beta}{2} \ell_s \partial_x \left(\frac{\delta H}{\delta \rho} - \frac{\delta H}{\delta m} \right) \Big|_{\rho(x,t),m(x,t)} \right) \right], \\
\text{rate} (+ \to -) &= \frac{\gamma}{L^2} e^{\beta \frac{\delta H}{\delta m}}.
\end{aligned}
\tag{A.8}
$$

Importantly, the leading order terms are symmetric hops with rate D_0. These dominate over both the asymmetric $\mathcal{O}(L^{-1})$ jump rate, and over the slower $\mathcal{O}(L^{-2})$ tumbling rate defined in section 2. This timescale separation guarantees that the measure is controlled by the simple symmetric hop rules corresponding to noninteracting random walkers [28]. It is then straightforward to write down the coupled hydrodynamics of the ρ_+ and ρ_- fields

$$
\begin{bmatrix} \partial_t \rho_+ \\ \partial_t \rho_- \end{bmatrix} = -\partial_x \begin{bmatrix} J_{\rho_+} \\ J_{\rho_-} \end{bmatrix} + \begin{bmatrix} -K \\ K \end{bmatrix}. \tag{A.9}
$$

The hydrodynamic fluxes are found by coarse-graining, using the hydrodynamic coordinates $x = \ell_s i / L$ and $t = \gamma \hat{t} / L^2$, with $\hat{t}$ the time variable of the microscopic dynamics:

$$
\begin{aligned}
J_{\rho_+} &= -\partial_x \rho_+ - \beta \rho_+ \partial_x \left(\frac{\delta H}{\delta \rho} + \frac{\delta H}{\delta m} \right) + \text{Pe}\,\rho_+, \\
J_{\rho_-} &= -\partial_x \rho_- - \beta \rho_- \partial_x \left(\frac{\delta H}{\delta \rho} - \frac{\delta H}{\delta m} \right) - \text{Pe}\,\rho_-,
\end{aligned}
\tag{A.10}
$$

and, following the same steps, the tumbling rate reads

$$K = m e^{\beta \frac{\delta H}{\delta m}}. \tag{A.11}$$

Note that, when sending $\beta \to 0$ in (A.10), one recovers the fluxes of AIM [13]. Finally, one obtains the expressions (6) and (7) in terms of ρ and m with a few steps of algebra. Interestingly, in the equilibrium limit (Pe = 0), the free energy F serves as a Lyapunov function:

$$
\frac{dF}{dt} = -\frac{L}{\ell_s} \int_0^{\ell_s} dx \left\{ \frac{1}{2} \partial_x \begin{bmatrix} \frac{\delta F}{\delta \rho} \\ \frac{\delta F}{\delta m} \end{bmatrix}^\dagger \mathbb{C}(\rho, m) \partial_x \begin{bmatrix} \frac{\delta F}{\delta \rho} \\ \frac{\delta F}{\delta m} \end{bmatrix} + 2M \frac{\delta F}{\delta m} \sinh\left(\frac{\delta F}{\delta m} \right) \right\} \leqslant 0, \tag{A.12}
$$

where we have used that $M > 0$ and $\mathbb{C}$ is positive semi-definite. The case $dF/dt = 0$ is achieved in steady state. In the presence of noise, this case corresponds to the system eventually relaxing towards the minimum $\delta F / \delta \rho = \delta F / \delta m = 0$. Notice, however, that the well-posedness of this hydrodynamic description must break in the equilibrium limit Pe = 0 for any non-convex free energy $\mathcal{F}$ (7). It is cured by retaining sub-leading interfacial terms $\mathcal{O}(L^{-2}|\nabla \rho|^2, L^{-2}|\nabla m|^2)$.

Appendix B. Spinodals and binodals in the equilibrium limit

As mentioned in section 3.1, the density spinodal (15), which is complementary to the magnetization one (13), marks the boundary of the convex region of the single-variable function $\mathcal{F}[\rho, m_0(\rho, T)]$ [42]. Indeed, this boundary is given by

$$\frac{\partial^2 \mathcal{F}}{\partial \rho^2}\bigg|_{\rho_0, m_0(\rho, T)} + \frac{\partial^2 \mathcal{F}}{\partial \rho \partial m}\bigg|_{\rho_0, m_0(\rho, T)} \frac{\partial m_0(\rho, T)}{\partial \rho}\bigg|_{\rho_0} = 0, \tag{B.1}$$

where we used the fact that, in the ordered state $m_0(\rho_0, T)$, we have

$$\frac{\partial \mathcal{F}}{\partial m}\bigg|_{\rho_0, m_0(\rho, T)} = 0. \tag{B.2}$$

Differentiating (B.2) with respect to ρ yields

$$\frac{\partial m_0(\rho, T)}{\partial \rho}\bigg|_{\rho_0} = -\left(\frac{\partial^2 \mathcal{F}}{\partial \rho \partial m} \bigg/ \frac{\partial^2 \mathcal{F}}{\partial^2 m}\right)\bigg|_{\rho_0, m_0(\rho_0, T)}. \tag{B.3}$$

Plugging the relation (B.3) into (B.1), it follows that finding the density spinodal is equivalent to locating the onset of local non-convexity of the two-variable free energy:

$$|\text{Hess}(\mathcal{F})|_{\rho_0, m_0(\rho_0, T)} = 0. \tag{B.4}$$

The density spinodal can be expressed in an implicit parametric form, since all partial derivatives entering (B.4) are explicit functions of ρ and m. Indeed, we can use the parametric representation (14)

$$z = \frac{m_0}{\rho_0}, \quad T(z, \rho_0) = \frac{z}{\text{arctanh}(z)} f(\rho_0), \tag{B.5}$$

to explicitly write (B.4) as

$$\left(\frac{1}{1 - z^2} - \frac{\text{arctanh}(z)}{z}\right)\left[1 - \frac{x\,\text{arctanh}(z)}{2} \frac{(\text{d}/\text{d}\rho)\left(\rho^2 f'(\rho)\right)}{f(\rho)}\right] = \text{arctanh}^2(z)\left(\frac{\rho f'(\rho)}{f(\rho)}\right)^2, \tag{B.6}$$

which gives $z = z(\rho_0)$ in an implicit form, from which follows the curve $T(\rho_0)$ for the density spinodal. This procedure allows us to determine the corresponding curves in figures 2(a), 4(a) and 5(a).

The binodals are found by the common-tangent construction over the single-variable free energy $\mathcal{F}[\rho, m_0(\rho_0, T)]$, which can result in several possible phase-diagram topologies [42]. The binodals correspond to the coexistence between dense and dilute phases. For monotonic $f(\rho)$ in section 3, it is always the dense and the dilute phases which are respectively ordered ($m = m_0(\rho, T)$) and disordered ($m = 0$). Yet, one can have the opposite case for non-monotonic $f(\rho)$, as explained in section 4. For illustration purposes, considering here the case of monotonic $f(\rho)$, the condition for common tangent is given by

$$\mathcal{F}[\rho_l, m_0(\rho_l, T)] = \mathcal{F}(\rho_g, m = 0) + \partial_\rho \mathcal{F}(\rho_g, m = 0)(\rho_l - \rho_g),$$
$$\partial_\rho \mathcal{F}[\rho_l, m_0(\rho_l, T)] = \partial_\rho \mathcal{F}(\rho_g, m = 0), \tag{B.7}$$

which explicitly reads

$$\frac{\rho_l}{2}\left[\ln\frac{\rho_l^2(1 - z^2)}{4} + z\ln\frac{1 + z}{1 - z}\right] - \rho_l - \beta\frac{\rho_l z^2}{2} f(\rho_l) = \rho_g\left(\log\frac{\rho_g}{2} - 1\right) + \log\frac{\rho_g}{2}(\rho_l - \rho_g),$$

$$\log\left(\frac{\rho_l}{2}\sqrt{1 - z^2}\right) - \frac{\beta z^2}{2}[\rho_l f'(\rho_l) - f(\rho_l)] = \log\frac{\rho_g}{2}, \tag{B.8}$$

where we have used the parametric representation $z = m_0(\rho_l, T)/\rho_l$. Substituting the temperature relation (B.5) into (B.8), we get

$$1 - \frac{\rho_g}{\rho_l} = \frac{z\,\text{arctanh}\,z}{2}\frac{\rho_l f'(\rho_l)}{f(\rho_l)}, \qquad \frac{\rho_g}{\rho_l} = \sqrt{1 - z^2}\, e^{-\frac{z\,\text{arctanh}\,z}{2}\left[\frac{\rho_l f'(\rho_l)}{f(\rho_l)} - 1\right]} \tag{B.9}$$

which finally gives the two binodals in an implicit parametric representation for $\rho_l(z)$, $\rho_g(z)$ and $T = T(z)$. The tricritical point is defined as the meeting point of the two binodals (B.9), which is also the meeting point of the two spinodals (B.6) and (13). This occurs at vanishing $z = m_0/\rho_0$. Indeed, expanding either one of these expressions at small x leads to (16).

Appendix C. Linear instability in the nonequilibrium case

To derive the expression of the density spinodal in (18), we examine the linear stability of the dynamics (5) and (6) around the homogeneous solutions $\rho(x,t) = \rho_0 + \delta\rho(x,t)$ and $m(x,t) = m_0(\rho_0, T) + \delta m(x,t)$, where m_0 is given by (14). Using the Fourier convention for an arbitrary field X:

$$\delta X(x,t) = \sum_n \delta X_n(t)\, e^{ik_n x}, \quad k_n = 2\pi n/\ell_s, \tag{C.1}$$

we arrive at the following mode dynamics

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} \delta\rho_n \\ \delta m_n \end{bmatrix} = \mathbb{M} \begin{bmatrix} \delta\rho_n \\ \delta m_n \end{bmatrix}, \tag{C.2}$$

where the stability matrix $\mathbb{M}$ reads

$$\mathbb{M} = \begin{bmatrix} -k^2 A & -k^2 B - \mathrm{Pe}\, ik \\ -k^2 C - \mathrm{Pe}\, ik - 2\alpha_2 & -k^2 D - 2\alpha_3 \end{bmatrix} \Bigg|_{\rho_0, m_0(\rho_0, T)} \tag{C.3}$$

with

$$\begin{aligned}
A &= \rho\partial_{\rho\rho}\mathcal{F} + m\partial_{\rho m}\mathcal{F}, & B &= \rho\partial_{\rho m}\mathcal{F} + m\partial_{mm}\mathcal{F}, & \alpha_2 &= M\partial_{\rho m}\mathcal{F}, \\
C &= m\partial_{\rho\rho}\mathcal{F} + \rho\partial_{\rho m}\mathcal{F}, & D &= m\partial_{\rho m}\mathcal{F} + \rho\partial_{mm}\mathcal{F}, & \alpha_3 &= M\partial_{mm}\mathcal{F},
\end{aligned} \tag{C.4}$$

and $M(\rho, m) = \sqrt{\rho^2 + m^2}$. The eigenvalues of (C.3) are of the form

$$\Lambda_{1,2} = \frac{1}{2}\left[\mathrm{Tr} \pm \sqrt{\mathrm{Tr}^2 - 4\mathrm{Det}} \right], \tag{C.5}$$

where

$$\mathrm{Tr} = -\left[2\alpha_3 + k^2(A + D) \right], \quad \mathrm{Tr}^2 - 4\mathrm{Det} = a^2 + ikb + k^2 c + \mathcal{O}(k^3), \tag{C.6}$$

with

$$a = 2\alpha_3, \quad b = 8\mathrm{Pe}\,\alpha_2, \quad c = 4\left[\alpha_3(D - A) + 2B\alpha_2 - \mathrm{Pe}^2 \right], \tag{C.7}$$

yielding

$$\mathrm{Re}\left[\sqrt{\mathrm{Tr}^2 - 4\mathrm{Det}} \right] = a + \frac{k^2}{a}\left(\frac{b^2}{8a^2} + \frac{c}{2} \right) + \mathcal{O}(k^4), \quad \mathrm{Im}\left[\sqrt{\mathrm{Tr}^2 - 4\mathrm{Det}} \right] = k\frac{b}{2a} + \mathcal{O}(k^3), \tag{C.8}$$

where we used that $\alpha_3 \propto \partial_{mm}\mathcal{F} > 0$ at the magnetized state $[\rho_0, m_0(\rho_0, T)]$. (Note that the expressions a, b, c in (C.7) are not to be confused with the Model C coefficients (11).) Then, the eigenvalue with the larger real part is given by

$$\Lambda_+ = ik\mathrm{Pe}\frac{\partial_{\rho m}\mathcal{F}}{\partial_{mm}\mathcal{F}} + k^2\left\{ \frac{\mathrm{Pe}^2}{2M\mathcal{F}_{mm}}\left[\left(\frac{\partial_{\rho m}\mathcal{F}}{\partial_{mm}\mathcal{F}}\right)^2 - 1 \right] + \frac{\rho_0\mathrm{Hess}(\mathcal{F})}{\mathcal{F}_{mm}} \right\}. \tag{C.9}$$

The condition of vanishing real part for the eigenvalue (C.9) yields the ordered spinodal (18).

Appendix D. Details on numerical solution of the hydrodynamics

To numerically solve the hydrodynamics (5) and (6) we employ a finite difference scheme. We solve it on a segment of length $\ell_s = 500$ with periodic boundary conditions. The interfacial width is $\mathcal{O}(1)$ and the resulting traveling wave profiles have well resolved bulk phases as seen in figures 2(c), 4(c) and 5(c). The profiles were initiated at a homogeneous state with small sinusoidal perturbation. We have confirmed that the final traveling wave profile at steady state is robust for other choices of initial conditions. For some of these the relaxation to steady state proceeded via a traveling wave with multiple phase boundaries which coarsen over time towards a single phase-separated profile with sharp interfaces.

Generic values for space and time discretization are $\mathrm{d}x = 0.1$ and $\mathrm{d}t = 0.001$. However, at low temperatures the interfaces become increasingly sharper and finer meshes were chosen. Indeed, in the limit $\beta \to \infty$ the free energy (7) diverges. Thus, the dynamics is dominated by the equilibrium free energy

derivatives entering (6). As noted in appendix A, the equilibrium hydrodynamics becomes ill-defined for any non-convex free energy $\mathcal{F}$ since this free energy is missing the standard interfacial energy terms and the width of the interfacial profile will vanish. The self propulsion terms in (6) cure this and assign a finite interface width, but this decreases with decreasing temperature and is hard to resolve numerically. Luckily, due to the same reasoning, close to vanishing temperatures the binodals of the non-equilibrium phase diagram Pe > 0 must approach those of the equilibrium Pe = 0 limit. As explained in appendix B, the latter are determined analytically without the need of solving the hydrodynamics (5) and (6), which for the equilibrium case are ill-defined. We use these equilibrium asymptotics to extend the binodal curves of the non-equilibrium phase diagrams all the way to zero temperature (the numerically computed curves at small temperatures are found to match them smoothly).

ORCID iDs

Robert L Jack ⬤ https://orcid.org/0000-0003-0086-4573
Étienne Fodor ⬤ https://orcid.org/0000-0003-1372-2195

References

[1] Toner J, Tu Y and Ramaswamy S 2005 *Ann. Phys., NY* **318** 170–244
[2] Chaté H 2020 *Annu. Rev. Condens. Matter Phys.* **11** 189–212
[3] Cavagna A and Giardina I 2014 *Annu. Rev. Condens. Matter Phys.* **5** 183–207
[4] Bain N and Bartolo D 2019 *Science* **363** 46–49
[5] Bricard A, Caussin J B, Desreumaux N, Dauchot O and Bartolo D 2013 *Nature* **503** 95–98
[6] Vicsek T, Czirók A, Ben-Jacob E, Cohen I and Shochet O 1995 *Phys. Rev. Lett.* **75** 1226–9
[7] Toner J and Tu Y 1995 *Phys. Rev. Lett.* **75** 4326–9
[8] Marchetti M C, Joanny J F, Ramaswamy S, Liverpool T B, Prost J, Rao M and Simha R A 2013 *Rev. Mod. Phys.* **85** 1143–89
[9] Mermin N D and Wagner H 1966 *Phys. Rev. Lett.* **17** 1133
[10] Halperin B 2018 *J. Stat. Phys.* **175** 521–9
[11] Grégoire G and Chaté H 2004 *Phys. Rev. Lett.* **92** 025702
[12] Solon A P and Tailleur J 2015 *Phys. Rev.* E **92** 042119
[13] Solon A P and Tailleur J 2013 *Phys. Rev. Lett.* **111** 078101
[14] Solon A P, Caussin J B, Bartolo D, Chaté H and Tailleur J 2015 *Phys. Rev.* E **92** 062111
[15] Kourbane-Houssene M, Erignoux C, Bodineau T and Tailleur J 2018 *Phys. Rev. Lett.* **120** 268003
[16] Scandolo M, Pausch J and Cates M E 2023 *Eur. Phys. J.* E **46** 103
[17] Fodor E, Nardini C, Cates M E, Tailleur J, Visco P and van Wijland F 2016 *Phys. Rev. Lett.* **117** 038103
[18] Fodor E, Jack R L and Cates M E 2022 *Annu. Rev. Condens. Matter Phys.* **13** 215–38
[19] Seifert U 2012 *Rep. Prog. Phys.* **75** 126001
[20] Ballerini M *et al* 2008 *Proc. Natl Acad. Sci. USA* **105** 1232–7
[21] Ginelli F and Chaté H 2010 *Phys. Rev. Lett.* **105** 168103
[22] te Vrugt M, Holl M P, Koch A, Wittkowski R and Thiele U 2022 *Modelling Simul. Mater. Sci. Eng.* **30** 084001
[23] Fischer A, Chatterjee A and Speck T 2019 *J. Chem. Phys.* **150** 064910
[24] Bebon R, Robinson J F and Speck T 2024 arXiv:2401.02252
[25] Aslyamov T, Avanzini F, Fodor E and Esposito M 2023 *Phys. Rev. Lett.* **131** 138301
[26] Martin D, Chaté H, Nardini C, Solon A, Tailleur J and Van Wijland F 2021 *Phys. Rev. Lett.* **126** 148001
[27] Lebowitz J L and Spohn H 1999 *J. Stat. Phys.* **95** 333–65
[28] Bodineau T and Lagouge M 2010 *J. Stat. Phys.* **139** 201–18
[29] Landau L D and Lifšic E M 1980 The impossibility of the existence of phases in one-dimensional systems *Statistical Physics* vol 5, 3rd edn (Butterworth-Heinemann) p 537
[30] Benvegnen B, Chaté H, Krapivsky P L, Tailleur J and Solon A 2022 *Phys. Rev.* E **106** 054608
[31] Gorbonos D, Puckett J G, Van Der Vaart K, Sinhuber M, Ouellette N T and Gov N S 2020 *J. R. Soc. Interface* **17** 20200367
[32] Bastien R and Romanczuk P 2020 *Sci. Adv.* **6** eaay0792
[33] King A E B T and Turner M S 2021 *R. Soc. Open Sci.* **8** 201536
[34] Wirth T D, Dachner G C, Rio K W and Warren W H 2023 *PNAS Nexus* **2** gad118
[35] Jona-Lasinio G, Landim C and Vares M E 1993 *Prob. Theory Relat. Fields* **97** 339–61
[36] Agranov T, Ro S, Kafri Y and Lecomte V 2021 *J. Stat. Mech.* 083208
[37] Agranov T, Ro S, Kafri Y and Lecomte V 2023 *SciPost Phys.* **14** 045
[38] Agranov T, Cates M E and Jack R L 2022 *J. Stat. Mech.* 123201
[39] Bray A J 2002 *Adv. Phys.* **51** 481–587
[40] Kaiser M, Jack R L and Zimmer J 2018 *J. Stat. Phys.* **170** 1019–50
[41] Blume M, Emery V J and Griffiths R B 1971 *Phys. Rev.* A **4** 1071–7
[42] Roux D, Coulon C and Cates M 1992 *J. Phys. Chem.* **96** 4174–87
[43] Pearce D J G, Miller A M, Rowlands G and Turner M S 2014 *Proc. Natl Acad. Sci. USA* **111** 10422–6
[44] Koynova R D, Boyanov A I and Tenchov B G 1987 *Biochim. Biophys. Acta Biomembr.* **903** 186–96
[45] Shim P S, Chun H M and Noh J D 2016 *Phys. Rev.* E **93** 012113
[46] Crosato E, Prokopenko M and Spinney R E 2019 *Phys. Rev.* E **100** 042613
[47] Ferretti F, Grosse-Holz S, Holmes C, Shivers J L, Giardina I, Mora T and Walczak A M 2022 *Phys. Rev.* E **106** 034608

[48] Yu Q and Tu Y 2022 *Phys. Rev. Lett.* **129** 278001
[49] Dadhichi L P, Maitra A and Ramaswamy S 2018 *J. Stat. Mech.* 123201
[50] Borthne Ø L, Fodor E and Cates M E 2020 *New J. Phys.* **22** 123012
[51] Markovich T, Fodor E, Tjhung E and Cates M E 2021 *Phys. Rev. X* **11** 021057
[52] Falasco G and Esposito M 2023 arXiv:2307.12406
[53] Proesmans K *et al* in preparation