This is life: some thoughts on self-organized structure formation in active liquids and biological systems

Hoffmann, L.A.

CHAPTER 1

Introduction to Liquid Crystals, Active Matter, and Differential Geometry

The goal of this thesis is to understand different aspects of self-organization that result in a spontaneous transition to order out of disorder. With no, or only minimal, external guidance, a large range of systems is found to structure itself in a reproducible fashion, creating large-scale structures and order out of seemingly random, microscopic chaos. Examples range from bacterial colonies, to flocks of birds, and human decision making processes. We use tools from elasticity theory, hydrodynamics, statistical physics, and differential geometry to built and analyze minimal models to learn about fundamental principles that drive these dynamics. In particular, we are interested in general laws that are independent of microscopic details. Rather than accurately describing a certain, specific system, we want to find rules that govern the dynamics of as broad a range of systems as possible. Analytical calculations are combined with simulations and experiments to guarantee that the models analyzed are realistic and predictive. Often biological systems are motivating the research questions presented in this thesis and we try to explain certain biological processes, like tissue dynamics.

In the dynamic space of the living Rocket, the double integral has a different meaning. To integrate here is to operate on a rate of change so that time falls away: change is stilled .... "Meters per second" will integrate to "meters." The moving vehicle is frozen, in space, to become architecture, and timeless. It was never launched. It will never fall.

T. Pynchon. *Gravity's Rainbow.*

To be a physicist in A-Io was to serve not society, not mankind, not the truth, but the State.

U. K. Le Guin. *The Dispossessed.*

We need a constant. You always need a constant. [...] Let d be any constant, for computational reasons the closer to 1 the better.

D. F. Wallace. *Infinite Jest.*

1

In this chapter we introduce some of the basic concepts that will be used in later chapters. First, in Sec. 1.1, we introduce the idea of a liquid crystal. These are systems that consist of rod-like particles which tend to align with each other, thus giving rise to an orientational order and an overall preferred direction. While the system would like to be perfectly aligned (all the rods pointing in the same direction) this is not always possible. So-called topological defects are particle-like distortions of the aligned state that are stable in time. After introducing these structures, we turn towards nematodynamics. We describe how liquid crystals behave in the presence of a flow field. Loosely speaking, what are the equations that should be used to describe the dynamics of a small branch that is thrown into a flowing river. In Sec. 1.2 active matter, a broad field of research that has recently attracted a lot of attention, is introduced. The overarching idea is to describe systems that are inherently out of equilibrium because its microscopic constituents are injecting energy into the system. An example are systems where each of the microscopic particles consumes energy and is motile on its own. We introduce three different models to describe different kinds of active systems. First, the active nematodynamic equations, which are the generalization of the passive nematodynamic equations for the case that each of the microscopic rods creates a microscopic flow. Thus, the flow field is now no longer externally applied as before, but created by the system itself. In the picture above, "branches" are thrown into an initially still river and these "branches" move the water such that the river starts flowing. We present results on how the presence of activity modifies the dynamics of topological defects. We furthermore introduce the Toner-Tu equations and the Vicsek model, both of which have been used to describe the flocking behavior of different kinds of systems, for example, flocks of birds or animal herds. The basic idea is that each particle (think bird) has a velocity and every particle (bird) tends to align with nearby neighbors. This can result in all the particles (birds) in the system moving in the same direction after some time. This process is the result of self-organization, that is no external guidance is needed for the system to organize. Lastly, in Sec. 1.3, we introduce some concepts from differential geometry. This is the mathematical language used to describe curved surfaces. We close this chapter with an outline of the following chapters of this thesis.

1.1 Passive Liquid Crystal

A liquid crystal is, as the name suggests, a state between a crystal and a liquid. A crystal is characterized by a regular arrangement of its microscopic constituents, meaning they are located on a regular lattice. If one then stands at point $\boldsymbol{x}$ in space and looks at another point $\boldsymbol{x}'$ far away, the density-density correlation function between these two points is a periodic function,

$$\lim_{|\boldsymbol{x}-\boldsymbol{x}'|\to\infty} \big\langle \rho(\boldsymbol{x})\rho(\boldsymbol{x}') \big\rangle = \text{ periodic in } \boldsymbol{x}-\boldsymbol{x}' \,. \tag{1.1}$$

This captures the fact that an infinitely large crystal is just an infinite amount of copies of the same basis and neighboring lattice points always have the same distance independent of where in the crystal one looks, see Fig. 1.1a. On the other

hand, for a liquid (Fig. 1.1b) this symmetry is completely broken and the position of two particles far away from each other is "minimally correlated" in the sense that one can express the probability of finding a particle at some given position far away from another only through the average particle density, that is

$$\lim_{|\boldsymbol{x}-\boldsymbol{x}'|\to\infty} \big\langle \rho(\boldsymbol{x})\rho(\boldsymbol{x}') \big\rangle = (\text{Average particle density})^2 \ . \tag{1.2}$$

Now, a liquid crystal is an intermediate state in the sense that the system has liquid-like order in at least one direction of space but where translational or rotational symmetry is still broken (like a crystal). The liquid crystal is characterized by how much symmetry is broken, i.e., how "crystalline" or "liquid" a system is. If the system is translationally invariant but not rotationally invariant, it is called a nematic liquid crystal if the system is still top-down symmetric, see Fig. 1.1c. If positional order is imposed in one or two dimensions, it is called smectic or columnar liquid crystal, respectively (Fig. 1.1d). If the system is not top-down symmetric, it is called a polar liquid crystal (Fig. 1.1e). We will only be concerned with liquid crystals that are translationally invariant (i.e., nematic or polar liquid crystals) and consider only this case in the following.

An uniaxial nematic liquid crystal is typically made up of elongated particles that have some direction, e.g., rods or rice grains, see Fig. 1.1c. Their centre of mass has no long-range order (thus the system is translationally invariant) but they tend to be aligned in the same direction (thereby breaking rotational symmetry). The average orientation is captured by defining a unit vector, the so-called director $\boldsymbol{p}$. In the nematic phase the particles are top-down symmetric such that every quantity describing the system is symmetric under $\boldsymbol{p} \to -\boldsymbol{p}$, see Fig. 1.1f. A biaxial nematic liquid crystal is made up of particles that have two distinct orientations (e.g., rectangles) such that in the corresponding ordered state particles are aligned along two directions. Finally, in a polar liquid crystal particles are not top-down symmetric and have a distinct orientation, i.e., they are not symmetric under $\boldsymbol{p} \to -\boldsymbol{p}$, see Fig. 1.1e,g. However, note that the kind of liquid crystal one deals with is determined by the macroscopic symmetry, not the microscopic symmetry of the particles. If top-down asymmetric particles align in a nematic fashion, that is for the alignment the particles' top-down asymmetry is irrelevant (Fig. 1.1h), this results in a nematic liquid crystal [1–4].

1.1.1 Order parameter and isotropic-nematic transition

To quantify the amount of alignment one computes an order parameter. This quantity vanishes in the liquid phase and is unity if all particles are aligned in the same direction. To define this order parameter that characterizes the liquid crystal phase one assigns to each particles an orientation or a direction $\boldsymbol{a}$, according to whether one is interested in the nematic or polar phase, see Fig. 1.1i. One then averages over all of the particles' orientations (directions) $\boldsymbol{a}$ which gives rise to a macroscopic orientation (direction) $\boldsymbol{p}$ that describes the average orientation (direction) in the system. To quantify how strongly the particles are aligned with this average orientation one then computes the order parameter by taking the average over the angle

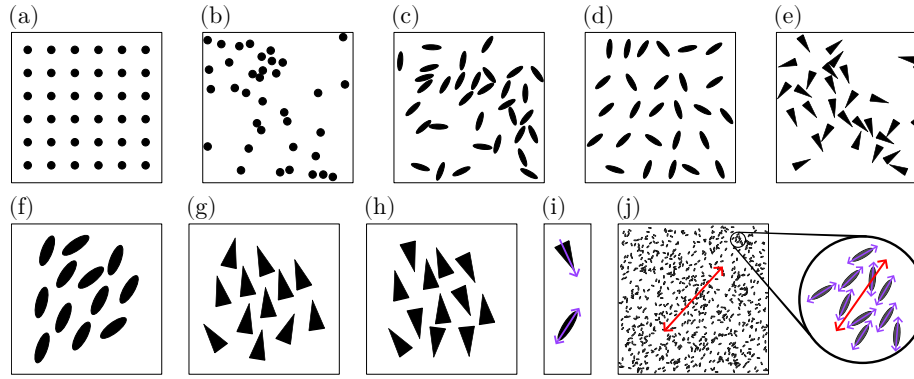

Figure 1.1: Liquid crystal.

(a) Sketch of a regular crystal lattice where particles are periodically arranged in space. (b) Sketch of a fluid where particles are randomly oriented in space. (c) Sketch of a nematic liquid crystal where the position of the center of mass of particles is random in space but they have a preferred orientation. (d) Sketch of a smectic liquid crystals where particles have a preferred orientation and, additionally, are forming layers. (e) Sketch of a polar liquid crystal where particles have a preferred direction but no preferred location. (f) Example of particles with top-down symmetry forming a nematic phase. (g) Example of particles with a preferred direction (broken top-down symmetry) forming a polar phase. (h) Example of particles with a preferred direction forming a nematic phase, their polar direction is irrelevant and particles only have a preferred orientation. (i) Sketch of a polar director (top) and nematic director (bottom) assigned to a particle. (j) Sketch of the coarse-graining procedure where each particle is assigned an orientation and a macroscopic director is defined from this by averaging over the orientation of all particles.

θ each of the particles makes with the average director $\boldsymbol{p}$ to obtain a single number to quantify the order in the phase. We follow Ref. [3] here. Choosing the coordinate system in three dimensions such that $\boldsymbol{p} = \{0, 0, 1\}$, one can write the orientation of each particle in spherical coordinates as

$$\boldsymbol{a} = \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix} . \tag{1.3}$$

A normalized distribution function $\mathcal{P}_3(\theta, \varphi)\mathrm{d}\Omega_3$ describes the probability for a given particle to have an orientation in an infinitesimal solid angle $\mathrm{d}\Omega_3 = \sin\theta\mathrm{d}\theta\mathrm{d}\varphi$. For an uniaxial nematic or polar liquid crystal the distribution function is independent of φ and it can be integrated out, $\mathcal{P}_3(\theta, \varphi) = \mathcal{P}_3(\theta)/2\pi$. In two dimensions, instead, we choose $\boldsymbol{p} = \{1, 0\}$,

$$\boldsymbol{a} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} , \tag{1.4}$$

and $\mathcal{P}_2(\theta, \varphi)\mathrm{d}\Omega_2 = \mathcal{P}_2(\theta)\mathrm{d}\Omega_2 = \mathcal{P}_2(\theta)\mathrm{d}\theta$. For a polar liquid crystal one can simply compute the average angle between each particle and the director $\boldsymbol{p}$. This defines the polar order parameter:

$$P := \langle \boldsymbol{a} \cdot \boldsymbol{p} \rangle = \langle \cos\theta \rangle = \int \cos\theta \mathcal{P}_d(\theta)\mathrm{d}\Omega_d . \tag{1.5}$$

For an isotropic system $\mathcal{P}_2(\theta) = 1/\mathcal{N}_2 = 1/2\pi$ and $\mathcal{P}_3(\theta) = 1/\mathcal{N}_3 = 1/2$ such that $\langle \cos\theta \rangle = 0$ and the order parameter vanishes. On the other hand, for a perfectly aligned system $\mathcal{P}(\theta) = \delta(\theta)/\mathcal{N}_d$ (that is $\theta = 0$ and all particles point in the same direction) such that $P = 1$.

For a nematic liquid crystal, however, the dipole contribution $\langle \boldsymbol{a} \cdot \boldsymbol{p} \rangle$ vanishes identically due to the top-down symmetry of the nematic since in this case $\mathcal{P}_d(\theta) = \mathcal{P}_d(\theta - \pi)$. Thus, one needs to consider the next-order term in the multipole expansion, $\langle (\boldsymbol{a} \cdot \boldsymbol{p})^2 \rangle$, which is non-vanishing. The nematic order parameter S in d dimensions is defined as

$$S := \frac{1}{d-1} \left\langle d\cos^2\theta - 1 \right\rangle . \tag{1.6}$$

Here, an overall prefactor is introduced and a constant is subtracted from $\langle (\boldsymbol{a} \cdot \boldsymbol{p})^2 \rangle$ in order to normalize the order parameter such that $S = 0$ in the isotropic case, and $S = 1$ in the perfectly aligned case. Indeed, in the isotropic case $\langle \cos^2\theta \rangle = 1/d$, and $S = 0$, while in the aligned case $\cos\theta = \pm 1$ and $S = 1$. These are the scalar order parameters for the polar and nematic phase, respectively. Note that for a polar liquid crystal $P = S$, while for a nematic liquid crystal the polar order parameter vanishes identically, $P = 0$. Apart from the scalar order parameters, there is the vectorial order parameter $\boldsymbol{p}$ determining the average orientation, and these two quantities together completely characterize the nematic and polar order present in a given system. However, since in the nematic phase there is an average orientation, but not an average direction, it is often convenient to define a quantity that takes

the top-down symmetry into account by definition, instead of stating that vectorial director $\boldsymbol{p}$ is to be viewed as symmetric under $\boldsymbol{p} \to -\boldsymbol{p}$. A simple way of achieving this is by considering $\boldsymbol{pp}$ instead of $\boldsymbol{p}$, as the former is inherently symmetric under $\boldsymbol{p} \to -\boldsymbol{p}$. Commonly, the tensorial order parameter constructed from this, the so-called $\boldsymbol{Q}$-tensor, is defined as

$$\boldsymbol{Q} := S\left(\boldsymbol{pp} - \frac{1}{d}\mathbb{1}\right) , \tag{1.7}$$

for uniaxial nematic liquid crystals. Here, $\mathbb{1}$ is the unit matrix which has been subtracted to make the tensor $\boldsymbol{Q}$ traceless. Note that for a biaxial nematic liquid crystal one still defines a symmetric, traceless $\boldsymbol{Q}$-tensor but its expression in terms of $\boldsymbol{p}$ is different. For uniaxial nematic liquid crystals one can use $\boldsymbol{Q}$ and $\boldsymbol{p}$ interchangeably, if one keeps the top-down symmetry requirement for the latter in mind while using it. Having defined an order parameter, one can then study the phase transition between an isotropic and an ordered phase in terms of this order parameter. For example, in the Landau-de Gennes approach for the isotropic-nematic transition, a free energy is written as a power series of the order parameter S,

$$F_{\mathrm{LdG}} = aS^2 + bS^3 + cS^4 , \tag{1.8}$$

with a, b, and c temperature-dependent numerical prefactors, such that, due to the presence of the cubic term, the transition from the isotropic ($S = 0$) to the nematic ($S \neq 0$) phase is seen to be first order. We do not consider this further, see, e.g., Refs. [1–7] for details. In the following sections and chapters we always assume that a liquid crystal phase is present. In fact, unless otherwise noted, we set $P = S = 1$. Thus, we are working with a highly order liquid crystal phase and often study perturbations of this state. Furthermore, we will denote by $\boldsymbol{p}$ the director field for both polar and nematic liquid crystals, from context it will be apparent whether $\boldsymbol{p}$ is inversion symmetric or not.

1.1.2 Frank free energy

So far we have only talked about the symmetry of the liquid crystal and how to quantify it. We now turn towards studying the dynamics of a liquid crystal. We first consider the case of a nematic liquid crystal. It is common to assign to a nematic liquid crystal a distortion free energy that vanishes if the nematic field is constant and can be expanded in powers of $\nabla\boldsymbol{p}$. This energy thus captures the energetic cost associated with a configuration where nearby particles are not aligned which would result in non-vanishing gradients of the director field. Assuming that the variations in the director field are slow it suffices to look at the lowest order terms in the free energy. The final free energy must be symmetric under $\boldsymbol{p} \to -\boldsymbol{p}$ and $\boldsymbol{x} \to -\boldsymbol{x}$. Furthermore, terms that are total derivatives can be ignored in the description of the bulk since they can be absorbed into a boundary term thanks to Stokes' theorem and $|\boldsymbol{p}|^2 = 1$. Due to these requirements there are no terms of linear order in $\nabla\boldsymbol{p}$. Thus, at lowest order in $\nabla\boldsymbol{p}$, the distortion free energy is of the form

$$F_{\mathrm{F}} = \int \mathrm{d}A\, \kappa_{\mathrm{F},ijkl} \nabla^i p^j \nabla^k p^l , \tag{1.9}$$

where $\kappa_{\mathrm{F},ijkl}$ is an unknown tensor that will, in general, depend on the local director field and material properties. It is possible to show that, due to the symmetry requirements, there are in the end only three independent terms such that the distortion free energy can be written as [1, 3]

$$F_{\mathrm{F}} = \int \mathrm{d}A \frac{\kappa_{\mathrm{F},1}}{2} \left(\nabla \cdot \boldsymbol{p}\right)^2 + \frac{\kappa_{\mathrm{F},2}}{2} \left(\boldsymbol{p} \cdot \nabla \times \boldsymbol{p}\right)^2 + \frac{\kappa_{\mathrm{F},3}}{2} \left(\boldsymbol{p} \times \nabla \times \boldsymbol{p}\right)^2 . \tag{1.10}$$

This elastic free energy of the nematic liquid crystal is called Frank energy. The material constants $\kappa_{\mathrm{F},i}$ have units of energy/length. The three terms of the free energy are the contributions from splay, twist, and bend deformations of the nematic field, respectively [1–3, 8]. These are visualized in Fig. 1.2. Note that in two dimensions the second term of the free energy (twist) always vanishes identically. In many instances this free energy is still too complex to handle and resulting equations are difficult to solve. Hence, to simplify analytical calculations one often considers the so-called one-elastic-constant approximation where the energetic difference between the different deformation modes is assumed to be negligible, $\kappa_{\mathrm{F},1} \approx \kappa_{\mathrm{F},2} \approx \kappa_{\mathrm{F},3} \approx \kappa_{\mathrm{F}}$. While this might not be quantitatively correct for many materials [3] it is often sufficient to gain qualitative insight. In this case the free energy simplifies (up to surface terms) to

$$F_{\mathrm{F}} := \frac{\kappa_{\mathrm{F}}}{2} \int \mathrm{d}A \left|\nabla \boldsymbol{p}\right|^2 . \tag{1.11}$$

The equations describing the equilibrium of the bulk are found from minimizing the free energy with respect to variations of the director field. From the resulting Euler-Lagrange equation we define the so-called molecular field $\boldsymbol{M}$ that is found to be

$$\boldsymbol{M} := -\frac{\delta F_{\mathrm{F}}}{\delta \boldsymbol{p}} = -\frac{\partial F_{\mathrm{F}}}{\partial \boldsymbol{p}} + \nabla \cdot \left(\frac{\partial F_{\mathrm{F}}}{\partial (\nabla \boldsymbol{p})}\right) = \kappa_{\mathrm{F}} \Delta \boldsymbol{p} . \tag{1.12}$$

In equilibrium the director and the molecular field must be parallel at each point.

A polar liquid crystal is similar to the nematic liquid crystal except that the inversion symmetry $\boldsymbol{p} \to -\boldsymbol{p}$ is broken. Thus, instead of rods the particles are usually sketched as arrows with a distinct head and tail. Due to this, terms that were previously ruled out because they would break the inversion symmetry of the nematic liquid crystal, are now allowed and should be added, e.g., to the Frank free energy, to describe a polar liquid crystal completely. However, in later chapters we will mostly assume that these terms are of sub-leading order and that the main dynamics can be captured by the nematic terms alone. Thus, we will use the same distortion free energy, Eq. (1.10), for both nematic and polar liquid crystals.

1.1.3 Defects in liquid crystals

It is not always possible for the system to be perfectly ordered and gradients in the orientation field result in a non-vanishing elastic free energy. There are two fundamentally different ways disorder can be introduced in the system. The first is by continuous deformation of an initially perfectly aligned configuration. This

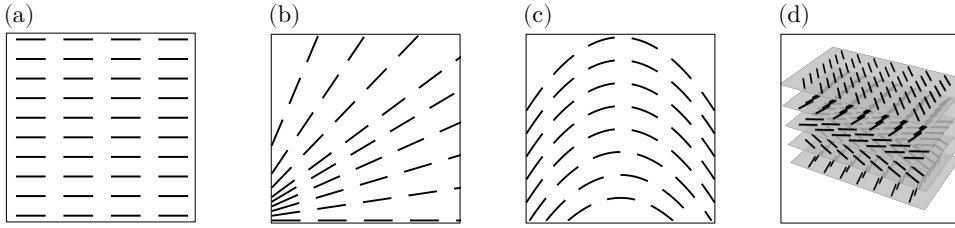

Figure 1.2: Splay, bend, and twist.

(a) A perfectly aligned nematic liquid crystal in two dimensions with no gradients in the director field. (b) Sketch of a splay deformation of the liquid crystal in two dimensions, the divergence of the director field is non-vanishing. (c) Sketch of a bend deformation in two dimensions, the projection of the curl onto the direction of the director is non-vanishing. (d) Sketch of the twist deformation in three dimensions, as there is no twist deformation in two dimensions. The director is rotated/twisted in a given two-dimensional cross-section relative to the director field in planes above and below.

deformation can be due to, for example, externally applied forces (e.g., an external shear flow or magnetic field), or thermal fluctuations. The second way is by creating so-called topological defects. These are fundamentally different in that it is not possible to continuously deform the defect configuration into an ordered configuration. Thus, since these structures are stable under continuous deformations, they are called topological. Defects can be present in a system, e.g., due to boundary conditions or they can be created through externally applied forces. However, due to their stability under continuous deformations, there are strict rules on how defects can be created and decay. To introduce these, we first introduce the concept of the topological charge. In the following subsection we introduce the concepts intuitively in two dimensions. Afterwards, we present a more rigorous definition in both two and three dimensions. Lastly, we compute the elastic energy of a defect.

Defect charge from line integral

As defects are stable against fluctuations and small deformations, they can often be considered as particle-like excitations. Similar to fundamental particles, they also carry a charge, their topological charge. Defects of different charges are distinct in that, given a defect with charge s_1 and one with charge s_2, it is impossible to continuously transform the defect of charge s_1 into the defect with charge s_2. The charge is defined as the number of rotation the director field undergoes as one tracks it along a closed circle [1–3, 9–11]:

$$s := \frac{1}{2\pi} \oint_{\mathscr{C}} \mathrm{d}\theta \,. \tag{1.13}$$

Here θ is the angle that describes the orientation of the two-dimensional director field which can be written as $\boldsymbol{p} = \{\cos\theta, \sin\theta\}$, it being a two-dimensional unit vector field. If no defect is present inside the circle $\mathscr{C}$, the charge is zero. For a perfectly aligned state the angle θ is trivially constant everywhere. Thus, in particular it is

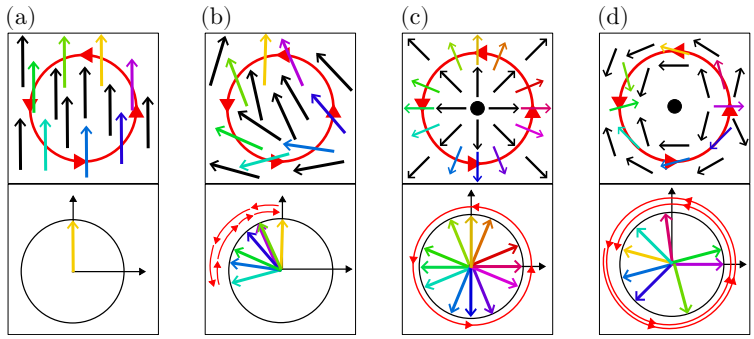

Figure 1.3: Aligned and defective configurations.

(a) In a perfectly aligned state, the director does not rotate at all when one measures the angle along a closed loop. (b) In a defect-free state without perfect alignment, as one measures the orientation along a closed loop, the director rotates by certain angle in counter-clockwise direction and then rotates back by the same angle in the clockwise direction. This traces out a closed line on the circle measuring the orientation. (c) If a $+1$ defect is present the director completes a full 2π-rotation along a closed loop encircling the defect. (d) For another defect (with charge $+2$) the director completes two full rotation, rotating by a total of 4π.

constant along any loop $\mathscr{C}$ one chooses for the contour integral, see Fig. 1.3a. If the system one considers is not perfectly ordered, then θ is not constant. However, in the case sketched in Fig. 1.3b, where the director configuration is found by slightly deforming the perfectly aligned state, the defect charge is still zero. This is because, while the angle θ varies along the contour line, the changes $\mathrm{d}\theta$ add up to zero. The director first rotates slightly counterclockwise but then exactly the same distance clockwise. However, for the configuration sketched in Fig. 1.3c this is not the case. As one travels along $\mathscr{C}$, the director continuously rotates counterclockwise, but never clockwise. Therefore, when one has traveled once along the circular contour loop, the changes in $\mathrm{d}\theta$ add up to 2π, not 0 as before; the vector $\boldsymbol{p}$ performed a full rotation. Thus, according to Eq. (1.13), this corresponds to a charge of $s = +1$. This non-zero charge reflects the fact that it is not possible to deform the configuration in Fig. 1.3c into the one in Fig. 1.3a by continuously deforming the director field. On the other hand, this is easily possible for the configuration in Fig. 1.3b which does not have a topological charge, $s = 0$. Thus, the configuration in Fig. 1.3c is a $+1$ defect. On the other hand, for the configuration in Fig. 1.3d one finds that the director field along the contour line rotates twice in the counterclockwise direction, thus this is a configuration with charge $s = +2$. It is not possible to continuously deform this configuration into either the defect-free configuration in Fig. 1.3a or the $+1$ defect in Fig. 1.3c. In this sense, defects with different charges are distinct structures. The two defect configurations presented so far had positive charge because the director field rotated in a counterclockwise direction. If the director field rotates clockwise, however, the charge is negative.

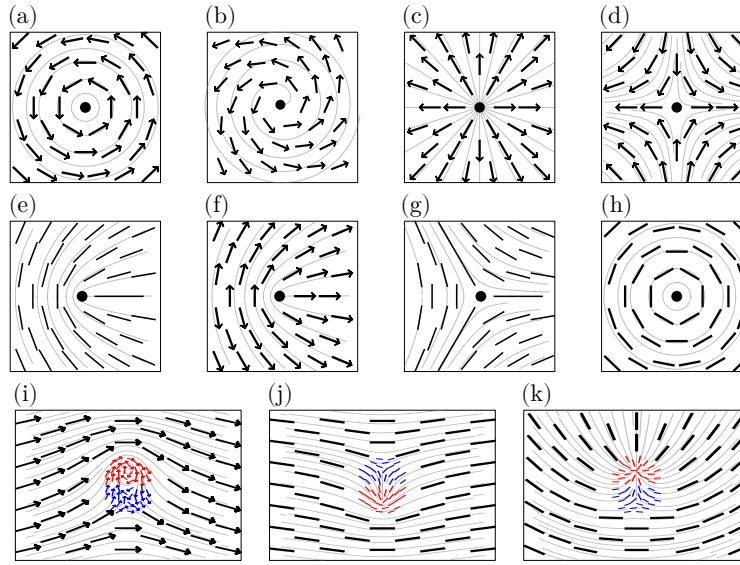

Figure 1.4: Defect charges.

(a) Vortex, (b) spiral, and (c) aster configuration of the +1 defect. They all have the same charge, i.e., in each case the director rotates by 2π (as in (Fig. 1.3c)). (d) Configuration of a -1 defect. (e) Configuration of a $+1/2$ defect. (f) $+1/2$ defects are forbidden in polar liquid crystals since they result in a discontinuity of the director field (here on the left half of the center line). (g) Configuration of a $-1/2$ defect. (h) A +1 defect is allowed for a nematic as the director field is still continuous everywhere outside the defect core at the center. (i) A +1 (red) and a -1 (blue) defect are invisible from far away and the director field is aligned there. (j) Same phenomenon but for a $+1/2$ (red) and a $-1/2$ (blue) defect. (k) A +1 (red) and a $-1/2$ (blue) defect look like a $+1/2$ defect from far away.

± 1 and $\pm 1/2$ defects

Before turning to the question of how to create and destroy defects in physical systems, we investigate the properties of some isolated defects a bit closer. From definition Eq. (1.13) of the defect charge, it is straightforward to find a parametrization for the director field $\boldsymbol{p}$ for an isolated defect of charge s, namely $\boldsymbol{p} = \cos\theta\, \hat{\boldsymbol{x}} + \sin\theta\, \hat{\boldsymbol{y}}$ with $\theta = s\varphi + \epsilon$, where φ is the angular coordinate in polar coordinates, and ϵ a constant. For a $+1$ defect the constant ϵ determines the local geometry of the defect such that the defect is an vortex (Fig. 1.4a) if $\epsilon = \pi/2 \mod \pi$, an aster (Fig. 1.4c) if $\epsilon = 0 \mod \pi$, and a spiral (Fig. 1.4b) for all values of $\epsilon = 0$ in between these two limiting cases. Note that it is possible to continuously deform these three configurations into each other since they all have the same defect charge. Thus, topologically they are the same, but energetically they could be different. If we were not working in the one-elastic-constant approximation, at equilibrium the minimal-energy configuration for a $+1$ defect is determined by the elastic constants in the Frank free energy and boundary conditions. Since we are considering a two-dimensional director field the twist-term in Eq. (1.10) vanishes. The $+1$ defects are energy-minimizing (if the boundary conditions are chosen appropriately) as asters if $\kappa_{\mathrm{F},1} < \kappa_{\mathrm{F},3}$ (splay is favored) and as vortices if $\kappa_{\mathrm{F},1} > \kappa_{\mathrm{F},3}$ (bend is favored). If $\kappa_{\mathrm{F},1} = \kappa_{\mathrm{F},3}$ the stable configuration are spirals. However, in the degenerate case of the one-elastic-constant approximation ($\kappa_{\mathrm{F},1} = \kappa_{\mathrm{F},3}$) that we will consider in following chapters all defects have the same energy, independent of the value of ϵ. For other defects, with $s \neq 1$, varying ϵ simply results in a global rotation of the entire director field, thus for an isolated defect it is irrelevant. In particular, there are not different geometries as for the $+1$ defect. A defect with charge $s = -1$ is shown in Fig. 1.4d. Note that the director field rotates in the clockwise direction now, giving rise to the negative sign. So far, we have only considered defects of integer charge and sketched polar director fields. However, for nematic director fields it is possible to find half-integer charge defects as well. The configuration for a $+1/2$ defect is shown in Fig. 1.4e. Note that it is not possible to have these configurations in polar liquid crystals since the top-down asymmetry would lead to a discontinuity as shown in Fig. 1.4f. Thus ± 1 defects are the lowest possible defect charges for polar liquid crystals, but in nematic liquid crystals one can have $\pm 1/2$ defects. The configuration for the $-1/2$ defect is shown in Fig. 1.4g. Topologically, integer charges are allowed for nematic liquid crystals as well (as there is no discontinuity, see Fig. 1.4h), but this is usually energetically not favorable as we will discuss at the end of this subsection.

Poincaré-Hopf theorem and defect annihilation

With the definition of topological charge at hand, we can now comment on how defects can be created and destroyed. For this, first note that if several defects are encircled by the contour $\mathscr{C}$ one chooses when computing the integral in Eq. (1.13), the result is the sum of all the charges of the defects that are encircled. There is now a deep connection between the geometry of the system one considers and the sum of all defect charges present in the system. Namely, the Poincaré-Hopf theorem states that [12–17]

$$\sum_i s_i = \chi(\mathscr{D}) \,, \tag{1.14}$$

where $\chi(\mathscr{D})$ is the Euler characteristic of the system (manifold) $\mathscr{D}$. For this to hold the director field is required to point normal to the boundary $\partial\mathscr{D}$ of $\mathscr{D}$ (if there is a boundary). The Euler characteristic is a topological constant that is unique for a given geometry. For example, $\chi(\text{Disk}) = 1$, $\chi(\text{Sphere}) = 2$, $\chi(\text{Torus}) = 0$. Thus, for example, if we have a polar liquid crystal on a disk there will always be at least one $+1$ defect present. For a nematic liquid crystal on a sphere, there will always be at least four $+1/2$ defects present (or two $+1$ or one $+2$). Thus, the theorem Eq (1.14) gives a lower bound for the number of defects present in a system and restricts the number of possible defect configurations by relating the sum to the geometry of the system.

However, the Poincaré-Hopf theorem is a statement about the sum of all defect charges, not the total number of defects present. Thus, it is allowed to create two additional defects if they have equal and opposite charge such that the total charge of the system does not change. Equally, it is possible for two defects of charge s_1 and s_2 to merge, resulting in a defect with charge $s_1 + s_2$. Thus, if $s_1 = -s_2$ the two defects will annihilate and leave behind a region behind a defect-free region. For example, a $+1$ and -1 defect look from afar like a defect-free configuration, see Fig. 1.4i, and if they annihilate they leave a defect-free region behind. Similarly for a $+1/2$ and $-1/2$ defect, see Fig. 1.4j. However, a $+1$ and a $-1/2$ defect merging creates a $+1-1/2 = +1/2$ defect, see Fig. 1.4k. The only way to achieve an ordered state form a defective configuration is for defects to annihilate. And if the geometry of the system is chosen the right way, for example a circle, a defect-free configuration can never be reached, according to Eq. (1.14).

Defect charge from homotopy group

We now introduce a more rigorous definition of topological charge that is easily extended to higher dimensions and that highlights the topological character. We first consider two-dimensional defects, but comment on three-dimensional defects afterwards. The starting point is to introduce the order parameter space $\mathscr{M}$. For a two-component spin, e.g., a polar liquid crystals, this order parameter space is the unit circle, $\mathscr{M} = S^1$. That is because there is one free parameter, the angle θ, that can take values $\theta \in [0, 2\pi]$ and is 2π-periodic, i.e., the director $\boldsymbol{p}$ can point in any direction in the plane, and this is described by the unit circle. Given a certain director field, one then has a mapping $\mathcal{F}$ from coordinate space $\mathscr{D}$, e.g., the xy-plane, to the order parameter space $\mathscr{M}$. Basically, in the two-dimensional case, at any point in the plane there is a director which has a certain angle $\theta(x, y)$ such that $\boldsymbol{p}(x, y) = \cos\theta(x, y)\ \hat{\boldsymbol{x}} + \sin\theta(x, y)\ \hat{\boldsymbol{y}}$. Since the angle depends on the coordinates (x, y) there is thus a mapping $(x, y) \to \theta(x, y)$ for every point in the plane, see Fig. 1.5a. Using this simple mapping, it turns out that one can study the topological properties of defective configurations. To see this, we can consider a few examples for different director field configurations. Again, we look at closed loops $\mathscr{C}$ in the coordinate space, however, we are now interested in how this loop $\mathscr{C}$ in coordinate space is mapped into the order parameter space, i.e., we investigate how the trajectory $\mathscr{L} = \mathcal{F}(\mathscr{C})$ of the angle θ in $\mathscr{M}$. For a two-dimensional director field, $\mathscr{C}$ thus lives in $\mathbb{R}^2$ while $\mathscr{L}$ lives in S^1. For an almost perfectly aligned configuration we find that $\mathscr{L}$ is a small closed loop around a single value for θ.

For a configuration that is less well aligned, but sill without defects, $\mathscr{L}$ is now a loop that covers a bigger range of values in θ. However, it does not cover the entire circle, see Figs. 1.3a,b and Fig. 1.5b. Now, if we consider a +1 defect, this changes. If $\mathscr{C}$ contains a defect, we find that $\mathscr{L}$ is closed, and does run once around the circle, i.e., does not "turn around" as before for the defect-free configurations, see Fig. 1.3c and Fig. 1.5c Furthermore, note that it runs in an anti-clockwise configuration. If a -1 defect is present, $\mathscr{L}$ again runs once around the circle, now, however, in a clockwise-direction. For a $+2$ (-2) defect we find that $\mathscr{L}$ runs around the circle twice in a counterclockwise (clockwise) direction, see Fig. 1.3d. Thus, $\mathscr{C}$ encircles a defect of charge $|s|$, if $\mathscr{L}$ wraps around the circle s-times. To see the connection with topology we have to consider the so-called homotopy group. While it is possible to continuously contract $\mathscr{L}$ to a single point if it only covers part of a circle (defect-free configuration) it is not possible if it wraps around the circle. Thus, these two configurations are fundamentally different and it is not possible to move continuously between one and the other. Similarly for higher defect charges. Note that the possible defect charges (s is an integer) are thus completely determined by the topology of the order parameter space $\mathscr{M}$. $\mathscr{L}$ wraps around the circle s-times, or it is homotopic to a single point[1]. The question how many different loops there are that are unique in that they cannot be continuously transformed into each other is answered by the (first) homotopy group π_1 tells us how many different[2] loops can be drawn on a given space. For the circle S^1, it is possible to show that $\pi_1(S^1) = \mathbb{Z}$, i.e., the possible defect charges are all integer. This is an example of a more general framework, where the defect charges that are possible in a given system are completely determined by the topology (in particular the homotopy group) of the corresponding order parameter space. Thus, to answer the question which defect charges are possible in nematic liquid crystals, we just have to identify the order parameter space $\mathscr{M}$ and then compute (look up) the homotopy group of this space. Unlike polar liquid crystals, where the angle θ is 2π-periodic, for nematic it is π-periodic, due to the top-down symmetry of the nematic particles. Thus, the order parameter space is a half-circle, where θ runs from 0 to π, and where 0 and π are identified, just as 0 and 2π are the same point on the circle. This space is known as the real projective line $\mathbb{R}\mathrm{P}^1$. Thus, the loop $\mathscr{L}$ only needs to wrap around half a circle and then cannot be continuously shrunk to a point. In agreement with this, one finds that $\pi_1(\mathbb{R}\mathrm{P}^1) = \mathbb{Z}/2$, i.e., the allowed defect charges are now half-integer, see Fig. 1.5d. References of the underlying mathematical concepts are, for example, Refs. [12, 13, 15–18], while a greater focus on the applications to liquid crystals can be found, e.g., in Refs. [1, 2, 9, 11].

There are higher-dimensional analogs of the first homotopy group. Instead of 1-spheres (that is circles) one considers how many unique 2-spheres, 3-spheres, ... there

[1] One says that a loop is homotopic to a single point if it can be continuously shrunk to a single point. For example, if a surface has a hole and one draws a loop around this hole, it is not possible to continuously shrink the loop to a point. Therefore, a loop that wraps once around a circle cannot be shrunk to a point because of the "hole" of the circle. A loop on a disk, on the other hand, can be continuously shrunk to a single point.

[2] This is to be understood in the homotopy sense: two loops are considered to be the same if they are homotopic, i.e., can be continuously deformed into each other.

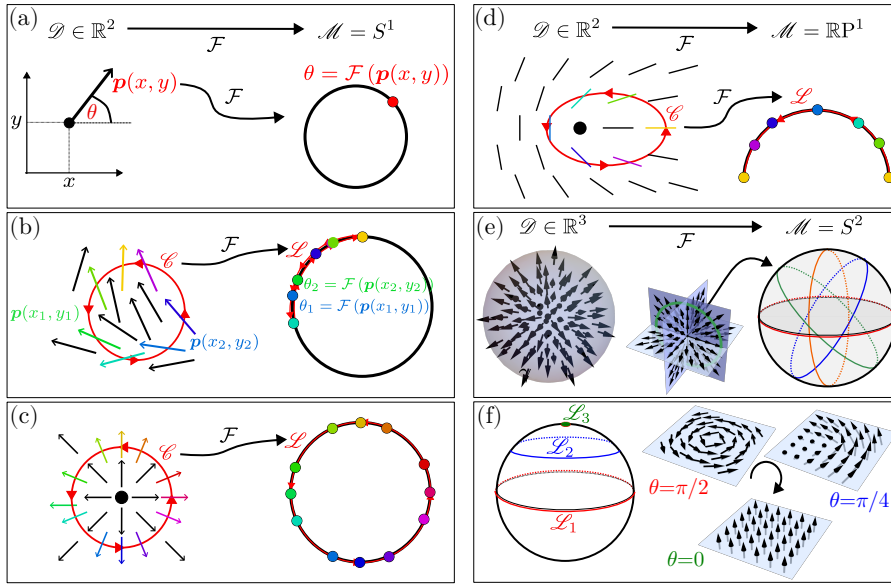

Figure 1.5: Homotopy groups and defect charge.

(a) Illustration of the mapping $\mathcal{F}$ explained in the main text of a two-dimensional polar director field at a single point (x, y) in real space $\mathbb{R}^2$ onto it's parameter space $\mathscr{M} = S^1$. (b) Mapping of the director of a defect-free configuration at several points onto the parameter space, as well as the mapping of a closed loop $\mathscr{C}$ from real space onto the corresponding loop $\mathscr{L}$ in parameter space. (c) The same mapping but for a structure containing a +1 defect. The loop $\mathscr{C}$ encircling the defect in real space is mapped onto the loop $\mathscr{L}$ that wraps once completely around the parameter space S^1. (d) The equivalent mapping but for a nematic liquid crystal, from real space into the parameter space $\mathscr{M} = \mathbb{RP}^1$. This space is a circle with antipodal points identified, thus it can be drawn as a half-circle whose two end-points are identical. In the +1/2-configuration illustrated here the yellow points are to be seen as the same point. (e) Mapping for a three-dimensional polar liquid crystal, from real space $\mathbb{R}^3$ to the parameter space $\mathscr{M} = S^2$. On the left-hand side a hedgehog configuration is shown. One can draw a circle in each cross-section of the hedgehog (as illustrated in the center) and map these onto the sphere. Doing this for each cross-section, the union of all the circles adds up to a sphere, such that one has a mapping of a sphere in real space onto the sphere that constitutes the parameter space. The hedgehog is assigned a charge +1 because the sphere that is mapped wraps around parameter-sphere exactly once. (f) If a two-dimensional polar liquid is allowed to rotate in the third dimension, point-defects are no longer topologically protected. Rotating all the vectors that form a +1 defect, it is possible to continuously transform the defective structure into a defect-free structure where all the arrows are parallel and point upwards. In parameter space, this rotation corresponds to contracting a loop on a sphere to a single point. Initially (red, $\mathscr{L}_1$) the structure is defective, the loop wraps around the equator of the sphere (at $\theta = \pi/2$). In the two-dimensional case this would be the loop wrapping once completely around the circle. However, if the director is allowed to rotate in the third dimension, the configuration space is a sphere, not a circle. Thus, it is possible to move the loop upwards, corresponding to a rotation of the director field by a certain angle (blue, $\mathscr{L}_2$). Finally, if the loop is contracted to a point (green, $\mathscr{L}_3$) the director field in real space is no longer defective but perfectly aligned.

are and this leads to the second, third, ... homotopy group. While more abstract, the basic idea is the same. Given a certain space $\mathscr{M}$, how many unique ways are there to embed n-spheres that cannot be continuously deformed into each other. In three-dimensional liquid crystals the higher-dimensional homotopy correspond to different kinds of defects. That is, unlike in two dimensions, where only point defects exists, there are now different kinds of defects which are measured using different homotopy groups. The first homotopy group measures line defects, which are "stretched out" point defects, i.e., each cross-section of the line looks like a point defect. A loop can be drawn around a line, and thus the first homotopy group corresponds to the possible defect charges of these disclination lines. The two-sphere S^2, on the other hand, is used to measure point defects. That is because if there is a single point, to completely encircle it one needs to draw a sphere around it. This is similar to how an electric charge is measured through Gauss's law by integrating the electric field on a sphere containing the charge. Thus, to learn which possible line and point defects there are in polar and nematic liquid crystals, we have to compute π_1 and π_2 for the respective parameter space. For a three-dimensional polar liquid crystal (three-component spin) the unit director can be parametrized in terms of two angles such that there are now two parameters $\varphi \in [0, 2\pi]$ and $\theta \in [0, \pi]$. These parameters span the two-sphere S^2 and thus $\mathscr{M} = S^2$ for three-dimensional polar liquid crystal. This simply reflects that $\boldsymbol{p}$ is a three-component spin of unit magnitude that can point in any direction in three-dimensional space. Then $\pi_2(S^2) = \mathbb{Z}$ is the analogy of the loop wrapping around the circle, with spheres wrapping around a sphere. This means that there are an infinite number of integer charge point defects. The unit charge point defect, called hedgehog, is shown in Fig. 1.5e. Each of its cross-sections looks like a two-dimensional +1 defect. Thus, to measure the charge one can take a contour integral in each cross-sectional plane which is mapped onto a loop on the parameter space, the 2-sphere. The union of all these lines results in a sphere wrapping round $\mathscr{M}$. Thus, to measure three-dimensional point charges one has to consider how 2-spheres can wrap around 2-spheres, that is $\pi_2(S^2)$. Again, either the sphere $\mathscr{L}$ can be continuously contracted to a single point, or it completely wraps around the sphere $\mathscr{M} = S^2$ a total of s-times, resulting in the allowed charges to be the integers $\mathbb{Z}$ again. On the other hand, $\pi_1(S^2) = 0$, i.e. all loops one draws on a sphere can be continuously deformed into each other, and in particular they can all be contracted to a single point, see Fig. 1.5f. Thus, there are no line defects in polar liquid crystals and only two distinct line defects in nematic liquid crystals. This is known as the "escape in the third dimension" and can be illustrated as follows. If one starts, for example from a two-dimensional +1 defect but the polar director lives in three dimensional, i.e., is allowed to rotate in z-direction, it is always possible to continuously rotate the director into an aligned configuration, see Fig. 1.5f. A "defect" line is just the same configuration but parallelly extended in z-direction. But since one can perform the same rotation operation in each cross-section of the line it is thus not topologically protected. For a three-component nematic director field, on the other hand, $\mathscr{M} = \mathbb{R}\mathrm{P}^2$. This space is called real projective plane and can be defined as a two-sphere with antipodal points identified. This is the two-dimensional analogue of the antipodal points of a circle being identified for a

two-dimensional nematic, resulting in the real projective line $\mathbb{R}\mathrm{P}$ (Fig. 1.5d). Again, to learn which kind of line and point defects are allowed one computes the first two homotopy groups, the result is $\pi_1(\mathbb{R}P^2) = \mathbb{Z}_2 = \{0, 1\}$, $\pi_2(\mathbb{R}\mathrm{P}^2) = \mathbb{Z}/2$. Thus, for a nematic liquid crystal, there is a single unique defect line. That is because it is possible to continuously deform a $+1/2$ into a $-1/2$ defect if the nematic director is allowed to rotate in z-direction. But it is not possible to reach the defect-free configuration as in the polar case, thus there is no escape in the third dimension. There is thus either the defect-free configuration or the disclination line of charge $s = 1/2$. There are no topologically protected lines of higher charges, since, e.g, a configuration that would have integer charge can be deformed into the defect-free configuration. This can easily be seen from $1 = 1/2 + 1/2 = 1/2 - 1/2 = 0$. In the second-to-last step we used that $+1/2$ can be deformed into $-1/2$, and thus the two are identical. Physically, this correspond to start from an integer line, split it into two half-integer lines, deform one of them such that each cross-section has the structure of a $-1/2$ point defect, and then combine the two lines again, resulting in an annihilation. In three-dimensional liquid crystals there are more exotic defects and one can consider higher-order homotopy groups. We refer to Refs. [12, 13, 15–18] for the mathematical background and to Refs. [1, 2, 9, 11, 19–25] for more details on the subject of higher-dimensional homotopy groups applied to liquid crystals.

Energy of defects

After this small detour we now return to the two-dimensional case and more basic notions, introducing some concepts of practical relevance. Here and in the following, we focus on $\pm 1/2$ and ± 1 defects because these are the minimal allowed defect charges in nematic and polar liquid crystals, respectively. Assuming an isolated point defect at the center of a disk of radius R, its elastic energy is found from the Frank free energy Eq. (1.11) with the director being given by $\boldsymbol{p} = \cos\theta\, \hat{\boldsymbol{x}} + \sin\theta\, \hat{\boldsymbol{y}}$ to be [1, 3]

$$F_\mathrm{F} = \frac{\kappa_\mathrm{F}}{2} \int \mathrm{d}A\, |\nabla \boldsymbol{p}|^2 = \frac{\kappa_\mathrm{F}}{2} \int \mathrm{d}A\, \frac{s^2}{r^2} = \pi \kappa_\mathrm{F} \int_a^R \mathrm{d}r\, r \frac{s^2}{r^2} \tag{1.15}$$

such that the energy of a defect can be written as [1–3]

$$E_s = \pi \kappa_\mathrm{F} s^2 \log \frac{R}{a} + E_\mathrm{c}\,, \tag{1.16}$$

where R is the size of the system and a the defect core radius, the small length scale cutoff where the continuous theory breaks down near the center of the defect. Physically, the order parameter goes to zero in this region, but in the simplified continuous theory we are using there is a singularity of the director field. See the following paragraph for a more detailed discussion. In the last step, going from Eq (1.15) to Eq. (1.16), we added a constant defect core energy E_c by hand to account for the energy associated with this region. In our theory it is an undetermined constant, but it is possible to model the defect core in order to get an expression for this energy, see Refs. [1–3] Thus, the elastic energy scales with the defect charge squared. In particular, the energy of two half-integer defects is smaller, by a factor

of two, than the energy of a single integer defect. If possible (that is for nematic liquid crystal), the elastic energy can hence be reduced by an integer defect splitting into two half-integer defects. Therefore, unless externally imposed through, e.g., substrate patterning or boundary conditions, one commonly only finds $\pm 1/2$ defects in nematic liquid crystals and ± 1 defects in polar liquid crystals. For this reason, $\pm 1/2$ and ± 1 are the most relevant defect charges for nematic and polar liquid crystals, respectively. Defects of greater charges are rarely encountered in practice.

Lastly, we briefly comment on a general feature of the defect geometry. We focus on two dimensions, but the same idea holds in three dimensions. One can split a defect into two regions, are core region and a far-field region. Single defects are system-wide structures in that the geometry of their far-field does not change with distances, if not for the presence of boundaries or other defects. In particular, the line integral, Eq. (1.13), yields the same charge s independent of how far away from the defect core one computes the line integral. In the far-field region the nematic order parameter is unity, $S = 1$, but the director field varies slowly in space in such a way that the Frank free energy is not minimal giving rise to the elastic energy of a defect written Eq. (1.16). The fact that the defect is a system-wide structure is reflected in the dependence of the elastic energy on the system size R. Near the defect core the structure of the defect changes, however. As can be appreciated, e.g., from the illustrations in Figs. 1.4a-h, the director field is not well-defined at the center of the defect. This results in a singularity in the director field at the origin, where the orientation of the director field is undefined. The singularity is an artifact of our approach of keeping the scalar order parameter at unity throughout the system. If the order parameter is allowed to vary, one finds that it decays to zero inside the defect core. Thus, directional order is locally destroyed, which removes the singularity. See, e.g., Refs. [1–3] for details and the resulting elastic energy associated with a defect core. However, if one is interested in the dynamics of defects away from the defect core it is sufficient to ignore the dynamics inside the defect core by introducing a short-distance cutoff a which reflects the fact that the assumption of unit order parameter is valid only on larger length scales, $r > a$. This cutoff was introduced when the free energy of a defect in Eq. (1.16) was calculated. It is easily seen that taking $a \to 0$ results in a divergent elastic energy, reflecting the breakdown of the model. We will proceed with using an ultraviolet cutoff in the following and ignore the dynamics inside the defect core.

1.1.4 Passive nematodynamics

So far we have only considered the elastic forces due to the Frank free energy (Sec. 1.1.2). In the presence of a flow field there are additional forces, different from the elastic relaxation driven by the Frank free energy, acting on the liquid crystal particles. The description of these problems goes under the name of nematohydrodynamics. The resulting dynamics is described by the following three

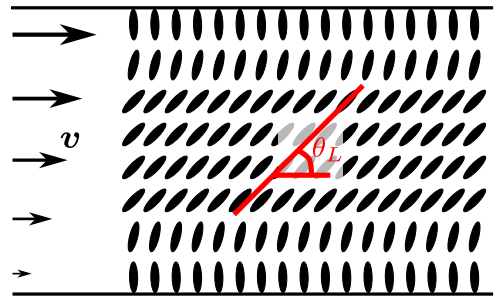

Figure 1.6: Flow alignment.

A liquid crystal is confined to a channel. At the edge of the channel the orientation of the liquid crystal is fixed. Here, the director field is set to be perpendicular at the boundary. If a shear flow is imposed (black arrow on the left), this flow interacts with the director field and rotates it. This results in a distorted director field, which does not minimize the Frank free energy, but instead is tilted in the direction of the flow by an angle $\pi - \theta_L$. The angle θ_L is called Leslie angle.

coupled equations [1–3, 26–28]:

$$D_t p_i + \omega_{ik} p^k = \lambda u_{ik} p^k + \Gamma^{-1} h_i \,, \tag{1.17a}$$

$$D_t v_i = \nabla^j (-P_\mathrm{h} g_{ij} + \sigma_{ij}) \,, \tag{1.17b}$$

$$\nabla \cdot \boldsymbol{v} = 0 \,. \tag{1.17c}$$

Here, $D_t \equiv D/Dt := \partial_t + \boldsymbol{v} \cdot \nabla$ is the material derivative while $u_{ij} = (\nabla_i v_j + \nabla_j v_i)/2$ and $\omega_{ij} = (\nabla_i v_j - \nabla_j v_i)/2$ are the strain-rate and vorticity tensor, respectively. The molecular field $\boldsymbol{M}$, defined in Eq. (1.12) in the one-elastic-constant approximation, governs the relaxation dynamics of the nematic director. The material parameters λ and Γ are, respectively, the flow-alignment parameter and the rotational viscosity of the nematic fluid. The rotational viscosity determines how quickly the liquid crystal relaxes back to its equilibrium configuration in the absence of flows, $\boldsymbol{v} = 0$. On the other hand, the flow-alignment parameter controls how the nematic director field responds to a shear flow. In general, a shear flow can be decomposed into a rotational and an extensional component. For $|\lambda| > 1$ the material is called flow aligning and the director field will tend to align with an angle $\theta_\mathrm{L} = \arccos(1/\lambda)/2$ to the direction of flow in which case the rotational and extensional parts of the shear flow are balanced, see Fig. 1.6. On the other hand, for $|\lambda| < 1$ it is called flow tumbling with the rotational part of the shear flow always dominating such that the director will rotate continuously under shear. Furthermore, the sign of λ is related to the shape of the liquid crystal particles. Namely, for $\lambda < 0$ particles are imagined to be rather disk-like whereas for $\lambda > 0$ they are more rod-like. $\lambda = 0$ corresponds to spheres.

Eq. (1.17a) can thus be understood as the evolution equation for the director field in the presence of a flow. The terms on the left-hand site constitute the comoving and corotational derivative. On the right-hand side there are two, potentially competing,

terms. The first one describes how the director field is deformed by the presence of a flow and the second one drives the liquid crystal back to its equilibrium configuration such that the molecular field vanishes if the Frank free energy is minimal. If no flow is present this equation reduces to a gradient descent describing how the director evolves to its equilibrium state over time.

Eq. (1.17b) additionally contains a hydrodynamic pressure P_h, the metric g_{ij} and the dissipative stress tensor $\sigma_{ij} = 2\eta u_{ij} + \sigma_{ij}^\mathrm{el}$. This equation is the momentum balance in the overdamped regime and reduces to the usual low-Reynolds-number Navier-Stokes equation if the elastic stress tensor is taken to vanish, i.e., if $\sigma_{ij}^\mathrm{el} = 0$. This elastic stress tensor is due to the presence of the liquid crystal which results in an additional stress. We will give an explicit expression for this stress tensor below. Through this equation the director field $\boldsymbol{p}$ can act back on the flow field and, in particular, it can source a flow if $\boldsymbol{M} \neq 0$. This phenomenon is termed backflow. Finally, we only consider incompressible fluids, which is enforced by Eq. (1.17c), and eventually results in a non-vanishing hydrodynamic pressure. If a fluid is incompressible its density is constant.

Eq. (1.17a) is called Leslie-Ericksen equation and describes the dynamics of the director field in the presence of a flow field that potentially distorts the liquid crystal field and drives it away from its equilibrium configuration. These equations were derived originally by Refs. [29–33] from symmetry and thermodynamic considerations in the form they are presented here. That is, all terms allowed by symmetry were written down and then relations between them were derived from thermodynamic principles. The derivation assumes local equilibrium and that the system is close to equilibrium. Furthermore, an uniaxial nematic state with uniform magnitude of order is assumed. These assumptions are valid far below the nematic-isotropic transition temperature. A more general set of equations, where only local thermodynamic equilibrium and the system being close to equilibrium, is assumed can be derived from linear irreversible thermodynamics. In this case the equations are written in terms of the $\boldsymbol{Q}$-tensor order parameter Q_{ij} such that uniaxiality and a constant scalar order parameter are no longer required. The derivation is similar to the one for the Leslie-Ericksen equation we sketch below except that the Frank free energy and the dynamic equation for the flow parameter are now written in terms of Q_{ij}. If written in term of $\boldsymbol{Q}$, Eq. (1.17a) is typically referred to as the Beris-Edwards equation. See Refs. [34–38] for details.

Derivation of nematodynamic equations

We will not consider the $\boldsymbol{Q}$-tensor approach further now, but instead very briefly sketch the derivation of Eqs. (1.17). We refer to Refs. [3, 39, 40] for more details of the derivation below. The total free energy has the form

$$F_\mathrm{tot} = \int \mathrm{d}V \; \frac{\rho}{2} v^2 + F_\mathrm{int} + F_\mathrm{F} \; , \tag{1.18}$$

where the first term represents the kinetic energy with ρ the density, and F_int is an internal free energy, e.g., enforcing incompressibility. Then, for an isothermal process, the dissipation (energy production) rate $T\dot{S}$ is given by $T\dot{S} = -D_t F_\mathrm{tot}$. Using

the general momentum conservation equation $\rho D_t v^i = \nabla_i \sigma^{ij}$ one finds (dropping surface terms)

$$-\frac{D}{Dt}\int \mathrm{d}V\ \frac{\rho}{2}v^2 = \int \mathrm{d}V\ \sigma_{ij}\nabla^i v^j\ , \tag{1.19a}$$

as well as

$$-\frac{D}{Dt}\int \mathrm{d}V\ F_{\mathrm{int}} + F_{\mathrm{F}} = \int \mathrm{d}V\ M^i D_t p_i - \sigma^{\mathrm{E}}_{ij}\nabla^i v^j\ . \tag{1.19b}$$

Here,

$$\sigma^{\mathrm{E}}_{ij} := -\frac{\partial F_{\mathrm{F}}}{\partial(\nabla^i p_k)}\nabla_j p_k - P_{\mathrm{h}} g_{ij} \tag{1.19c}$$

is the so-called Ericksen stress which is the sum of the distortion stress tensor and a pressure term which is due to incompressibility. The first term on the right-hand side of Eq. (1.19b) is due to a variation in the orientation of the liquid crystal particles while keeping their center of mass fixed. The second term is due to the displacement of the center of mass of the liquid crystal particles while keeping their orientation fixed. Thus, adding the two contributions Eq (1.19a) and Eq. (1.19b), the total dissipation is given by

$$T\dot{S} = \int \mathrm{d}V\ \sigma^{\mathrm{v}}_{ij}\nabla^i v^j + M^i D_t p_i\ , \tag{1.20}$$

where we defined the viscous stress σ^{v}_{ij} as the difference between the total stress and the equilibrium stress, $\sigma^{\mathrm{v}}_{ij} = \sigma_{ij} - \sigma^{\mathrm{E}}_{ij}$. In general this stress is asymmetric and the antisymmetry is characterized by a vector $L_k = -\epsilon^{ij}{}_k \sigma^{\mathrm{v}}_{ij}$. It is found that $\boldsymbol{L} = \boldsymbol{p}\times\boldsymbol{M}$, i.e., it is the torque exerted by the fluid on the director, and this torque vanishes in equilibrium where director and molecular field are parallel. Denoting by σ^{s}_{ij} the symmetric part of σ^{v}_{ij} it is then easily seen that

$$T\dot{S} = \int \mathrm{d}V\ \sigma^{\mathrm{s}}_{ij}u^{ij} + p^i\ , N_i\ , \tag{1.21}$$

where $N_i := D_t p_i + \omega_{ij}p^j$ describes the rate of change of the director field with respect to the background fluid. Thus, the two sources of dissipation are seen to be shear flow and rotation. Assuming that the fluxes u_{ij} and N_i are weak on the molecular scale it is possible to write the forces σ^{s}_{ij} and M_i in terms of these fluxes:

$$M_k = \mathscr{R}_{ijk}u^{ij} + \mathscr{U}_{kl}N^l\ , \tag{1.22a}$$

$$\sigma^{\mathrm{s}}_{ij} = \mathscr{P}_{ijkl}u^{kl} + \mathscr{R}_{ijk}N^k\ , \tag{1.22b}$$

where Onsager relations where used to simplify these expressions. The yet to be determined tensors $\mathscr{P}_{ijkl}$, $\mathscr{R}_{ijk}$, and $\mathscr{U}_{ij}$ have dimensions of viscosity. To determine these tensors it is used that they have to be symmetric under the local symmetry of the nematic (symmetry group $D_{\infty\mathrm{h}}$), such that the only vector appearing in their expressions is the local director. Additionally, the final equations must be invariant

under $\boldsymbol{p} \to -\boldsymbol{p}$. The most general expressions, assuming incompressibility, are then found to be

$$M_i = \zeta_1 N_i + \zeta_2 p^j u_{ij} \tag{1.23a}$$

$$\sigma^{\mathrm{v}}_{ij} = \nu_1 p_i p_j p^k p^l u_{kl} + \nu_2 p_i N_j + \nu_3 p_j N_i + \nu_4 u_{ij} + \nu_5 p_i p^k u_{kj} + \nu_6 p_j p^k u_{ki} \tag{1.23b}$$

with $\zeta_1 = \nu_3 - \nu_2$ and $\zeta_2 = \nu_2 + \nu_3 = \nu_6 - \nu_5$. The coefficients ν_i are called Leslie coefficients and there are thus five independent coefficients with the dimension of viscosity. Note that, at lowest order in $\boldsymbol{p}$ the total stress tensor just reduces to $\sigma_{ij} = -P_{\mathrm{h}} g_{ij} + \nu_4 u_{ij} + \mathcal{O}(\boldsymbol{p})$ and that is the (passive) stress tensor we will use in the following chapters. Furthermore, in the absence of any flow field, i.e. $u_{ij} = 0$, the viscous stress can be written as

$$\sigma^{\mathrm{v}}_{ij}\big|_{u_{ij}=0} = \frac{\zeta_2}{2\zeta_1}\left[p_i M_j + p_j M_i\right] - \frac{1}{2}\left[p_i M_j - p_j M_i\right] \tag{1.24}$$

which is then the stress due to the director field not minimizing the Frank free energy. To make connection with Eqs. (1.17) we define $\Gamma = 1/\zeta_1$ and $\lambda = -\zeta_2/\zeta_1$. Thus, we immediately recover Eq. (1.17a) from Eq. (1.23a). Furthermore, we define $\nu_4 = 2\eta$ and thus find the expression for the stress tensor in Eq. (1.17b) that was stated above, where we now also have an explicit expression for the stress tensor $\sigma^{\mathrm{el}}_{ij}$ as was promised before.

1.2 Active Matter

We will now turn towards active systems. So far we have introduced liquid crystals that are passive, that is the individual particles respond only to external forces and are not motile themselves. However, the thesis is concerned with active systems. These are systems where the individual constituents making up the system inject energy into the system themselves. For example, the rods making up a nematic liquid crystal are now motile and move around on their own. This drives the system as a whole away from equilibrium and allows numerous new dynamic structures [41–53]. The field of active matter is nowadays very large and we focus on three of the main models, namely active liquid crystals, the Toner-Tu equations, and the Vicsek model. All three are models used to explain macroscopic pattern formation in systems with underlying orientational order and where microscopic particles are motile. The former two are both coarse-grained, hydrodynamic theories used to describe the large-scale behavior of systems. Thus, these models are employed when one is interested in the dynamics on the level of the entire system and not the dynamics of its microscopic constituents. The hydrodynamic theories are typically classified according to the symmetry of the microscopic particles (are they nematic or polar) as well as the relevance of momentum conservation. If momentum is not conserved, the systems (or rather models used to describe a system) are called “dry”. On the other hand, if momentum conservation is relevant the system is referred to as “wet” [42]. This convention is due to the fact that systems in which momentum conservation is important are often systems in which active particles are suspended in a surrounding fluid such that solvent-mediated hydrodynamic interactions are

relevant and must be taken into account, with the total momentum in these systems thus being conserved. In a dry system, on the other hand, the dynamic of the surrounding fluid is irrelevant for the dynamics of the particles, e.g., because of a separation of timescales, and its presence can be incorporated by a simple friction term. For example, the active particles can be modeled as a system with overdamped dynamics where momentum is simple transferred to the surrounding fluid. In this case momentum is not conserved in the system of active particles one considers, though of course globally it is still conserved. A successful example of model used to describe wet systems is that of active nematodynamics whereas dry active matter is often described by the Toner-Tu model. In the following, we first introduce active liquid crystals and after that we turn towards introducing the Toner-Tu equations. These are the active hydrodynamic models used in later chapters. The other class of systems, agent-based models, are instead built up by considering not macroscopic, hydrodynamic quantities but the dynamics of individual microscopic particles. A famous example of these is the Vicsek model which we will briefly discuss thereafter as well. At the end of the section, we introduce some applications of active matter theory to biological systems.

1.2.1 Active nematodynamics

The equations used in the framework of active liquid crystal theory are very similar to the ones derived/described in the previous subsection. The only addition is that there is an additional term in the stress tensor, the so-called active stress. This stress is a macroscopic quantity found by coarse-graining over the energy-consuming microscopic particles and drives the system out of equilibrium on the macroscopic scale. We present a version of this microscopic derivation in Chapter 3. Here we just state the commonly used expression for the active stress tensor and justify it from symmetry considerations. In the following subsection we present some examples to highlight different features of active liquid crystals.

Active hydrodynamic equations

The active stress tensor we (mostly) use reads $\sigma^{\mathrm{a}}_{ij} = \alpha Q_{ij}$. Here, α is the proportionality constant between the nematic order parameter and the active stress. It is called activity. This stress is in general anisotropic, thus cannot be absorbed into a pressure, and the symmetry is broken by the presence of the director field. In this sense, this expression for the stress tensor is the obvious choice in that it is the lowest-order symmetric tensor in the director field that respects the nematic symmetry and it is not isotropic (because of the aligned liquid crystal particles breaking rotational symmetry). Thus, as we will show in Chapter 3, the microscopic particles are modeled as force dipoles with a the direction of force given by their direction, see Fig. 1.7a. Depending on its sign, one calls the activity either extensile, if $\alpha < 0$ or contractile, if $\alpha > 0$. For extensile activity the force is pushing outwards, in the direction of the orientation, while for contractile activity it is pulling inwards. See Fig. 1.7b for a sketch of the active force dipole. This model of a microscopic force dipole was already used by, e.g., Pedely and Kessler [54] and an active version of Eqs. (1.17) was first written down by Simha and Ramaswamy [55]. They derived

the active stress by coarse graining over a model of microscopic force dipoles; see also e.g. Refs. [41, 42, 48, 56] for more details. Shortly after, the same equations were derived in the framework of linear irreversible thermodynamics. This approach is very similar to the derivation of the passive case sketched in the previous subsection [57–59]. From almost the very beginning, a large effort has been made to connect these hydrodynamic equations to microscopic models and the statistical mechanics of non-equilibrium systems, e.g., in Refs. [55, 60–63]. We do not consider this further, see, e.g., Refs. [42, 64, 65] and references therein for an introduction and more details.

We will use this active stress tensor and drop the elastic stress tensor because, while it does affect the details of the dynamics, it is often unimportant for the phenomenology of active nematics. Since it complicates the structure of the equations considerably, we treat it as a subleading correction and do not consider it in the following chapters. Thus, to summarize and for future reference, our dynamic equations of an active liquid crystal (for both nematic and polar liquid crystals) are

$$D_t p_i + \omega_{ik} p^k = \lambda u_{ik} p^k + \Gamma^{-1} h_i \,, \tag{1.25a}$$

$$D_t v_i = \nabla^j (-P_\mathrm{h} g_{ij} + \sigma_{ij}) = \nabla^j (-P_\mathrm{h} g_{ij} + 2\eta u_{ij} + \sigma^\mathrm{a}_{ij}) \,, \tag{1.25b}$$

$$\nabla \cdot \boldsymbol{v} = 0 \,. \tag{1.25c}$$

Note that the pressure P_h now includes an active pressure P_a. See Fig. 1.7b for the flow that a single microscopic particle can drive due to the coupling between active force and velocity field. Even though Eqs. (1.25) look very similar to the passive equations, Eqs. (1.17), the resulting dynamics are much richer and more complicated. For example, for flow aligning systems ($\lambda > 1$), extensile active liquid crystals are unstable to splay instabilities while contractile active systems enhance bend instabilities; and vice versa for flow tumbling systems ($\lambda < 1$) systems. We now present two examples to illustrate some aspects and for future reference.

Spontaneous flow transition

First predicted by Voituriez et al. [26], the spontaneous flow transition in active liquid crystals is a transition between a stationary and a flowing state in a channel. This transition is analogous to the Fréedericksz transition in standard liquid crystal theory [3, 40] with the twist that instead of an externally applied magnetic field the transition is now driven by the active forces produced on the microscopic level[3]. The ground state considered is an active, two-dimensional film in a channel that is infinitely long in x-direction (such that the whole problem is translationally symmetric in this direction) and has width L in y-direction[4]. The boundary condition imposes that the director field is parallel to the edge of the channel, i.e., pointing in x-direction there.

[3] In fact, in the case that there is no flow alignment ($\lambda = 0$) the problem can be mapped exactly onto the Fréedericksz transition.

[4] Note that the system is rotated by 90° with respect to Ref. [26] to make a comparison with Chapter 3 easier.

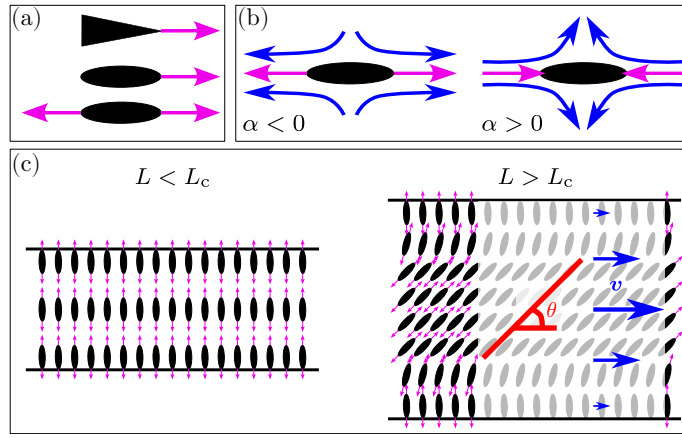

Figure 1.7: Activity and active flow transition.

(a) A polar particle (top) exerts a force (pink arrow) along its own direction. A nematic particle (center) can also exert a directed, polar force, or it can exert a non-directed, nematic force (bottom) (b) In our convention, if the forces are pointing outward, the activity α is negative and we call the activity extensile. If the forces are pointing inwards the activity is positive and we call it contractile. The blue arrows illustrate the fluid flow that the respective activities can drive. (c) Active particles are aligned in a channel (left) of width L smaller than some critical length scale L_c. If the width of the channel is increased to some value greater than this critical length scale, the particles away from the edge of the channel (where perpendicular anchoring is enforced) rotate by an angle $\pi - \theta$ and a flow (blue arrows) is created in this spontaneous flow transition.

They consider an active polar liquid crystal whose dynamic is essentially described by Eqs. (1.25) but including the elastic stress tensor as written in Eq. (1.24) and with $\kappa_{\mathrm{F},2} = 0$ (no twists since we are working in two dimensions) in the Frank free energy, Eq. (1.10), as well as an additional term $\propto p_i$ included in the Leslie-Ericksen equation that is related to the activity. However, the general phenomenon is independent of these small modifications, as will be seen in Chapter 3 as well. There we repeat the calculation of Voituriez et al. [26] in a slightly more general context (including a chiral active stress) and thus we present the detailed calculation in Chapter 3 and only state the results of Ref. [26] now. The onset of the transition is essentially controlled by the sign of $\alpha(1 + \lambda)$, which has to be positive for a spontaneous flow transition to occur. In this case the system acquires a non-vanishing velocity field and starts flowing, if, for a given activity, the width of the channel exceeds a critical size L_c. Alternatively, for a given channel thickness this leads to a critical activity. The critical length scale is proportional to the so-called active length scale defined as $\ell_\mathrm{a} = \sqrt{\kappa_\mathrm{F}/|\alpha|}$, quantifying essentially the relative strength of elastic versus active forces. Below the critical threshold, $L < L_\mathrm{c}$, the elastic terms dominate and the polarization field is constant with the orientation set by the boundary condition. Above the transition, $L > L_\mathrm{c}$, the stationary state $\boldsymbol{v} = 0$ is unstable and a new solution exists with a non-trivial flow field $\boldsymbol{v} \neq 0$. Due to the interaction with the flow-field, the constant polarization field is disrupted and rotates away from the edge of the channel. If the integration constant is chosen such that the total net flow vanishes, the minimum tilt angle reached by the polarization field for flow-aligning systems is essentially given by the Leslie angle θ_L. The structure of the velocity field and the exact point at which the flow transition occurs depend on the boundary conditions of either the velocity or the stress field. See Fig. 1.7c for an example with perpendicular anchoring of the nematic at the boundary and no-slip boundary conditions (i.e., the velocity field vanishes at the boundary). A more detailed/extended look at this phenomenon can be found e.g. in Refs. [27, 28] and we will return to this problem in detail in Chapter 3.

The take-away message is that the transition know for passive liquid crystals is recovered in active systems. But, whereas in the passive case an external field had to be applied to distort the uniformly aligned system, in the active case the transition is driven by the system itself. The presence of activity alone can create a non-vanishing flow field that distorts the director field.

Active defects

We will now consider the affect of activity on the topological defects in the director field introduced in the passive case in Sec. 1.1.3. Again, we merely state the results here and go into more detail for some of the calculations in the following chapters. In particular, we consider four different defect charges, namely $\pm 1/2$ and ± 1, in two dimensions because these are the minimal allowed defect charges in nematic and polar liquid crystals, respectively. In general, the idea is to choose the configuration of the director field $\boldsymbol{p}$ such that the defect one wants to study is present. That is one chooses $p^i = \cos\theta\, \hat{\boldsymbol{x}} + \sin\theta\, \hat{\boldsymbol{y}}$ with $\theta = s\varphi + \epsilon$, where φ the angular coordinate in polar coordinates, s the defect charge, and ϵ a constant, c.f. Sec. 1.1.3. For this configuration one then solves Eqs. (1.25) for the velocity field. The presence of

activity combined with the presence of symmetry-breaking due to the defect results in numerous interesting phenomena that have been studied in detail. In particular, there is mounting evidence that active defects have biological functions, see below in Sec. 1.2.4 for more on this.

Historically, the case of an active +1 defect in polar systems was first studied by Kruse et al. [57]. It is the simplest case to study because the +1 defect is rotationally symmetric which simplifies the calculations considerably. As explained in Sec. 1.1.3 the constant ϵ determines the local geometry of the defect such that the defect is an aster if $\epsilon = 0 \mod \pi$, a vortex if $\epsilon = \pi/2 \mod \pi$, and a spiral for all values of ϵ in between these two limiting cases. At equilibrium, which configuration is the one minimizing the elastic energy is determined by the elastic constants in the Frank free energy and boundary conditions. In an active system, which is away from equilibrium, the stability of the different configurations is determined from solving Eqs. (1.25). Since the problem is rotationally symmetric the velocity field always only has a azimuthal component v^φ and the radial velocity field v^r vanishes. It was found by Kruse et al. [57] that

$$v^\varphi(r) = \alpha\omega_0(\eta, \epsilon) r \ln \frac{r}{r_0} \,, \tag{1.26}$$

where r_0 is a constant set by the boundary conditions, and $\omega_0(\eta, \epsilon)$ is a function of the constant ϵ and viscosity η. Imposting no-slip boundary conditions at the boundary of a disk of radius R results in $r_0 = R$, see Fig. 1.8b. Including friction results in a new effective length scale that sets r_0. Thus, spiral defects are rotating for any non-vanishing value of activity. For $\kappa_{\mathrm{F},1} = \kappa_{\mathrm{F},3}$ it is found that a steady state solutions exists if $|\lambda| > 1$ and then $\cos 2\epsilon = 1/\lambda$. Thus, the spiral defects rotate for any non-vanishing velocity. Not assuming the one-elastic-constant approximation, i.e., for $\delta\kappa_\mathrm{F} = \kappa_{\mathrm{F},3} - \kappa_{\mathrm{F},1} \neq 0$, a stability diagram can be obtained [57]. For flow-aligning systems, if contractile activities are sufficiently small, asters (vortices) are stable if $\delta\kappa_\mathrm{F} > 0$ ($\delta\kappa_\mathrm{F} < 0$). Above some critical value of activity there is a dynamic instability and only rotating spirals are stable. An in-detail follow-up numerical investigation of the phase diagram was carried out by Elgeti et al. [66].

We now move on to the nematic case where $\pm 1/2$ defects are relevant. This was first systematically investigated theoretically by Giomi et al. [67]. These are no longer spherically symmetric and thus the expressions are more complicated. In particular, a steady-state analytical solution of the velocity field can be found only if one only solves the Stokes equation, Eq. (1.25b), but not the Leslie-Ericksen equation, Eq. (1.25a). There is thus no feedback of the flow on the orientation of the director. Since the constant ϵ does not change the geometry of the defects (unlike in the case of a +1 defect) but only results in a total rotation one can thus set $\epsilon = 0$, see Sec. 1.1.3. The flow fields generated by the $\pm 1/2$ defects are shown in Figs. 1.8c-f. The most interesting new phenomenon is that the active $+1/2$ defect, and only this one, is self-propelling. That is, a $-1/2$ defect has a three-fold rotational symmetry (symmetric under rotation of $2\pi/3$) in the director field and a six-fold rotational symmetry in the velocity field. This results in the velocity field vanishing at the defect core. The same was found in the case of a +1 defect, see Eq. (1.26). However, the $+1/2$ defect has an orientation and it is only

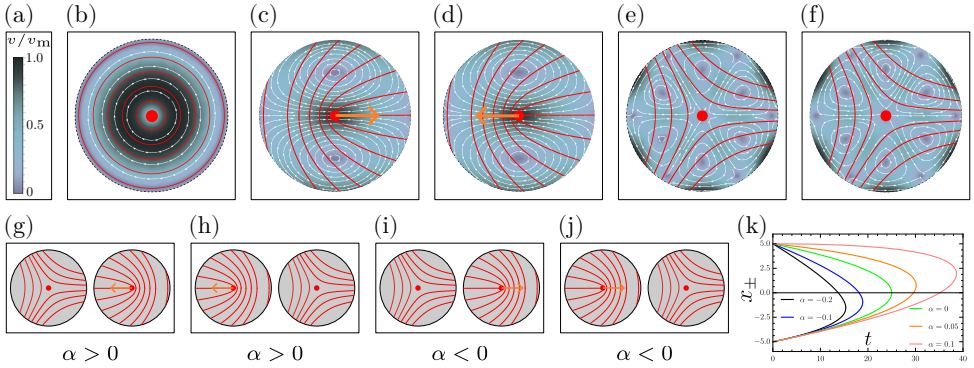

Figure 1.8: Active defects.

(a) Color bar of the speed relative to the maximal speed used in following panels. (b) Velocity field (white stream lines) of a +1 defect (red stream lines are director field) for contractile activity.. The flow is purely azimuthal and the defect is stationary. (c) Velocity field of a $+1/2$ defect for contractile activity. The flow created by the defect results in the defect moving in the direction of the orange arrow, i.e., towards its "tail". (d) If the activity is extensile, the velocity field is mirrored and the defect moves towards its head. (e) Velocity field of a $-1/2$ defect for contractile activity. Like the +1 defect it is stationary. (d) Velocity field $-1/2$ defect for contractile activity; the velocity changes sign compared with extensile activity. (g) If a $+1/2$ and a $-1/2$ defect are present in the same system, there are four fundamentally different configurations possible. First, in (g), we show the configuration for contractile activity and where the half-integer defects are oriented such that the velocity of the $+1/2$ defect points towards the $-1/2$ defect. (h) Contractile activity but the velocity of the $+1/2$ defect points away from the $-1/2$ defect. (i) Same configuration but for extensile activity. (j) Extensile activity but the velocity of the $+1/2$ defect points towards the $-1/2$ defect again. (k) Annihilation trajectories for different values of activity for defects originally oriented as in (h) and (j). The positive defect is initially located at $x_+ = 5$ while the negative defect is located at $x_- = -5$. Plotted is the position of the defects as a function of time for different activities. When the two branches intersect, the defects are at the same position and annihilate. The passive case is shown in green, the trajectories are symmetric and the defects annihilate at $x = 0$. For negative (extensile) activity, the velocity of the $+1/2$ defect points towards the $-1/2$ defect. This results in a speed-up of the annihilation, which now occurs at smaller times compared with the passive case, and an asymmetry. The annihilation occurs not at $x = 0$, but at $x < 0$. The greater the absolute value of activity, the faster the speed of the $+1/2$ defect and the annihilation process. For positive (contractile) activity the velocity of the $+1/2$ defect points away from the $-1/2$ defect. This slows the annihilation process down and the defects now annihilate at a position $x > 0$. Time and position are in arbitrary units.

mirror symmetric, with the symmetry axis being its orientation, not rotationally symmetric. This results in a velocity field that is non-vanishing at the defect core. Namely, for a defect in a disk of radius R and with no-slip boundary conditions, the velocity at the defect core is given by $v_0 = \alpha R/(4\eta)$. This is the velocity with which the defect core (and therefore the entire defect) is moving. Thus, due to the presence of activity the defect effectively creates a flow field that then in turn transports the defect, effectively resulting in a self-propelling particle. The direction of motion is set by the sign of activity. For contractile activity ($\alpha > 0$) the defect moves in direction of its "tail" while for extensile systems ($\alpha < 0$) it moves in direction of its "head'. That is the defect moves towards the left and right, respectively, for a director configuration as in Fig. 1.8c,d.

So far we have considered only a single isolated defect. Knowing the flow fields created by these it is then, e.g., possible to study the influence of activity on the defect annihilation mentioned in the passive case in Sec. 1.1.3. We consider the case of one $+1/2$ and one $-1/2$ defect at an initial distance $\Delta(t = 0)$. It turns out that both the initial configuration (that is, the relative orientation of the defects) and the sign of the activity are relevant. If the system is initialized such that the tail of the $+1/2$ defect points in the direction of the $-1/2$ defect, then the active flow results in a speed-up of the $+1/2$ defect towards the $-1/2$ defect in contractile systems such that effectively the annihilation time decreases. On the other hand, in extensile systems the $+1/2$ defect moves towards its head, that is away from the $-1/2$ defect, hence effectively resulting in a repulsive interaction. If the $+1/2$ defect is oriented initially in the opposite direction, then the behavior is reversed[5]. See Figs. 1.8g-j. To obtain an equation describing the defect dynamics, one proceeds as follows. The energy of a defect pair is given found similarly to the energy of an isolated defect (Eq. (1.16)), but with

$$\theta = s_1 \arctan\left(\frac{y - y_1}{x - x_1}\right) + s_2 \arctan\left(\frac{y - y_2}{x - x_2}\right) + \epsilon \,. \tag{1.27}$$

Choosing the coordinate system such that $y_1 = y_2 = 0$, and $x_2 = -x_1 = \Delta/2$, and integrating in polar coordinates over $\theta \in [0, 2\pi]$ and $r \in [0, \Delta/2 - a] \cup [\Delta/2 + a, R]$ one finds the elastic energy

$$E_{s_{1+2}} = E_{s_1} + E_{s_2} + 2\pi\kappa_{\mathrm{F}} s_1 s_2 \ln\frac{R}{\Delta} + 2E_{\mathrm{c}} \,, \tag{1.28}$$

i.e., the energy consists of the sum of the energy of the isolated defects plus an interaction term. This expression can be rewritten as [1]

$$E_{s_{1+2}} = \pi\kappa_{\mathrm{F}} (s_1 + s_2) \ln\frac{R}{a} + 2\pi s_1 s_2 \ln\frac{a}{\Delta} + 2E_{\mathrm{c}} \,. \tag{1.29}$$

Thus, the divergent term $\ln R$ vanishes if the defect charges add up to zero, $s_1 + s_2 = 0$. This reflects the fact that if in a region the defect charges add up to zero, from

[5] If the two defects are not as perfectly aligned as in the configurations sketched in Figs. 1.8g-j then there are additional torques acting on the defects, see [68, 69].

far away the director field looks homogenous and not defective, see Figs. 1.4i-k. The elastic force exerted by defect 1 on defect 2 can be found from this expression to be

$$\boldsymbol{F}_2 = -\nabla_2 E_{s_{1+2}} = (2\pi\kappa_{\mathrm{F}} s_1 s_2 \Delta)\,\hat{\boldsymbol{x}}, \tag{1.30}$$

thus the force is attractive if the defects have charge of opposite sign (resulting in the annihilation), and repulsive if they have the same sign. We now consider the case of two half-integer defects in more detail, with a $+1/2$ defect at position $\boldsymbol{r}_+ = \{x_+, 0\}$ and a $-1/2$ defect at position $\boldsymbol{r}_- = \{x_-, 0\}$. In general, assuming an overdamped equation of motion, and in the absence of backflow, the resulting equations of motion for a single defect is given by $\zeta \mathrm{d}\boldsymbol{r}/\mathrm{d}t = \boldsymbol{F}$, where $\boldsymbol{F}$ is the net force acting on the defect, and ζ an effective drag coefficient. In our case, this results in the equation of motion

$$\frac{\mathrm{d}x_\pm}{\mathrm{d}t} = \mp\frac{\mathcal{C}}{x_+ - x_-}\,, \tag{1.31}$$

where $\mathcal{C} = 2\pi\kappa_{\mathrm{F}} s_1 s_2/\zeta$, such that

$$\frac{\mathrm{d}\Delta}{\mathrm{d}t} = -\frac{2\mathcal{C}}{\Delta}\,. \tag{1.32}$$

It is straightforward to integrate this equation and one finds $\Delta(t) = 2\sqrt{\mathcal{C}}\sqrt{t_{\mathrm{a}} - t}$, where t_{a} is an integration constant. It is the time at which the defect separation vanishes and thus the time at which the defects annihilate. See Fig. 1.8k for an example of an annihilation trajectory. This dynamics was studied in detail in Refs. [70–74], among others. So far we have only considered the passive case, where the annihilation is purely due to the gradient force derived from the elastic energy. To investigate the effect activity has on the annihilation dynamics we include the self-propulsion speed v_0 of the $+1/2$ defect by simply adding v_0 on the right-hand side of Eq. (1.31) for the equation for x_+. Thus, the equation for the time evolution of the distance between the two defects takes the simple modified form [67, 75]

$$\frac{\mathrm{d}\Delta}{\mathrm{d}t} = v_0 - \frac{2\mathcal{C}}{\Delta}\,. \tag{1.33}$$

It is found that for sufficiently large activity and rotational viscosity there are two regimes. For large separations of the defects the dynamics is dominated by activity and $\Delta(t) \propto \alpha t$. If the defects are close to each other, they behave as in the passive case, $\Delta \propto \sqrt{t_{\mathrm{a}} - t}$. Some resulting trajectories for different values of activity are shown in Fig. 1.8k. See Refs. [67, 75] for more details. In particular, two $+1/2$ defects can effectively attract (as long as activity dominates over elastic repulsion) if the initial configuration is chosen properly.

Active turbulence

If activity is sufficiently large, it cannot only modify defect annihilation but can result in the creation of defects. This is due to the activity creating large deformations of the director field which results in the generation of defect pairs that can

separate, instead of instantly annihilating, due to the motility of the +1/2 defect, as described above. Defect pairs thus constantly are created and annihilate, with the overall sum of charges being constant, according to the topological requirement of the Poincaré-Hopf theorem. The system can reach a dynamical steady state where the mean number of defects is on average constant in time [67]. This state of high activity has been called active turbulence, due to the turbulent-like looking flows in the system, or active chaos. However, despite a similar phenomenology the state is very different from classical turbulence, which occurs at high Reynolds numbers [76] unlike the active turbulence at low Reynolds numbers. See e.g. Refs. [77, 78] for more details on this topic.

3D active liquid crystals

So far, we have only considered the case where the director field is a two-component field. However, recently the case of fully three-dimensional active liquid crystals has attracted attention as well, both theoretically and experimentally. Similar to the two-dimensional case, one of the main points that is being investigated is the behavior of defects in these systems. Instead of point defects being present, however, the number of defective structures is much larger in three dimensions (see Sec. 1.1.3 above). The dominant excitations now are disclination lines and loops (that is, closed lines) and point defects. But other, more exotic structures such as knotted disclination lines could be present as well. First steps have been taken towards understanding the dynamics of disclination lines, both experimentally [79–82] and theoretically [83–90], but the understanding is still very incomplete.

1.2.2 Toner-Tu equations

We now introduce the Toner-Tu equations in two dimensions. As mentioned as the beginning of this section, these are used to model dry systems, that is systems where momentum is not conserved. They were introduced in 1995 by J. Toner and Y. Tu in Refs. [91, 92]. In this model there are two dynamical fields, the number density field $\rho(\boldsymbol{r}, t)$ and the velocity field $\boldsymbol{v}(\boldsymbol{r}, t)$. Both of these macroscopic, continuous fields can be defined from the position $\boldsymbol{r}_n(t)$ and velocity $\boldsymbol{v}_n(t)$ of the microscopic constituents, such that

$$\rho(\boldsymbol{r}, t) := \sum_n \delta(\boldsymbol{r} - \boldsymbol{r}_n(t)) \,, \tag{1.34a}$$

$$\boldsymbol{v}(\boldsymbol{r}, t) := \frac{1}{\rho(\boldsymbol{r}, t)} \sum_n \boldsymbol{v}_n(t) \delta(\boldsymbol{r} - \boldsymbol{r}_n(t)) \,. \tag{1.34b}$$

As the velocity of each microscopic particle, $\boldsymbol{v}_n(t)$, gives rise to an orientation of this particle, namely $\boldsymbol{p}_n(t) = \boldsymbol{v}_n(t)/|\boldsymbol{v}_n(t)|$, it is possible to write all of the following equations either in terms of the macroscopic polarization or the velocity. The model then consists of two equations for the two dynamical fields, namely the mass conservation equation, and an equation of motion for the velocity field. Following

Ref. [42] we write these as

$$\partial_t \rho + v\nabla \cdot (\rho \boldsymbol{p}) = 0 \ , \tag{1.35a}$$

$$\partial_t \boldsymbol{p} + \lambda_1 (\boldsymbol{p} \cdot \nabla) \boldsymbol{p} = -\frac{1}{\Gamma} \frac{\delta F_p}{\delta \boldsymbol{p}} \ . \tag{1.35b}$$

where F_p is a free-energy functional, v is the self-propulsion speed of the individual particles, while Γ is a kinetic coefficient. Note that we have ignored noise terms here. In writing the equations in this way it is easy to distinguish the equilibrium contribution (arising from the free energy) from the non-equilibrium ones. Note the difference in prefactor for the second terms on the left-hand side of Eqs. (1.35). The parameter λ_1 has dimensions of speed and determines the strength of the convective term that is similar to the advective term in the Navier-Stokes equation. However, since the system is out of equilibrium and momentum is not conserved, Galilean invariance is broken such that it is possible to have $\lambda_1 \neq v$. Comparing with the corresponding equations for the active liquid crystal, Eqs. (1.25), we see that there is an important difference in that the orientation field $\boldsymbol{p}$ is both the orientational order parameter (as in the active liquid crystal case) and also a current (by being equal to the direction of velocity). Thus, the director field acts on itself through the advection term $(\boldsymbol{p} \cdot \nabla) \boldsymbol{p}$ as well as through the flow alignment term $\delta F_p / \delta \boldsymbol{p}$ in Eq. (1.35b). The former is different from the active liquid crystal model where velocity and polarization were not proportional, and the term reads $\boldsymbol{v} \cdot \nabla \boldsymbol{p}$. The free energy is given by

$$F_p := \int \mathrm{d}\boldsymbol{r} \left[\frac{\tilde{\xi}(\rho)}{2} |\boldsymbol{p}|^2 + \frac{\tilde{\zeta}(\rho)}{4} |\boldsymbol{p}|^4 + \frac{\tilde{\kappa_{\mathrm{F}}}}{2} |\nabla \boldsymbol{p}|^2 - \omega \frac{\rho - \rho_0}{\rho_0} \nabla \cdot \boldsymbol{p} \right] \ , \tag{1.36}$$

where ρ_0 is the average density such that $\rho - \rho_0$ describes the density fluctuations. Using this free energy, the Toner-Tu equations can be written as:

$$\partial_t \rho + v\nabla \cdot (\rho \boldsymbol{p}) = 0 \ , \tag{1.37a}$$

$$\partial_t \boldsymbol{p} + \lambda_1 (\boldsymbol{p} \cdot \nabla) \boldsymbol{p} = -\left[\xi(\rho) + \zeta |\boldsymbol{p}|^2 \right] \boldsymbol{p} + \kappa_{\mathrm{F}} \nabla^2 \boldsymbol{p} - \mu \nabla \rho \ , \tag{1.37b}$$

where $\xi = \tilde{\xi}/\Gamma$, $\zeta = \tilde{\zeta}/\Gamma$, $\kappa_{\mathrm{F}} = \tilde{\kappa_{\mathrm{F}}}/\Gamma$, and $\mu = \omega/(\Gamma \rho_0)$. As before, this free energy is found from symmetry considerations, considering all terms that are in harmony with the polar symmetry of the system. Note that here we do not write all possible terms allowed by symmetry, and that this is a minimal model that will be used in Chapter 2 and has proven to capture the relevant dynamics there. By not including all the terms we disregard possible long-range hydrodynamic couplings. For a further discussion on this see Refs. [93–95]. The full free energy can be found, e.g., in Ref. [42]. The second-to-last term of Eq. (1.36) is the Frank free energy contribution in the one-elastic-constant approximation which penalizes gradients in the director field. The first two terms in Eq. (1.36) control the order-disorder transition that occurs at $\xi = 0$. The parameter is a function of the density and there is a critical density ρ_c at which the transition occurs. For $\rho < \rho_c$, that is $\xi > 0$, the system is isotropic such that $\boldsymbol{p} = 0$ and consequently the mean velocity vanishes

as well. Above the critical density, for $\rho > \rho_c$, i.e., $\xi < 0$, the system assumes an ordered state with uniform orientation and non-vanishing mean velocity, namely $|\boldsymbol{p}_0| = \sqrt{-\xi/\zeta}$ and $\boldsymbol{v} = v_0\boldsymbol{p}_0$. It is then customary to write $\xi = \xi_0\,(1-\rho/\rho_c)$ near the transition. Note that ξ_0/Γ is a rotational diffusion rate and that $\zeta > 0$ is required for stability. Remarkably, this equation allows for long-range order even in two dimensions, evading the Mermin-Wagner theorem, due to its non-equilibrium nature and the convective term generating long-range interactions. This model has been studied in great detail, see, for example, Refs. [42, 96, 97] and references therein.

1.2.3 Vicsek model

The Toner-Tu equations can be seen as a hydrodynamic theory of the agent-based Vicsek model, and coarse-graining the equations of the Vicsek model one can recover the Toner-Tu equations [60]. Unlike the models for active matter described so far, the Vicsek model, introduced by Vicsek et al. in 1995 [98], is not a hydrodynamic theory but describes the evolution of the microscopic active particles in a system. As before for the Toner-Tu equations, each of the N particles in the system has a position $\boldsymbol{r}_n(t)$ and a velocity $\boldsymbol{v}_n(t)$ which gives rise to a polarization $\boldsymbol{p}_n(t) = \boldsymbol{v}_n(t)/|\boldsymbol{v}_n(t)|$. While there are slightly different versions, a commonly used version of the equations of motions in two dimensions is [98, 99]

$$\boldsymbol{r}_n(t+\Delta t) = \boldsymbol{r}_n(t) + v_0\Delta t\boldsymbol{p}_n(t+\Delta t) \tag{1.38a}$$

$$\theta_n(t+\Delta t) = \arg\left[\sum_j C_{nm}(t)\boldsymbol{p}_m(t)\right] + \eta\xi_n(t)\,. \tag{1.38b}$$

Here, Δt is a time step such that the position of the nth particle changes by $v_0\Delta t$ in the direction of its orientation $\boldsymbol{p}_n$. Thus, Eq. (1.38a) is just a deterministic motion with each particle moving with velocity v_0 in the direction of their orientation $\boldsymbol{p}_n = \{\cos\theta_n, \sin\theta_n\}$. The interaction with other particles is included in the equation of motion for the orientation, Eq. (1.38b). Here C_{nm} is the connectivity matrix defined as

$$C_{nm} = \begin{cases} 1 & \text{if } |\boldsymbol{r}_n(t) - \boldsymbol{r}_m(t)| < \mathcal{R} \\ 0 & \text{if } |\boldsymbol{r}_n(t) - \boldsymbol{r}_m(t)| > \mathcal{R} \end{cases}\,, \tag{1.39}$$

where $\mathcal{R}$ is a free parameter that determines the length scale of the interaction. Thus, the connectivity matrix is unity for all particles that are closer than $\mathcal{R}$ to a given particle and vanishes for all particles that are further away. In this way, $\arg\left[\sum_m C_{nm}(t)\boldsymbol{p}_m(t)\right]$ is the angle of the average orientation of all particles that are within a radius $\mathcal{R}$ of a given particle n. Incorporated into the equations of motion for the nth particle as in Eq. (1.38b), this results in an alignment interaction. Each particle tends to align with all the particles near it. In the absence of noise this results in all particles having the same orientation after some time, even if the orientations of all particles are initially random. The presence of the noise term introduces some randomness, and, depending on its magnitude η, acts against the alignment interaction. The Gaussian white noise term $\xi_n(t)$ is uniformly distributed

in $[-\pi, \pi]$. $\eta = 1$ is the maximal noise for which the orientation of each particle is randomized at each time step Δt. In this case there is no alignment at all and the model reduces to a collection of N independent random walkers. Away from this extreme case, there is a competition between the alignment interactions and the noise, and depending on the relative importance of each term the particles in the system are more or less aligned. This can be quantified by the polar order parameter

$$P := |\boldsymbol{P}(t)| = \frac{1}{N} \left| \sum_{n=1}^{N} \boldsymbol{p}_n(t) \right| . \tag{1.40}$$

Since one can fix length and time scales of the model by setting $\mathcal{R} = 1$ and $\Delta t = 1$, there are effectively three free parameters in the model, the density $\rho = N/V$ (where V is the area of the system), the noise amplitude η, and the deterministic velocity v_0. Essentially, at a fixed velocity there is an ordered state if the density is sufficiently high and the noise sufficiently low. The critical noise below which collective motion occurs is found to be $\eta_c \sim \sqrt{\rho}$. Despite the simple equations Eq. (1.38) there has been a lot of debate and research surrounding the model since its introduction over 25 years ago. We refer to Refs. [97–100] and references therein for further details.

1.2.4 Active matter and collective motion in biology

As the last part of this section, we briefly comment on some applications of the models introduced above to biological systems, as this will be one of the main motivating factors in the following chapters. Biological systems are inherently non-equilibrium and active. Often they develop in a situation where many microscopic particles (e.g., cells or bacteria) have to coordinate without, or with only minimal, external guidance. Examples of minimal guidance include the presence of gravity or chemical gradients (chemotaxis). It is then natural to use the theories introduced above, where each microscopic particle is active, i.e., moving on its own, but where macroscopic order still spontaneously develops on system-spanning scales, to explain various phenomena observed in biological systems. The models described above can be adapted to include environmental guidance through the addition of external fields. As we mentioned, which of these models should be used depends on properties of the system one wants to describe, e.g., if momentum is conserved or not. In general, the procedure is to assign to each microscopic particle an orientation and/or a velocity, coarse-grain these quantities over a certain system-dependent scale (think ten particles as the order of magnitude), and then use these averaged quantities in the hydrodynamic active matter models. For the Vicsek model, it being an agent-based and not a hydrodynamical theory, coarse-graining is not necessary. The advantage of this approach is that the resulting equations of motions are relatively simple and easy to analyze, compared with other models that take many microscopic details into account. Furthermore, the generality and abstraction of this approach allows it to extract fundamental principles and guiding mechanisms for self-organization that are universally valid. Many different kinds of pattern formations and collective motions, spanning several magnitudes of size, have been successfully described with models of active matter. For example, dry active models (Toner-Tu equations and Vicsek model) have been used to study and explain the flocking behavior of

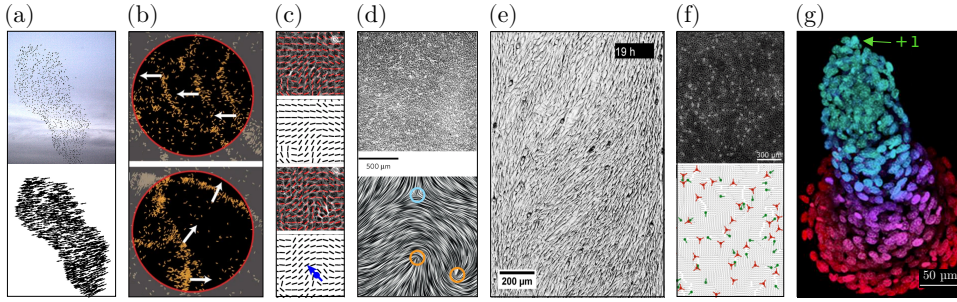

Figure 1.9: Active matter and biology.

(a) Example of a flock of birds, taken from Ref. [104]. (b) Traveling waves in actin motility assays, taken from Ref. [110]. (c) Top: To each MDCK cell (grey) an orientation is assigned, after coarse-graining over the size of a few cells this results in an orientation field (red sticks). This orientation field can be extracted (black sticks) and one can study its dynamic in isolation. Bottom: Using this procedure, one can find topological defects (here $+1/2$ defect, blue arrow). Taken from Ref. [120]. (d) Using the same procedure for cells in a NIH 3T3 monolayer (top), one can find (bottom) $-1/2$ (light blue circle), and $+1/2$ defects (orange circles), the latter of which are motile. Taken from Ref. [121]. (e) The spontaneous flow transition is observed using RPE1 cells in a channel. The cells flow towards the bottom on the left edge of the channel, and towards the top on the right edge. Away from the edges the cells are tilted, compare with Fig. 1.7c. Taken from Ref. [122]. (f) A HBE cell monolayer monolayer (top) is found to be in a state that matches well with the descriptions of active turbulence. There are numerous defects present (bottom, green are $+1/2$ and red are $-1/2$ defects) that constantly annihilate and nucleate. From Ref. [123]. (g) C2C12 myoblasts are confined to a disk in a flat layer. It is observed that a $+1$ defect is present and that after some time the cells move upwards, forming a three-dimensional tower of cells. From Ref. [124].

animal herds, birds, fish, or bacteria [42, 43, 96, 97, 101–109], see Fig. 1.9a, thereby investigating and understanding fundamental principles that underly the dynamics of collective motion and flocking. These principles are independent of many of the microscopic details of the system one considers. It is sufficient that it is possible to assign to each particle (be it a fish, bird, bacteria, etc.) an orientation and a velocity, and that there is some alignment interaction between nearby particles. Apart from flocking, other applications have been to explain traveling waves in actin motility assays [110] (Fig. 1.9b) or cluster formation in myxobacteria [111]. Numerous studies of artificial active systems, such as active colloids or self-propelled hard rods, have also be performed and used to study flocking behavior and collective motion in the lab. See, e.g., Refs. [93, 112–119]. Despite their simplicity (relative to the complexity of biological systems) they have often be able to recreate structures found in biological systems.

Regarding wet active matter, described using the active nematodynamics equations, an important example that much research has focused on is the modelling of cells as active nematics. There are many other applications, such as swimming bacteria [49, 125–131] and, especially, kinesin-driven microtubule solutions [75, 80,

81, 132–135]. However, we now focus on describing some facts about applications in cells and tissues, since this will be very important for many of the following chapters. For an overview over this topic see, for example, the review articles Refs. [42, 49, 52, 136, 137] and references therein. In the following we just outline the general idea and highlight a few successful applications. In confluent cell monolayers one essentially has a two-dimensional layer of cells that covers the substrate almost completely, i.e., there are no holes in the layer. The cells interact via hydrodynamic interactions, thus they are considered wet. The starting point is, again, to assign to each cell a direction. After coarse-graining this results in a continuous director field whose evolution can be tracked over time, see Fig. 1.9c,d. Since each of the cells is motile, the liquid crystal field obtained through coarse-graining is active. Many of the predictions and characteristics of active nematic liquid crystal theory have been recovered in cell monolayers, indicating that it is a decent theory to describe their dynamics. For example, the spontaneous flow transition described in Sec. 1.2.1 has been experimentally observed in cells (Fig. 1.9e) [122, 138], as has active turbulence [123] (Fig. 1.9f), with the theory introduced above being able to qualitatively describe many of the observed dynamics. Furthermore, active $\pm 1/2$ defects have been observed in cells by, e.g., Refs [121, 123, 138, 139], and their dynamics can be described using the active nematic theory as explained in Sec. 1.2.1 (Fig. 1.9c,d). These defects in particular have been linked to biological functions. For example, it has been observed that at the position of defects cell apoptosis (dead cells being expelled from the cell monolayer) takes place [120], or that the growth of protrusions and structure formation in growing embryos is related to the presence of defects [140, 141] (Fig. 1.9g). The latter will be the topic of Chapters 5,6 and we will describe more details about this there.

Lastly, we mention that the field of active matter is much too broad to ever do it justice here, fields like active polymers [46], active solids and metamaterials [142–144], or applications to plants [145–148], were not even mentioned here.

1.3 Differential Geometry

In the last section of this chapter we turn away from active matter and towards differential geometry. The purpose of this section being to introduce some concepts and notations that we will use in some of the following chapters to investigate how active matter interacts with geometry and creates different kinds of shapes. In these chapters we are concerned with two-dimensional surfaces that are embedded (i.e., live in) three-dimensional space. This is different from, e.g., general relativity where, besides dealing with higher-dimensional spaces, there is no surrounding space in which the surface can be embedded. This results in a slightly different approach than one might is familiar with from other fields like general relativity. On the other hand, the fact that the surface is embedded in three-dimensional space makes it very easy to illustrate definitions and concepts. We therefore now introduce definitions used in later chapters in a (hopefully) intuitive and illustrative manner, and refer to other works for a more mathematically rigorous treatment. In the appendix we briefly outline how a connection can be made between the “intrinsic geometry” view

(used, e.g., in general relativity) and the "extrinsic geometry" view commonly used in soft matter.

1.3.1 Tangent vectors, normal vectors, metrics, and all that

We follow in this subsection mainly the review by M. Deserno [149], especially for the parts about external curvature, which is also discussed in Refs. [16, 150]. Literature on the internal curvature is numerous, with basically every book on differential geometry or general relativity introducing it, some references are [14–17, 150–152] As mentioned, we are interested in studying a two-dimensional surface $\mathscr{M}$ that is embedded in three-dimensional Euclidean space $\mathbb{R}^3$. As the surface $\mathscr{M}$ is two-dimensional, it is possible to parametrize it in terms of two coordinates $\{u^1, u^2\}$. For this one can imagine that a subset $\mathscr{S}$ of $\mathbb{R}^2$, i.e., a small, flat surface, is equipped with an usual cartesian coordinate system with coordinates u^1, u^2. This flat surface can then be mapped onto a small part of the curved surface $\mathscr{M}$. That is, the flat surface can be "bent and stretched into shape" by this mapping, and the coordinate system on $\mathscr{S}$ is deformed together with the surface, see Fig. 1.10a. In this way, it is possible to obtain a coordinate system on (a part of) $\mathscr{M}$. Through introducing curvature, we map the two-dimensional coordinate system into three dimensional space and we can write

$$\boldsymbol{X}(u^1, u^2) := \begin{pmatrix} X(u^1, u^2) \\ Y(u^1, u^2) \\ Z(u^1, u^2) \end{pmatrix} . \tag{1.41}$$

This is the parametrization of the surface $\mathscr{M}$, and to each coordinate $\{u^1, u^2\}$ we assign a point $\{X, Y, Z\} \in \mathbb{R}^3$. This mapping is called a chart. We use bold-face characters to denote vectors in $\mathbb{R}^3$. Thus, Eq. (1.41) simply defines a different vector in $\mathbb{R}^3$ for each value of $\{u^1, u^2\}$, and for each different choice of $\{u^1, u^2\}$ the vector $\boldsymbol{X}(u^1, u^2)$ ends at a different point. The union of all these points forms $\mathscr{M}$. For example, for a sphere we can choose $u^1 = \theta \in [0, \pi]$ and $u^2 = \varphi \in [0, 2\pi]$ such that

$$\boldsymbol{X}(\theta, \varphi) = R \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix} . \tag{1.42}$$

Note, however, that in general a single chart is not sufficient to describe the entirety of a given surface. Even for a sphere, this is not possible. Instead, several charts have to be "patched together" in an "appropriate way". The union of several charts is called an atlas. The precise definition of what is meant by "appropriate way" can be found in any introduction to differential geometry, see e.g. Refs. [14–17, 150–152]. Very abridged, the requirement is for different charts to overlap at least partly and in a smooth way. Having chosen a parametrization, it is possible to define several important geometric quantities. First, we define a coordinate system at every point of the surface $\mathscr{M}$. The tangent vectors are defined as

$$\boldsymbol{e}_a := \frac{\partial \boldsymbol{X}}{\partial u^a} \equiv \partial_a \boldsymbol{X} , \tag{1.43}$$

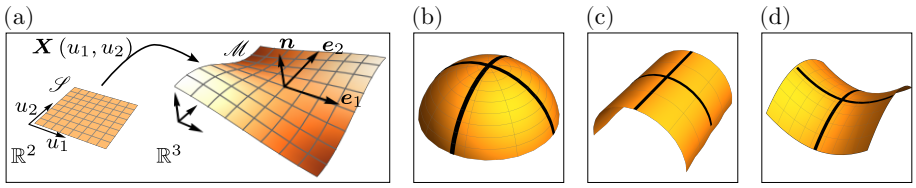

Figure 1.10: Differential geometry.

(a) We illustrate some of the quantities defined in the main text, namely the parametrization $\boldsymbol{X}(u_1, u_2)$ of a two-dimensional surface surface $\mathscr{M}$, as well as the coordinate system $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{n}\}$ defined on $\mathscr{M}$. (b) Cap of a sphere with circles (black) drawn in the direction of maximal and minimal curvature, that is the curves that have the largest and smallest curvature/radius are drawn. As the sphere is symmetric, the two curves drawn have the same curvature. (c) However, for a cylinder one of the curves is completely flat, i.e., the curvature in this direction vanishes. (d) For a saddle surface, the curvature in one direction is positive while the curvature in the other direction is negative. However, the magnitude of the curvature in both directions is identical.

and $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ spans the local two-dimensional tangent plane at each point of $\mathscr{M}$. A third vector, the surface normal, can be constructed from these:

$$\boldsymbol{n} = \frac{\boldsymbol{e}_1 \times \boldsymbol{e}_2}{|\boldsymbol{e}_1 \times \boldsymbol{e}_2|} \,, \tag{1.44}$$

with "$\times$" the ordinary vector product in $\mathbb{R}^3$. Evidently, the surface normal $\boldsymbol{n}$ has unit norm while, in general, the tangent vectors are not normalized when defined as above. These three vectors, $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{n}\}$ form a local coordinate system in $\mathbb{R}^3$, see Fig. 1.10a. They are just ordinary vectors in $\mathbb{R}^3$ that are based at the point $\boldsymbol{X}(u^1, u^2)$ since each of these vectors is (implicitly) a function of the coordinates $\{u^1, u^2\}$. From this coordinate system we can directly define two central quantities, the metric tensor (sometimes called first fundamental form)

$$g_{ab} := \boldsymbol{e}_a \cdot \boldsymbol{e}_b \,, \tag{1.45}$$

and the curvature tensor (sometimes called second fundamental form)

$$K_{ab} := \boldsymbol{e}_a \cdot \partial_b \boldsymbol{n} = -\boldsymbol{n} \cdot \partial_b \boldsymbol{e}_a = -\boldsymbol{n} \cdot \partial_{ab} \boldsymbol{X}. \tag{1.46}$$

Here, "$\cdot$" is the scalar product in $\mathbb{R}^3$. In Eq. (1.46) we used $\boldsymbol{e}_a \cdot \boldsymbol{n} = 0$ in the second step. Note that, strictly speaking, g_{ab} and K_{ab} are not tensors themselves but components of tensors in the coordinate system we chose, with the full (index-free) tensor given by $\boldsymbol{g} = g_{ab} \mathrm{d}u^a \mathrm{d}u^b$ and $\boldsymbol{K} = K_{ab} \mathrm{d}u^a \mathrm{d}u^b$. These are the truly coordinate-independent quantities that are invariant under coordinate transformation. The metric is used to raise and lower indices, e.g.,

$$\boldsymbol{e}^a = g^{ab} \boldsymbol{e}_b \,. \tag{1.47}$$

The tangent and normal vectors are related via the Weingarten equation $\nabla_a \boldsymbol{n} = K_a^b \boldsymbol{e}_b$ and the Gauss equation $\nabla_a \boldsymbol{e}_b = -K_{ab} \boldsymbol{n}$. The second fundamental form satisfies the Mainardi-Codazzi equation $\nabla_a K_{bc} = \nabla_c K_{ab}$.

From the metric and curvature tensor one can define the following quantities. The area element $\mathrm{d}A$ of the surface $\mathscr{M}$ is defined as

$$\mathrm{d}A := \sqrt{g}\mathrm{d}u^1\mathrm{d}u^2 \,, \tag{1.48}$$

were $g = \det g_{ab}$ is the determinant of the metric. The area element quantifies the size of an infinitesimal area element of $\mathscr{M}$ at each point of the surface. The trace and determinant of the curvature tensor are two important scalars that quantify the curvature of a given surface. We call

$$H := \frac{1}{2} g_{ab} K^{ab} = \frac{1}{2} K^a_a \tag{1.49}$$

the mean curvature and

$$K_\mathrm{G} := \det K^a_b \tag{1.50}$$

the Gaussian curvature, both of which are defined at each point of the surface. Note that the mean curvature H is sometimes defined without the factor $1/2$. These quantities can be interpreted as follows. Thinking back to the two-dimensional flat patch of $\mathbb{R}^2$ that is being bent and stretched into the shape of a part of $\mathscr{M}$, the metric describes exactly how the patch, and the coordinate system $\{u^1, u^2\}$ on the patch, have to be stretched and deformed in order to match the surface. If the patch is not deformed at all, the metric reduces to the identity matrix. If the diagonal elements are > 1 (< 1) the patch is being stretched (compressed) in this direction. Non-zero off-diagonal elements correspond to twisting. In this sense, the metric quantifies how the coordinate system of $\mathscr{M}$ behaves and it thus quantifies how to measure distances and angles on $\mathscr{M}$. Remember that in Eq. (1.45) the metric is defined using the tangent vectors on $\mathscr{M}$ which are not orthogonal. The curvature tensor, on the other hand, describes how the surface is embedded in the surrounding space. In particular, it is straightforward from its definition (Eq. (1.46)) to see that the curvature tensor quantifies how much the surface normal changes in the direction of the tangent vector. Since the surface normal has always norm one, only its direction changes. Thus, if the surface is flat, the direction of the surface normal does not change and the curvature tensor trivial. If the surface is bending towards the direction of $\boldsymbol{n}$ in one direction, this component of the curvature is negative, while if it bends in the opposite direction, it is positive. Both the metric tensor and the curvature tensor thus quantify different aspects about how $\mathscr{M}$ curves and bends. Note that the definition of the metric does not contain any reference to the surface normal while the curvature tensor does. Therefore, it is possible to define the metric even if the surface is not embedded in another space, here $\mathbb{R}^3$, while the definition of the curvature tensor relies on this embedding. One thus says that the metric quantifies internal curvature (which one would see and be able to measure if one lived on $\mathscr{M}$ without knowing about the surrounding space), while the curvature tensor quantifies external curvature, measuring how a given surface is bending and curving in the surrounding space. For the latter one does not need to know anything about distances on the surface, one simply can consider from an outside point-of-view how the surface (normal) is changing in space in different directions. However,

the two are, in fact, not independent from each other. To see this note that it is possible to construct the so-called Riemann tensor from the metric alone, thus it is a purely internal quantity, independent of, and not relying on, the surrounding space. Nevertheless, in two dimensions the Riemann tensor is found to be

$$R_{abcd} = K_{\mathrm{G}} \left[g_{ac} g_{bd} - g_{ad} g_{bc} \right] \ , \tag{1.51}$$

thus the Gaussian curvature (an external quantity) enters, despite the original definition of the Riemann tensor not containing any reference to the curvature tensor. From contracting two indices one finds the Ricci tensor

$$R_{ab} = g^{ac} R_{abcd} = K_{\mathrm{G}} g_{ab} \ . \tag{1.52}$$

Taking the trace of the Ricci tensor, one obtains a simple equation relating internal and external curvature, namely

$$R = g^{ab} R_{ab} = 2K_{\mathrm{G}} \ . \tag{1.53}$$

This relation is called Theorema Egregium. Thus, metric and curvature tensor are not as independent as could be thought at first glance.

The mean curvature can be interpreted as follows. At each point, one can associate with each direction in the tangent plane a curvature that states how much the curvature of $\mathscr{M}$ changes in this direction. The mean curvature is the average of the largest and smallest value at each point. The Gaussian curvature, on the other hand, is the product of the two. For a sphere both curvatures are the same, namely $1/R$, due to the symmetry of the sphere, see Fig. 1.10b. However, for a cylinder the Gaussian curvature is zero (Fig. 1.10c) while for a saddle point the mean curvature is zero (Fig. 1.10d). As mentioned above, Gaussian curvature is related to internal curvature. Thus, if the Gaussian curvature is zero there is no internal curvature and the shape can be obtained by bending a flat surface. That is, no stretching of the flat surface is required. It is, for example, possible to bent a flat sheet of paper into a cylinder but not into a sphere, because the paper cannot be stretched. As the sphere has non-zero Gaussian curvature, to deform a flat surface into a sphere the original surface has to be stretched and distances are distorted. From another point of view, if the curvature is zero a single chart is sufficient to obtain a coordinate system on the entire surface $\mathscr{M}$. Thus, mean curvature can be seen as making a statement about how a surface is embedded in $\mathbb{R}^3$, while Gaussian curvature contains information about the internal curvature. If one was living on a surface $\mathscr{M}$ without having a notion of $\mathbb{R}^3$, then one would not be able to distinguish between a cylindrical and a flat $\mathscr{M}$.

A brief comment on the bold-face and index notion is in order which hopefully clarifies the potentially somewhat confusing notation. Every symbol that is bold-face is a vector (or matrix) in $\mathbb{R}^3$, thus it can be written as a three-component vector, as we have done, for example, when defining the position vector $\boldsymbol{X}$ in Eq. (1.41). They are thus in reference to the three-dimensional coordinate system of $\mathbb{R}^3$. Bold-face quantities just being ordinary vectors (or matrices) in $\mathbb{R}^3$, they can be multiplied through the scalar product or the vector product with each other.

However, some quantities, e.g. the tangent vectors, have an additional index. In the definition, these are just used to enumerate and distinguish the two different tangent vectors that are present. The indices are in reference to the two-dimensional coordinate system on the surface $\mathscr{M}$. It is then possible to define another, different scalar product with respect to the indices, namely $\boldsymbol{y} := \boldsymbol{x}^a \boldsymbol{x}_a = g^{ab} \boldsymbol{x}_a \boldsymbol{x}_b$, where $\boldsymbol{x}_a$ is some arbitrary quantity that could be taken to be, e.g., the tangent vector. This is different from the scalar product above in that the resulting quantity $\boldsymbol{y}$ is a scalar in that it does not carry any index. However, it being bold-faced, it is still a vector in $\mathbb{R}^3$. From the "surface-$\mathscr{M}$-point-of-view" a bold-face quantity without index is just a scalar, for example $\boldsymbol{X}$ is just a point on the surface $\mathscr{M}$. If one lived exclusively on the surface $\mathscr{M}$ without being aware of the surrounding space, one would not notice the vector-character that one sees when one lives in $\mathbb{R}^3$, looking onto $\mathscr{M}$ from outside. When taking the scalar product in $\mathbb{R}^3$ one thus ignores potential indices while, when taking the scalar product on $\mathscr{M}$, one ignores the potential bold-face character.

As an illustrative example we consider the sphere. From Eq. (1.42) it is straightforward to find the tangent vectors

$$\boldsymbol{e}_\theta = R \begin{pmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ -\sin\theta \end{pmatrix}, \qquad \boldsymbol{e}_\varphi = R \begin{pmatrix} -\sin\theta\sin\varphi \\ \sin\theta\sin\varphi \\ 0 \end{pmatrix}. \tag{1.54a}$$

The metric is found from Eq. (1.45) to be

$$g_{ab} = R^2 \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}. \tag{1.54b}$$

To see the difference in the two scalar products note that, for example, by definition

$$\begin{aligned} g_{\theta\theta} = \boldsymbol{e}_\theta \cdot \boldsymbol{e}_\theta &= R \begin{pmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ -\sin\theta \end{pmatrix} \cdot R \begin{pmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ -\sin\theta \end{pmatrix} \\ &= R^2 \left(\cos^2\theta\cos^2\varphi + \cos^2\theta\sin^2\varphi + \sin^2\theta\right) = R^2 . \end{aligned} \tag{1.54c}$$

On the other hand,

$$\boldsymbol{e}^a \boldsymbol{e}_a = g^{ab} \boldsymbol{e}_a \boldsymbol{e}_b = g^{\theta\theta} \boldsymbol{e}_\theta \boldsymbol{e}_\theta + g^{\varphi\varphi} \boldsymbol{e}_\varphi \boldsymbol{e}_\varphi = \frac{1}{R^2} \left[\boldsymbol{e}_\theta \boldsymbol{e}_\theta + \frac{\boldsymbol{e}_\varphi \boldsymbol{e}_\varphi}{\sin^2\theta} \right] \tag{1.54d}$$

does not carry any index, but is a 2-tensor in $\mathbb{R}^3$. One could thus say that $\boldsymbol{e}_a$ is a 3×2 matrix and not just a tangent vector. This is connected to a more rigorous connection between the two viewpoints which we briefly explain in Sec. 1.A below. For completeness, we note that $H = 1/R$ and $K_\mathrm{G} = 1/R^2$ for a sphere.

Finally, as the last part of this subsection, we introduce the covariant derivative. When taking an ordinary derivative, one compares how a given quantity changes between two infinitesimally close points. Usually one considers them to be at a distance h and then takes the limit $h \to 0$. When taking derivatives on a curved

surface it has to be taken into account that not only the quantity might change from point to point, but that the surface is changing as well. For example, as mentioned above, the metric relays informations about how to measure distances on a surface. If the metric is not constant, than this has to be taken into account when taking derivatives to ensure that, very loosely speaking, one always "compares the same infinitesimal distances" when taking derivatives at different points. "h now depends on the position on the surface" and is not constant anymore. One thus introduces the covariant derivative ∇_a, replacing partial derivatives in flat space. For a scalar quantity (that is, no index) the two agree, $\nabla_a X = \partial_a X$, for a one-index quantity one has

$$\nabla_a X^b = \partial_a X^b + X^c \Gamma^b_{ac} \,, \tag{1.55a}$$

$$\nabla_a X_b = \partial_a X_b - X_c \Gamma^c_{ab} \,. \tag{1.55b}$$

Here Γ^a_{bc} are the so-called Christoffel symbols that are the corrections due to curvature. Accordingly, they are defined from the metric

$$\Gamma^a_{bc} := \frac{1}{2} g^{ad} \left[\partial_b g_{cd} + \partial_c g_{bd} - \partial_d g_{bc} \right] \,. \tag{1.56}$$

If the surface is flat and the metric constant, then the Christoffel symbols vanish and the partial and covariant derivatives agree. Taking the covariant derivative of higher rank-tensors (more indices) results in more correction terms, see, e.g., Ref. [16, 149, 152]. However, the metric is special and $\nabla_a g_{bc} = 0$ is always true, i.e., the metric is covariantly constant (by construction). Note that a scalar quantity in this context is one without indices, thus $\nabla_a X = \partial_a X$ and $\nabla_a \boldsymbol{X} = \partial_a \boldsymbol{X}$, even though the latter is bold-face and thus a vector in $\mathbb{R}^3$. However, since we are taking derivatives on $\mathscr{M}$, we are interested in if a quantity is a scalar, vector, ... from the internal-geometry point-of-view. These are the basic geometric quantities we will need in later chapters. More specialized concepts will be introduced there when needed.

1.3.2 Hydrodynamic equations on curved surfaces

In the preceding sections we explained the hydrodynamic equations that we will use in following chapters. However, we wrote these equations only for a flat, two-dimensional surface. While sufficient for Chapters 2, 3 in Chapters 5, 6, 7 we will need the equivalent of these equations on curved surfaces. A straightforward way of finding these equations is to simply replace all quantities in the flat equations with their curved counterparts. That is, partial derivatives become covariant derivatives, indices are raised and lowered using the metric of $\mathscr{M}$, vectors are considered vectors on the surface $\mathscr{M}$ etc. While this is quit ad-hoc, there are more rigorous justifications for this procedure which result in the same equations. It is possible to find additional terms which are non-trivial only for curved surfaces, that are not found using this method. For example, explicit couplings between the nematic director and the extrinsic curvature have been studied [153–159]. We will ignore such terms for simplicity in the following chapters.

1.3.3 Helfrich Hamiltonian

An intimate connection between differential geometry and soft matter physics is found in the study of lipid membranes. These are two-dimensional surfaces that are liquid in the sense that in-plane shear stresses are negligible. Often, when one is interested in the dynamic on the scale of the membrane, one can ignore microscopic details and study a coarse-grained model. A popular example is the Helfrich model. The Helfrich free energy reads [149, 160, 161]

$$F_{\mathrm{H}} := \int \mathrm{d}A \left[\gamma + \kappa_{\mathrm{B}} \left(H - H_0 \right)^2 + \kappa_{\mathrm{G}} K_{\mathrm{G}} \right] , \tag{1.57}$$

where H_0 is the so-called spontaneous curvature. The elastic constants of this model are the surface tension γ, the bending modulus κ_{B}, and the Gaussian bending modulus κ_{G}. This free energy can be found as an expansion up to terms of $\mathcal{O}\left(\mathrm{length}^{-3}\right)$ in the curvature [149, 162]. The different terms can be interpreted as follows. The surface tension γ couples only to the area element $\mathrm{d}A$ and thus penalizes changes in the area. If it is positive (as is usually the case) it works towards minimizing the surface area. The second term punishes deviations of the mean curvature H from the spontaneous curvature H_0, which can be seen as the "preferred" mean curvature. The elastic constant associated with this term is the bending modulus κ_{B}. Finally, the third term contains the Gaussian curvature. Remarkably, using the Gauss-Bonnet theorem, it can be shown that this term is equal to a topological invariant (Euler characteristic) of the membrane. Thus, the value of $\int \mathrm{d}A K_{\mathrm{G}}$ is constant for arbitrary continuous deformations of the surface as long as the topology does not change, e.g., through the rapture of the membrane or the creation of holes. This is true for a closed vesicle, if the membrane has a boundary one needs to consider an additional boundary term. Thus, if one considers only changes that keep the topology constant one can ignore this term. Accordingly, the Gaussian bending modulus κ_{G} does no enter into the so-called shape equation which is found from minimizing the Helfrich free energy, and whose solution yields the equilibrium shape of the membrane. Introducing a pressure to ensure volume conservation through adding a Lagrange-muliplier term $-PV$ to the free energy Eq. (1.57), one finds [149, 162–166]

$$2\gamma H - \frac{\kappa_{\mathrm{B}}}{2} \left\{ \Delta H - 2 \left(H - H_0 \right) \left[\left(H - H_0 \right) H - 2H^2 + K_{\mathrm{G}} \right] \right\} = P \, . \tag{1.58}$$

The derivation of this equation is quite lengthy, and we refer to Chapter 7 where we present the derivation. This is a very complex equation, with fourth-order derivatives when written in terms of the surface parametrization, and can be solved exactly only in a few cases in the presence of symmetries that reduce the complexity of these equations. In the case of vanishing bending modulus it reduces to the well-known Young-Laplace law $2\gamma H = P$ which on its own is highly non-trivial, encompassing, as a special case, the minimal surface equation $H = 0$. Even only this very special case of the full shape-equation gives rise to the complex field of minimal surfaces [167, 168]. The Helfrich model has been studied in great detail and much more can be said about modifications to this simple model, connections to microscopic

models for membranes, bilayers, etc. We refer to the extensive literature for further details, see, e.g., Refs. [149, 169–172] and references therein. Furthermore, the Helfrich Hamiltonian is closely related to the Willmore functional $\int \mathrm{d}AH^2$ which has a long and fruitful history as well. Lastly, a brief comment about the assumption of a lipid membrane. As mentioned above, this entails that no in-plane shear stresses are present. Basically, there are no in-plane elastic stresses in the membrane but only normal stresses that bend and curve it[6]. If one relaxes this assumption, for example because one wants to derive the shape equation of an elastic solid, one needs to take additional elastic stresses into account. These arise because curving a surface results in a change in the metric which corresponds to a change of internal distances. If the material is elastic these give rise to additional energetic contributions that depends on the Young and Poisson moduli of the material. See, e.g., Refs. [1, 10, 173, 174] for more on this. We will return to this point when considering a liquid crystal coupled to a deforming surface, in Chapters 5, 6, 7. This concludes this introductory chapter we now give a brief outline of the following chapters of this thesis.

1.4 Outline

In **Chapter 2** we ask the question how ordered states arise in a system where particles tend to align their directions with their neighbors. That is, what are the self-organizing dynamics and principles due to which an initially completely disordered state evolves, over time, into a state with almost perfect order and with only minimal or no environmental guidance. Comparing experimental results of active colloids with simulations and analytics, we identify the crucial role topological defects play in this process. The results are not specific to the experimental system considered but give insight into general principles of self-organization of systems that can be described by the Toner-Tu equations, for example flocking animal herds or bacteria.

In **Chapter 3** we consider an active nematic liquid crystal but start from a microscopic picture where the microscopic particle exerts forces not along its orientation but at an angle to it. That is, the force dipole is not aligned with the orientation of the particle. Coarse-graining this picture, we find that a new term, the active chiral stress, is added to the standard equations of active nematodynamics introduced above. We investigate how this new term influences the dynamics of active $\pm 1/2$ defects and the spontaneous flow transition. We use the results to explain recent experimental observations of chirality in cell and tissue dynamics.

In **Chapter 4** we investigate the question of the possible origin of chirality in cells. It appears from experimental observations that some kind of cells always appear with a certain chirality (handedness). However, it is unclear why evolutionarily one chiral state should be preferred over the other. In this chapter we describe a simple model that can be used to explain the transition to homochirality. That is, starting from a system where both chiral states are present, we find that over time

[6] This is closely related to the diffeomorphism invariance of the surface and the reason one can write the Helfrich free energy in terms of geometric quantities only.

one of the two states disappears. This occurs without the need for explicit symmetry breaking and which of the two possible chiral states dies out is purely due to chance. More generally, our model can be used as a starting point to investigate how properties persist or disappear during the evolution of a system that can be described using the Vicsek model.

Chapter 5 is concerned with the interaction of active topological defects and elastic surfaces. This research is motivated by experimental evidence that topological defects might play the role of topological morphogens. That is, topological defects could be responsible for the formation of protrusion or limbs during morphogenesis, the biological process of shape-development in tissues and organisms. We develop an analytical model to investigate the interaction of activity, topological defects, and shape deformations of elastic surfaces. In a simple setting (an initially flat disk) we can show exactly how the presence of activity can favor the buckling instability of surfaces. We complement the analytical model with simulations to go beyond the linear instability and investigate the system for larger values of activity.

We continue investigating the interaction of active defects and elastic surfaces in **Chapter 6**. Instead of a flat disk we now consider a sphere as the initial geometry. We investigate both polar and nematic active liquid crystals. We find a wealth of different dynamics, which differ greatly for different signs of activity and whether the liquid crystal has polar or nematic symmetry. For example, we find that the sphere flattens and can even transition to a torus, or that the sphere elongates into a spindle-like shape.

Finally, in **Chapter 7**, we present the full derivation of the model used in Chapters 5, 6. As the derivation is rather lengthy, we have included a brief version of the derivation of the equations of motion in the respective chapters, and present the full derivation in this separate chapter, rather than as an appendix. Starting from previous work on passive elastic surfaces, we systematically include active stresses and liquid crystalline order to obtain a general set of equations that can be used to describe active liquid crystals coupled to an elastic surface. From these general equations we derive the equations used in Chapters 5, 6. Furthermore, we present some preliminary results for the active-buckling instability for defects different from charge +1, which was the case analyzed in detail in Chapter 5.

1.A Connection Between Extrinsic and Intrinsic Geometry Point of View

In this section we briefly introduce another point-of-view on the tangent vectors $\boldsymbol{e}^a$. As before, there is a mapping from a subset of $\mathbb{R}^2$ to the space $\mathbb{R}^3$ (or generally, another three-dimensional manifold). This mapping is defined as

$$u^i \to X^\mu\left(u^i\right). \tag{1.59}$$

Here, as before, X^μ is the position vector defining the manifold $\mathscr{M}$ that is embedded in $\mathbb{R}^3$. Now, in this section, latin indices take values $\in [1,2]$ and refer to components of tensors on $\mathscr{M}$. Greek indices, on the other hand, refer to components of tensors on $\mathbb{R}^3$ and consequently take values $\in [1,2,3]$. The metric of $\mathbb{R}^3$ is just the identity

matrix $\delta_{\mu\nu}$ which is used to raise and lower *greek* indices. The components of the tangent vectors of $\mathscr{M}$ are defined as before as

$$e_i^\mu := \partial_i X^\mu \ . \tag{1.60}$$

Note that this matrix has both greek and latin indices. It can thus be used to convert between latin and greek indices, i.e., project quantities defined in the coordinate system of $\mathbb{R}^3$ onto $\mathscr{M}$. For example, the induced metric on $\mathscr{M}$ is defined as the projection of the metric of $\mathbb{R}^3$ as

$$g_{ij} := \delta_{\mu\nu} e_i^\mu e_j^\nu \ . \tag{1.61}$$

This is the metric on $\mathscr{M}$ and is used to raise *latin* indices. Thus, while they are called tangent vectors, e_i^μ actually carries two indices and is thus a 2×3 tensor. However, instead of greek and latin indices it is common in the (soft matter) literature to use only one set of indices and use bold-face characters instead of latin indices. Thus,

$$X^\mu \longrightarrow \boldsymbol{X} \qquad e_i^\mu \longrightarrow \boldsymbol{e}_i \qquad e_i^\mu e_{\mu j} = \delta_{\mu\nu} e_i^\mu e_j^\nu \longrightarrow \boldsymbol{e}_i \cdot \boldsymbol{e}_j \tag{1.62}$$

etc. This is easily generalized to cases where a surface is embedded in a manifold different from $\mathbb{R}^3$ by considering a more general metric than the identity matrix. We will use the bold-face notation in the following chapters.