Bird Flocks as Condensed Matter

Andrea Cavagna[1,2] and Irene Giardina[1,2]

[1]Institute for Complex Systems, Consiglio Nazionale delle Ricerche, 00185 Rome, Italy; email: irene.giardina@roma1.infn.it

[2]Department of Physics, Sapienza University of Rome, 00185 Rome, Italy

Annu. Rev. Condens. Matter Phys. 2014. 5:183–207

First published online as a Review in Advance on January 10, 2014

The *Annual Review of Condensed Matter Physics* is online at conmatphys.annualreviews.org

This article's doi:
10.1146/annurev-conmatphys-031113-133834

Abstract

Flocking is a paradigmatic case of self-organized collective behavior in biology and a living example of active matter. Several models and theories have been developed in recent years to address these kinds of systems. However, unlike granular materials and biological systems at the microscale, experiments have been scarce until recently, preventing the necessary comparison between theory and data. In this review, we discuss a novel approach to flocking, in which experimental data are used as a starting point to empirically characterize flocking as a collective phenomenon—as the term is understood in statistical and condensed matter physics—and build models directly from the data.

1. INTRODUCTION

Collective behavior is a fundamental concept in physics, lying at the core of phase transitions in condensed matter. The emergence of order and the progressive onset of long-range correlations have been deeply investigated, and sophisticated theoretical approaches, experimental protocols, and numerical techniques have been developed. This methodology represents a powerful tool of investigation to also address collective phenomena in other fields, and physicists have become progressively interested in applying the physics framework to problems of interdisciplinary origins.

Biology presents unique challenges in this respect. Collective behavior is widespread in biological systems, occurring at several scales and levels of complexity. Cooperative dynamics has indeed been observed in such diverse systems as bacterial colonies, swimming cells, insect swarms, fish schools, bird flocks, and wildebeest herds, spanning an enormous range of length scales from micrometers (bacteria) to kilometers (mammals) (1). In many cases, collective behavior arises without centralized control. There is no group leader, internal compass, or external stimulus to guide individuals toward the common pattern. Rather, coordination occurs spontaneously as a consequence of the local interactions between individuals. The example of flocking birds is paradigmatic: Each individual bird flies in the same direction as its neighbors, and this local tendency to align gives rise to a coherent moving flock. This mechanism, which produces global patterns from local rules, is known in the biological literature as self-organization (2–5) and very much resembles ordering in condensed matter systems.

Understanding the process of group formation, from individual dynamics to collective motion, is a key issue in biology. Animal aggregations exhibit remarkable coordination and adaptability, and group behavior is often important in biological functions, such as foraging, mating, antipredation, and migration (6, 7). Individuals in a group must interact in a relatively simple way, determined by their limited sensory and cognitive abilities. How does self-organization emerge in a population of simple interacting individuals? What are the dynamical rules followed by individuals that lead to global coordination? To what extent do mutual interactions regulate collective efficiency and group function? Beyond the biological aspects of collective animal behavior, these questions are also relevant from a control point of view: Given a network of units, be they living, natural, or artificial, how can we design an optimal distributed control protocol?

Unfortunately, there are not yet clear answers to these questions. The intriguing aspects of collective animal behavior have fostered an intense interdisciplinary effort in the past ten years. Many models and theories have been developed to describe these systems. Biologists (8–15) have focused on more-detailed models, which can better pinpoint the scaling of interactions from single to multiple individuals in small groups. Physicists, on the contrary, used a different perspective by looking at flocking as a case of nonequilibrium long-range ordering in a system of active interacting units (see 16–18 for recent reviews). They introduced minimal models in which animals are described as self-propelled particles and studied the statistical properties and the ordering transition in fluids of such particles (19–25). Hydrodynamic theories have also been developed that theoretically predict the large-scale, long-time behavior of self-propelled liquids (20, 26–34). All these models and theories greatly improved our understanding of the mechanisms that lead to ordering and information propagation in active systems, giving rise to a whole area of research in which active matter, both living and inanimate (e.g., granular), is treated within a unified framework. Still, several problems remain unsolved when addressing animal groups.

All models and theories are based on assumptions about the microscopic interactions between individuals. The main ingredients (attraction, alignment, and self-avoidance between neighboring individuals) are reasonable but generic. Very little is known about the sensory-cognitive processes

that regulate mutual adaptation between animals in a group. As physicists, we can argue that the complexity of individuals is redundant, and only a few features of mutual interactions are sufficient to describe the collective behavior of large aggregations. We can expect, for example, that coarse-grained models are correct at asymptotically large scales and lengths once the essential interaction mechanisms and conservation laws are taken into account. However, we must not forget that flocks are finite groups of thousands of individuals, and much of the interesting collective phenomena occur on timescales of a few seconds. It is not evident that the asymptotic regime in which standard hydrodynamic theories are valid is relevant in this case. There can be quantities that are conserved on the scales of the flock dynamics and that must be accounted for. Thus, understanding what is crucial and what is not is in fact not trivial in this case and represents a major open issue.

In this respect, experiments represent an indispensable step in testing and validating theoretical models against data and in inferring information about the microscopic rules obeyed by individuals. Although experiments have been performed extensively on granular materials and living systems at the microscale (35–38), the same is not true for natural animal groups on the move. This is particularly so for groups moving in three dimensions, where the most stunning examples of collective behavior, such as bird flocking, fish schooling, and insect swarming, actually occur (see **Figure 1**). Performing field experiments on three-dimensional (3D) animal aggregations is in fact a formidable task and requires technical skills that go beyond a standard background in biology. As a consequence, empirical studies have been limited for a long time to very small groups (few tens of individuals), restraining the possibility of a reliable statistical analysis at the collective scale. In a way, we can say that the precious and fruitful interplay between physics and biology, and experiments and modeling, which has marked the biophysics revolution at the molecular scale for the past fifty years, only recently involved behavioral biology.

In this review, we describe and summarize a novel approach to collective animal behavior and flocking that is based on experiments and empirically grounded modeling. The main philosophy is to start from experimental data and quantify the collective patterns using concepts and methodologies from statistical and condensed matter physics. We then gather information, directly or

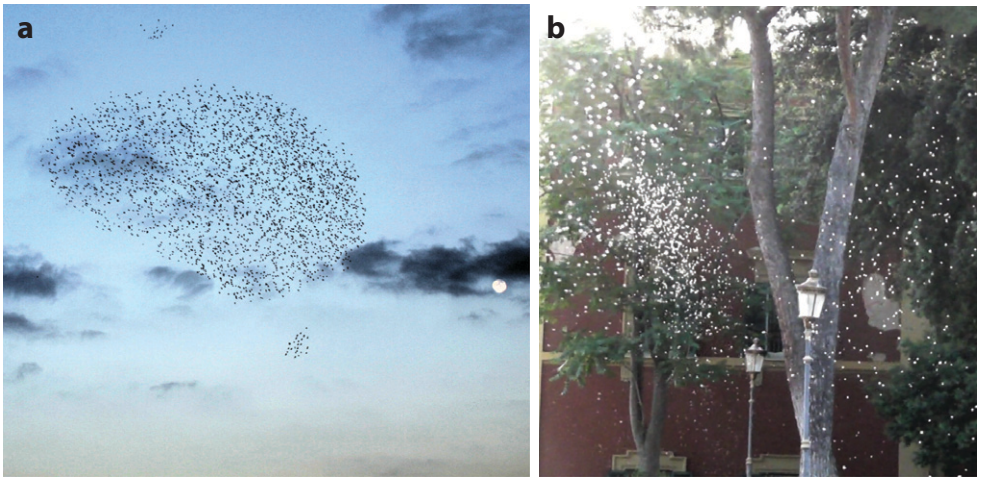

Figure 1

Examples of animal groups. (*a*) A flock of starlings. Copyright FP6-NEST 12682 STARFLAG project, INFM-CNR. (*b*) A swarm of midges. Copyright ISC-CNR.

indirectly, on the microscopic interactions using techniques of statistical inference. Finally, we build models based on these experimental findings.

The review is organized as follows. In Section 2, we summarize the state of the art of empirical studies on animal aggregations, with particular focus on flocking behavior and recent experiments on large groups. In Section 3, we describe how experimental data can be used to characterize flocking from a statistical physics point of view. In Section 4, we introduce methods of statistical inference to build models directly from the data. Next, we examine from a more abstract perspective the requirements needed for a good model. In the final section, we discuss how experimental findings connect with existing theories of flocking, the novel theoretical features that emerge from the data, and the open problems that remain to be addressed at the experimental and theoretical levels.

2. EXPERIMENTS

Collective behavior is a qualitatively different phenomenon, with emerging patterns and statistically stable properties, only when the system is large. For this reason, to investigate collective animal behavior we need empirical data on large aggregations of individuals. This poses serious technical issues, as appropriate experimental techniques must be developed. In standard condensed matter, where the number of particles is very large, collective properties are usually characterized in Fourier space by scattering experiments. The same approach cannot be easily applied to animal groups, where group sizes are orders of magnitude smaller, and movement occurs on the macroscale. There is only one case in which experiments in Fourier space have been performed (39), using acoustic wave-guide remote sensing on vast, oceanic shoals of fish. However, the technique is specific to underwater probing and can hardly be exported. For most animal groups, experimental investigations occur in real space. Experiments adopt a Lagrangian perspective and consist of retrieving information (positions, velocities, trajectories, etc.) about each individual organism in the group, much as is done in PTV (particle tracking velocimetry), where passive tracer particles are used to study turbulence in complex fluid flows (40, 41).

Collecting data of this kind on 3D animal groups is a difficult problem (7). Insects moving in two dimensions are small and can be kept under laboratory control even in large numbers. The same is true when working at the microscale, as with bacteria suspensions (38). However, when dealing with animals moving in three dimensions, individuals often have larger sizes and naturally move in a much larger environment. Laboratory control can be problematic, and techniques for field observations may prove extremely complicated. Several empirical analyses in the past were performed on fish because experiments could be performed in the laboratory under variable conditions (42, 43). Obtaining data for birds is more difficult because this must be done in the field (44, 45, 46).

In all past experiments, the number of individuals in the group was rather small (a few tens), and the group arrangements were often loose. The reason is essentially a technical one. To reconstruct the 3D position of an object, optical techniques (stereometry, orthogonal method, shadow method, etc.) require putting in correspondence (matching) images of it (47). To better understand the nature of these problems, let us consider in more detail the basic principles of stereoscopy. The fundamental idea of stereoscopy is that by taking synchronous images of the same object from two different positions in space, we have enough information to reconstruct its 3D coordinates, provided that we know the mutual position of the two cameras (see **Figure 2*a*,*b***). The case of flocks is, however, more complex because the system to be reconstructed is not a single object but an ensemble of many objects: the individual birds. Given two (or more) images of the flock taken from different points of view, every bird in one photograph must be matched with the

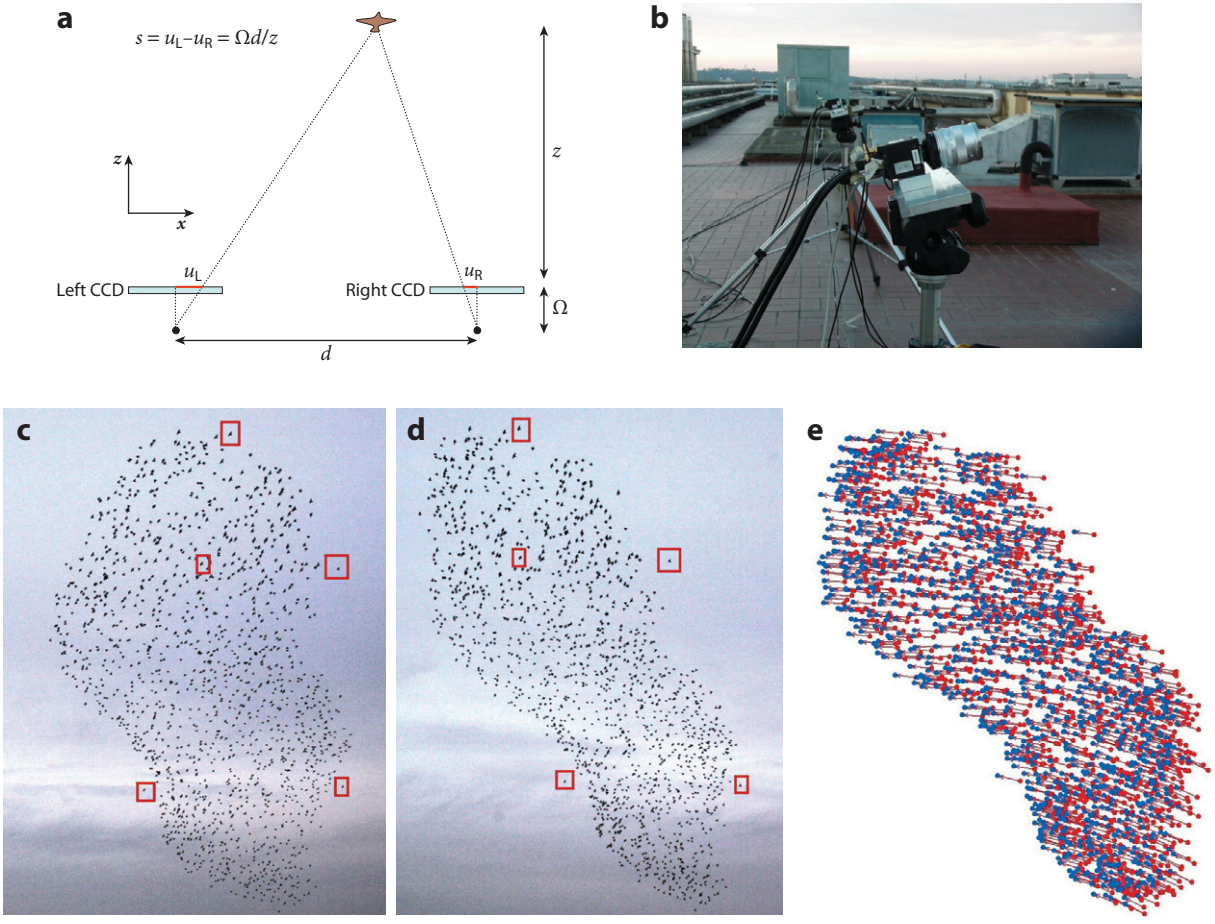

Figure 2

Example of a stereoscopic experiment. (*a*) Simple scheme of a two-camera setup: z represents the distance of the object from the observation point, Ω is the focal length of the lens, s is the stereoscopic shift, and d is the distance between the two cameras (51). (*b*) Three-camera stereoscopic setup during real experiments on starling flocks performed by Cavagna & Giardina and their group. Copyright ISC-CNR. (*c*, *d*) Stereoscopic images of one flock of 1,246 individuals. Cameras were 25 m apart and the flock was approximately 80 m away from the observation points. Red boxes indicate the same birds in each one of the two images. Reprinted from Reference 48. (*e*) Velocity field for the same snapshot as in *c* and *d*, as retrieved with the tracking algorithm. The red points correspond to the 3D positions of the individual birds obtained from the processing of images *c* and *d*, and are displayed with a similar perspective to the one of experiments. The blue points correspond to the 3D positions of the individual birds at the consecutive instants of time. The arrows from red to blue points correspond to the dynamic links established by the tracking algorithm and identify (once divided by the time interval) the individual velocities. In this example, the time resolution is 10 frames per second (fps). Reconstruction of the full individual trajectories typically requires much larger frame rates to deal with occlusions (see Reference 49, in which shootings at 80 and 170 fps were used).

corresponding bird in the second photograph. At this point, stereometric formulas can be applied to recover from each pair of matched individuals in the photo the physical 3D coordinates of the individual in space. This correspondence (or matching) problem becomes extremely severe when groups are large and compact, as flying flocks are, causing photographic images to typically be very dense sets of almost featureless points. An example is given in **Figure 2*c*,*d*** that shows two images of a flock taken from different observation points. Telling which bird is which can be easily done for

birds that appear isolated (e.g., on the boundary of the flock) but is impossible to do with the naked eye for individuals in the dense regions of the flock.

Additional problems are encountered in trajectory retrieval, in which the 3D positions of the same individual at consecutive instants of time must be collated together in a unique sequential path (see **Figure 2*e***). Owing to these limitations, only small groups were analyzed until recently, which is unfortunate in several respects. First, collective behavior is a truly emergent phenomenon only in large aggregations. Second, the statistics for small groups are rather poor, and the surface-to-volume ratio is very large, with many individuals on the border and very few inside. This produces a strong bias in the statistical analysis, which does not allow pinpointing the microscopic origin of the observed collective behavior.

A step forward in this respect was obtained in References 48–52. In these works, the authors performed stereoscopic photographic experiments on large starling flocks in the field and were able to obtain the 3D reconstruction of individual coordinates, velocities, and trajectories in groups of as many as 4,000 individuals. Interestingly, the manner in which the matching problem was solved is another example of interdisciplinary application of physics methods. Standard computer vision techniques address the stereo correspondence by looking for the mathematical operator that transforms one point image into another (47). If the apparatus is perfectly calibrated [i.e., one knows with infinite precision all the internal parameters of the cameras (focal length, position of the image sensor, etc.) as well as the mutual distances and orientations between different cameras], then this operator can be computed and the stereo match can be solved. In practice, calibration is never achieved with infinite precision. Even with very good experimental standards, the transformation operator is only roughly estimated. As a consequence, when there are many objects to be retrieved, the transformation from image to image becomes ambiguous because of noise, high density of points on the images, and multiple occlusions: Typically, for every individual on one photo, there are many possible candidates on the second photo. In this case any naive or greedy approach fails miserably, and the stereo match must be addressed as a global optimization problem. This is what was done in Reference 50. First, a measure was built on the basis of a pattern-recognition principle and epipolar constraints, quantifying the likelihood of any possible match between image points on the two photographs. Then, the matching problem defined by this measure was solved using global assignment.

A further level of complexity arises when addressing 3D tracking, i.e., retrieving individual trajectories in time: The above discussed stereo match intermingles with the matching problem between temporal instances. It turns out that in this case the best performing strategy is also to recast the full matching as a global optimization problem, in which the search space where optimization takes place is now the space of all possible trajectories. This can be done by building a measure, this time between pairs of 2D trajectories, quantifying the likelihood that stereo, dynamical, and geometrical constraints are satisfied (53). The optimal matching is then found using linear programming optimization methods.

3. DATA ANALYSIS

Stereoscopic experiments on flocks in the field provide detailed information on bird dynamics during flocking. More precisely, one has access to the set of 3D positions, velocities, and trajectories for all the individuals in large groups. This allows statistical characterization of the collective properties of the group with the same approach one would adopt in a physical system of interacting particles. Before describing the interesting results one can get in this way, let us make a few methodological comments. Biologists looking at collective motion have mainly focused on aggregate observables: the polarization of the group and its density and volume (everything that

could quantify the global degree of ordering). However, we know from condensed matter and off-equilibrium systems that most of the nontrivial features that arise at collective levels are encoded in multipoint functions and correlations, i.e., quantities that describe how a system is robust with respect to ordering and how it reacts to perturbations. We must also keep in mind that for flocks we do not know a priori what the microscopic interactions are. If we want to get information on the interactions, we need to identify the observables that are most affected by them, and use these observables to probe the interactions themselves.

3.1. Topological Interactions

The statistical analysis of experimental data can conceptually be organized in steps. First, we can look at the structural properties of flocks, namely at how birds are positionally arranged in space. We find that flocks are sparse homogeneous systems with very low packing fractions and trivial gas-like pair correlation functions (49, 51, 54). Thus, standard structural quantities that depend on mutual distances are poorly sensitive to (and informative about) the nature of interactions: Judging from the pair correlation function, a flock would appear more or less as a gas. However, there are other observables, dependent on angles rather than distances, that very clearly characterize a flock as a system of interacting units. For example, given a bird, the angular distribution of its closest neighbor is strongly anisotropic, with the neighbor being more likely located on the sides than along the direction of motion. The degree of this anisotropy can be quantified in a sharp way (48, 55) and measured for first, second, and third nearest neighbors, and so on. An example is shown in **Figure 3*a***, where one can see that anisotropy decays upon increasing the order of the considered neighbors. Given that this anisotropy is a direct consequence of the interactions between individuals (the neighbors' distribution is completely uniform for a noninteracting system), it can be used as a proxy for the interaction itself. This allows an estimate of the range of the interaction as the point—in terms of order of the neighbor (topological range) or, alternatively, of the distance (metric range)—beyond which the anisotropy disappears. For a given flock, the topological and the metric range are equivalent measures of the interaction decay. Indeed, in a homogeneous system of density ρ, the distance $r(n)$ of the n^{th} neighbor is naturally related to the order n of the neighbor as $r(n) \sim \rho^{-1/3} n^{1/3}$ (see **Figure 3*b***). However, the two ranges behave differently when looking at flocks of different densities because, precisely owing to this relationship, either the topological or the metric range can remain constant but not both of them. What the data show (see **Figure 3*c*,*d*** and 48, 55) is that the topological range does not depend on density, whereas the metric one does. In other words, in a flock, individuals interact with the same number of neighbors (approximately seven) regardless of the density the flock.

This kind of topological, density invariant interaction is completely different from all interactions usually studied in physical systems, where forces depend on metric distances. On the one hand, we must not be surprised by this: Birds are not particles, and the way they interact with one another is based on cognitive rather than merely physical processes. On the other hand, given that these interactions are not defined by natural laws but are the outcome of long evolutionary processes, one might wonder whether the topological nature of the interaction and even the value of the range are somehow connected to beneficial behaviors at the collective or individual level. This is certainly a difficult question, but some aspects of it can be quantitatively investigated. For example, numerical analysis of self-propelled particle (SPP) models with either metric or topological interactions shows that topological models are much more robust, in terms of cohesiveness and resilience to fragmentation, to noise and external perturbations (25, 48, 56), a feature that is crucial for antipredatory success. This result is quite natural. A metric interaction is not robust under perturbations simply because each time a bird is driven out of the (metric) interaction range,

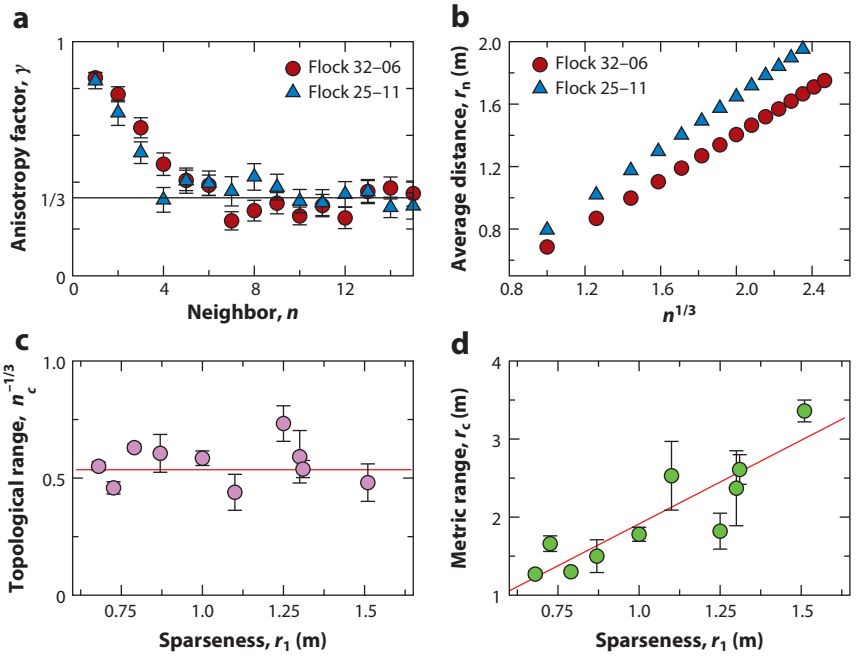

Figure 3

Evidence of topological interactions. (*a*) Degree of anisotropy versus order of neighbor for two different flocks: flock 32-06 and flock 25-11 (numbers indicate the experimental session and event; see Reference 48). The value for a noninteracting system is 1/3. (*b*) Average distance r_n of the n^{th} neighbor plotted against $n^{1/3}$. (*c*) Topological interaction range n_c versus sparseness $r_1 \sim \rho^{-1/3}$. The red line corresponds to the average value of n_c over all events, i.e. $n_c = 6.5$. (*d*) Metric interaction range versus sparseness. The red line is a linear fit of the data. Reprinted from Reference 48.

it loses contact with the flock. It effectively evaporates. This evaporation causes a gradual loss of cohesion of the flock, which is a distinct adaptive disadvantage. Indeed, the predation success rate of a falcon against an isolated starling is much larger than against a united flock. Although the reasons for this are not quite clear, that is a consolidated empirical fact (1, 57–59), showing that cohesion is a strong antipredatory advantage. A topological interaction is ideal in this respect, as interaction contact is not established by a criterion based on distance but on the number of interacting neighbors. Hence, even if the bird is pushed farther from the flock by the perturbation, it is always interacting, hence reuniting swiftly to the flock after the perturbation. Understanding why the number of interacting neighbors is on the order of seven individuals requires more care.

3.2. Why Seven Neighbors?

Experiments on starling flocks not only reveal that interactions between individuals have a topological nature but also seem to point out that starlings interact with (on average) seven neighboring birds. Why seven? There are three possible explanations.

First, this number can be a cortical bottleneck. Interacting with neighbors implies keeping track of several individuals at the same time during the rather complicated business that flying in a flock undoubtedly is. Hence, it is reasonable to expect that there is a cortical limit to the number of neighbors that an individual bird can simultaneously track. Although interacting with more neighbors would grant a more robust cohesion of the group, birds must stick to the largest number

they can manage from a cognitive point of view. Whether this limit is due to prenumeric abilities or other cognitive reasons, it is hard to say (60). Moreover, although seven seems a reasonable number for a cognitive limit (61), we are certainly far from proving that it is the correct number.

A second hypothesis is that the optimal number of interacting neighbors is the product of a trade-off between timeliness and accuracy of the received information (5, 62). According to this hypothesis, interacting with many neighbors is good because each bird receives the information (e.g., falcon attacking from a given direction) quickly. However, the information coming from a few birds is averaged together with many uninformed correlated individuals who follow their original polarized motion, so the information will be washed out due to this social force. Conversely, when interacting with only a few neighbors one gets very clear information but also very late information. The idea is that there may be a number of interacting neighbors that optimizes this balance. As in the previous hypothesis, though, it is very hard to provide hard proof that this trade-off is optimized by seven neighbors in three dimensions. Again, it seems reasonable, but it is definitely not the only possibility.

The third hypothesis, recently formulated in Reference 63, argues that there is another trade-off, that being the trade-off between the advantage of interaction (fostering cohesion and robustness) and the cost of interaction (some cortical price, as in the first hypothesis). The benefit of this third hypothesis is that, under some simplifying assumptions, calculations can actually be made, and hard numbers can be worked out from the experimental data. The main reason why birds within a flock interact is to create and strengthen consensus (in the direction of travel). Without consensus there would be no flock, just total chaos. The first idea in Reference 63 then is to define a mathematical metric that quantifies the disagreement within the network due to disturbances and noise. The inverse of the expected disagreement owing to each node (bird), the nodal robustness, is large when individuals contribute a small amount of disagreement and can be measured in the empirical data of starling flocks. The second idea is to introduce the cost of interaction: Each bird interacts with a fixed number of neighbors (n_c), but each interaction poses a cost because a bird must do several things beyond interacting with its neighbors, such as avoid predators, check for the position of the roost, etc. The key idea of this approach is therefore to compute the per-neighbor contribution to robustness as a function of n_c. The result is quite startling: The relative robustness is maximal between six and seven neighbors. Considering how this approach is different from the anisotropy approach of Section 3.1, the agreement is quite remarkable.

3.3. Scale-Free Velocity Correlations

As a further step in the investigation of collective behavior in flocks, we can look at velocities. Flocks are very polarized groups, with a large degree of global alignment. This can be easily quantified by looking at the so-called polarization, an analog of the global magnetization in ferromagnetic systems. Given that individual speeds can vary from individual to individual, polarization is usually defined in terms of the individual flight directions:

$$\Phi = \left| \frac{1}{N} \sum_i \frac{\mathbf{v}_i}{|\mathbf{v}_i|} \right|. \qquad 1.$$

Polarization is larger than 0.9 in natural flocks, indicating that ordering is very strong. More interesting are the properties of fluctuations and, in particular, their mutual correlations. Fluctuations can be defined in the usual way, i.e., by subtracting from each velocity the global velocity of the flock $\mathbf{u}_i = \mathbf{v}_i - \mathbf{V}$. From a biological point of view, fluctuations describe how much the

behavior (in this case, motion) of an individual differs from the behavior of the group. Snapshots of fluctuation fields clearly show the presence of large correlated domains (see 52). A better description is encoded in the correlation of fluctuations as a function of distance,

$$C(r) = \frac{\sum_{ij} \mathbf{u}_i \cdot \mathbf{u}_j \delta(r_{ij} - r)}{\sum_{ij} \delta(r_{ij} - r)}, \qquad 2.$$

which measures how much a behavioral change of state of an individual in one place of the flock influences individuals located at distance r from it. We know that in thermal equilibrium physics, correlation functions are intimately connected to response: In this respect, by looking at the correlations' intensity, decay, and extension, we can understand how individuals influence each other and how the group as a whole would react to either endogenous or external perturbations. In flocks, the correlations of the velocity fluctuations decay almost linearly from strong correlation at short distances to strong anticorrelation at large ones (see **Figure 4*a***). The zero of the correlation therefore quantifies the extension of the correlated domains and is a faithful finite size estimator of the correlation length ξ. When looking at different flocks, one finds that the correlation length ξ scales very neatly with the size of the flock (**Figure 4*c***), indicating that

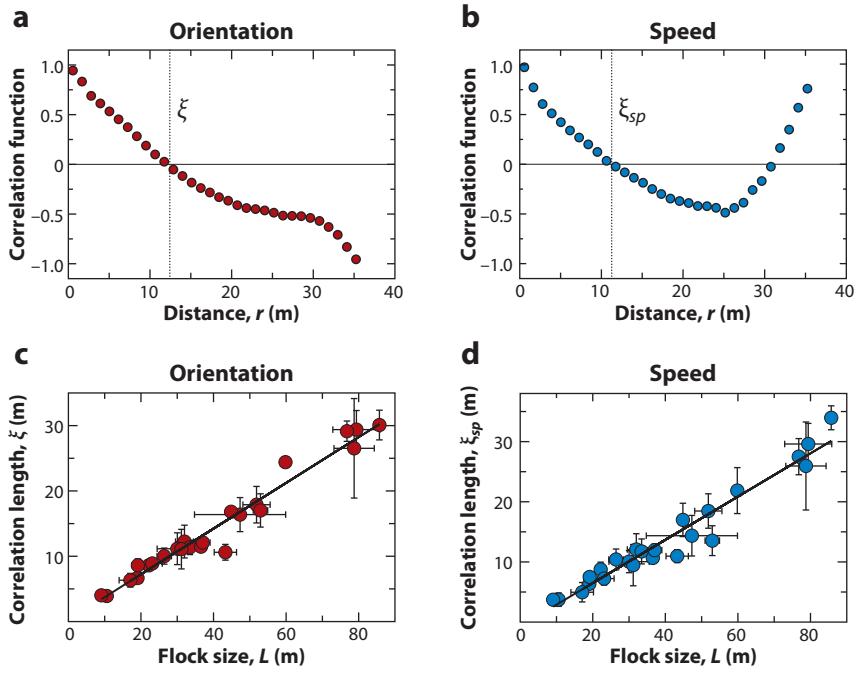

Figure 4

Statistical properties of the velocity fields. (*a*) Correlations of the fluctuations of the velocity orientations as a function of distance for a single instant of a single flock. The vertical line marks the distance where the correlation function crosses the zero axis, and identifies the correlation length ξ. (*b*) Correlation of the speed fluctuations as a function of distance (same flock as in panel *a*). The vertical line marks the distance where the correlation function crosses the zero axis, and identifies the correlation length ξ_{sp}. (*c*) Correlation length of orientational correlations (see panel *a*) versus the flock's size for 22 flocks. (*d*) Correlation length for the speed correlations (see panel *b*) versus flock size. Reprinted from Reference 52.

velocity correlations obey scale-free behavior (52). In other words, there is no intrinsic length scale other than the size of the system, and correlations extend over the whole group. We know that this kind of behavior corresponds asymptotically to a power-law decay of the correlation, which is modulated by finite size effects in finite groups. A scaling analysis reveals that the exponent γ of this power-law decay is in fact very close to zero in flocks: Not only do correlations extend over the whole system, but their intensity remains strong even at large distances. Interestingly, the same scale-free behavior and a similar value of the decay exponent are also found when considering the correlations of the fluctuations of the speed (see **Figure 4*b*,*d***).

From a biological point of view, these findings are telling us that individuals in a flock influence each other very effectively and are fully consistent with the common observation that flocks exhibit a remarkable collective response to attacks and disturbances. From a mechanistic point of view, they are on one hand expected and on the other rather surprising. When thinking of birds moving together, the first intuitive analogy with physical systems is the one of magnetism: There are individual vectorial degrees of freedom, such as the birds' velocities, that tend to align with each other, much as spins in a ferromagnet. In fact, mutual alignment between neighbors is a crucial ingredient of all models of collective motion (2), and imitative allelomimetic behavior[1] is unanimously considered the main mechanism for self-organized behavior (4). Alignment is a rotationally invariant interaction, and we know from equilibrium statistical mechanics that when a continuous symmetry is spontaneously broken, there are residual soft modes in the system that behave in a scale-free manner (64). For example, this is what happens in the Heisenberg model: When the system orders at low temperature, fluctuations of the order parameter that are perpendicular to the spontaneous magnetization (the so-called Goldstone modes) are massless and decay with a power law in the whole ordered phase. This is precisely what happens in flocks: There is spontaneous ordering (i.e., a finite polarization) and velocity fluctuations, which are mainly perpendicular to the global direction of motion, are scale free. In other words, the scale-free behavior that we observe experimentally in velocity correlations is a natural consequence of the fact that flocks are ordered and rotational invariance is broken. This argument can be generalized by taking into account the off-equilibrium nature of flocking, namely the fact that birds do not sit on a static lattice like spins in a ferromagnet but actually move in space. Indeed, hydrodynamic theories (20) have demonstrated that hybridized Goldstone modes also persist in moving fluids of self-propelled particles.

Much less expected is, however, the scale-free behavior found in speed correlations. Speed is not related to any evident symmetry of the interactions and is not a soft degree of freedom, as orientation is. For a flock, all directions of motion are a priori equivalent, whereas not all possible speeds are equally achievable. Energetic and aerodynamics criteria put strong constraints on individual and collective values of speed. For this reason, scale-free speed correlations cannot be explained in terms of symmetry arguments. Some other mechanism must be at play. Hydrodynamic theories of flocking (20) describe a coupling between local density and local speed, and predict long-range fluctuations between densities that would therefore imply long-range speed correlations. Existing analyses on flocks of hundreds to thousands of individuals indicate that such flocks are rather homogeneous in density (54). However, natural flocks have a thin and irregular shape, and available data correspond to short (a few seconds) flocking sequences: It is possible that larger groups and/or longer sequences are necessary to reveal the occurrence of anomalous density

[1]In the behavioral sciences, allelomimetic behavior indicates a range of activities in which the performance of a behavior increases the probability of that behavior being performed by other nearby animals. The term is commonly used to describe imitative behavior in social animals (4).

fluctuations, as those found in SPP models (25, 27). There is, however, another way to explain the scale-free behavior in the speeds. We know from condensed matter physics that scale-free behavior is a signature of criticality. It is possible that in this case as well, some form of critical behavior is at play. Understanding precisely how this might occur and what the control parameters are, are interesting open issues (65).

Another surprising feature of correlations is the value of the decay exponent, which for flocks is very close to zero for both orientations and speed. In equilibrium magnetic systems in three dimensions, this exponent is equal to 1 (64), and in SPP systems it has been predicted to be even larger [$\gamma = 6/5$ in $D = 3$ (27)]. In both cases, it corresponds to a sharper decay of correlation in space. What is the origin of this behavior? A possible mechanism is discussed in Reference 66, where it is shown that off-equilibrium perturbations acting on the boundary of the flock can enhance fluctuation modes of long wavelengths, producing stronger correlations. The premise of this work is that flocks are finite groups moving in a changing environment. This makes a difference with respect to a bulk system. Although it is reasonable that individuals inside the group only regulate their behavior because of interactions with neighbors, the same is not true for individuals on the border who are constantly exposed to external cues or might have border-specific behaviors. In a system with a continuous spontaneously broken symmetry, these border dynamics affect the behavior of the whole system. Thus, the correlation that we observe may be a combined effect of a bulk standard correlation and non-equilibrium (endogenous or exogenous) processes occurring at the boundary. In Reference 66, a simple example of this interplay is given: An idealized flock is described as a sphere of Heisenberg spins, and the perturbations at the boundary are modeled as a field evolving on the surface of the sphere. If the field rearranges on fast-enough timescales and is strong enough, its effect is to pin some long-range modes, making the spatial decay of the correlation slower (see **Figure 5*a*,*b***).

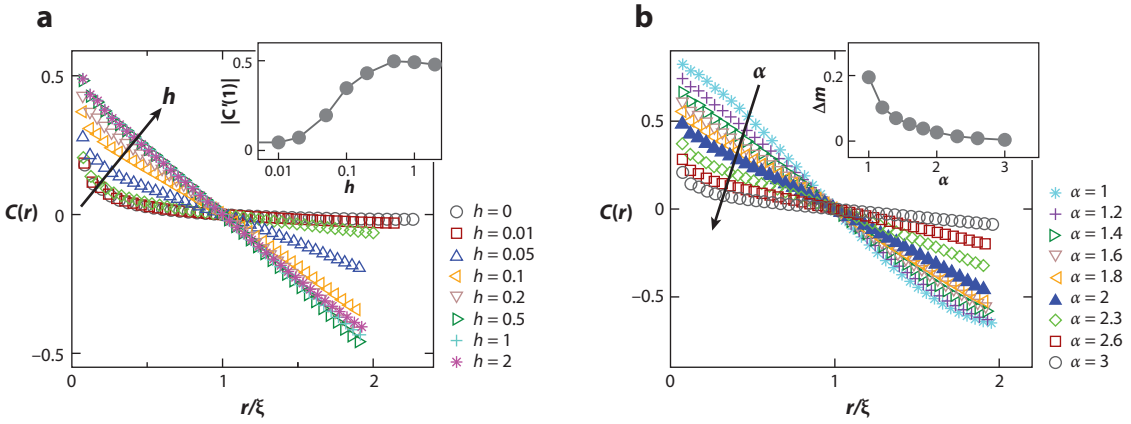

Figure 5

Boundary perturbations: a sphere of Heisenberg spins is considered under the effect of a dynamical field acting on the boundary. The field follows a uniform random walk on the surface of the sphere, and reverses its direction on a timescale, $\tau = R^{\alpha}$, with R being the radius of the sphere. (*a*) Correlation function computed for $\alpha = 2$ at various values of the field intensity: The decay is slower the larger the field (*the arrow indicates the direction of increasing field*). Inset: slope of the correlation in $r = \xi$ as a function of field intensity. A finite size analysis shows that $\gamma = 1$ for $h = 0$ (equilibrium Heisenberg), whereas $\gamma \sim 0$ for $h = 0.5$. (*b*) Correlation function at different values of α (*the arrow indicates the direction of increasing* α). Inset: fluctuations of the magnetization as a function of α. Reprinted with permission from Reference 67.

3.4. Diffusion

The analyses of positions and velocity fields are the first important characterizations of collective properties of flocks. However, one of the preeminent features of flocking is the fact that individuals move, not only in an absolute reference frame but with respect to each other. When naively thinking of birds as flying magnets, we are disregarding the fact that the network on which these "magnets" lie is in fact a dynamical one, continuously changing in time according to the mutual movement of individuals. The role of movement has been greatly investigated in self-propelled particle models and in hydrodynamic theories of flocking. In fact, in two-dimensional SPP systems, movement of individuals is crucial to produce polarized collective motion. Alignment between spins on a fixed 2D lattice is not able to sustain long-range ordering because fluctuations of the order parameter are too strong; however, in SPP systems, convection of individuals contributes to directly transporting information, depressing fluctuations, and stabilizing order (26). This effect is certainly less important in three dimensions, as ordering would be present anyway, but it can change the long-time, large-scale exponents that govern dynamical correlation functions in a nontrivial way (27).

To investigate this issue in real flocks one needs to look at trajectories, which require a much more complex stereodynamical matching than do positions or velocities. Work on this is still ongoing (49); however, the first analysis on trajectory reconstructions (67) shows that reciprocal motion and mutual diffusion are in fact present in natural flocks. Birds obey super-diffusive behavior both in the center of mass reference frame and with respect to each other, with the mean-square displacement growing as $\langle(\delta x)^2\rangle = Dt^{\alpha}$ with an exponent $\alpha \sim 1.7$ (**Figure 6*a***). This suggests that on long timescales, the convective effect described by hydrodynamic theories is indeed present. However, the diffusion coefficient D is very small, causing a very slow rearrangement of the network. For example, in **Figure 6*b*** it is shown that in a few seconds approximately 70% of the first eight neighbors of a given bird remain among that bird's first eight neighbors. If we consider that birds have a reaction time of milliseconds and move at approximately 10 m/s, a few seconds is a long time lag, during which many processes take place. For example, turns typically occur on the order of seconds (68). Some additional mechanisms might be required to explain information propagation on such timescales. Further investigation on retrieved trajectories is awaited to elucidate this issue (68).

To summarize, experimental findings discovered so far indicate that flocks are sparse liquid-like systems in which individuals move and exchange positions, coordinating with the first

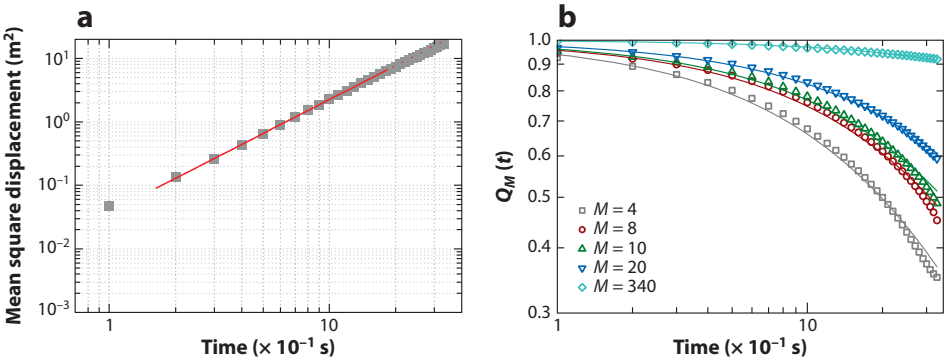

Figure 6

Diffusion of individual birds. (*a*) Mean-square displacement in the center of mass reference frame as a function of time, in logarithmic scale. The diffusion exponent is $\alpha = 1.77 \pm 0.02$. (*b*) Overlap, defined as the number of birds that remain among the first M neighbors of a given individual after a time t, as a function of time. Reprinted from Reference 65.

nearest neighbors with a topological density invariant interaction. Results on velocity fields support the idea that alignment and directional adaptation are key ingredients of the interactions that produce scale-free velocity correlations, as expected in systems that break the rotational symmetry. However, scale-free behavior in the speed fluctuations and the surprisingly low decay exponent of correlations indicate that flocking is also driven by microscopic mechanisms that are novel, nontrivial, and yet to be explained.

4. FROM DATA TO MODELS

4.1. Analysis of Small Groups

Statistical analysis of empirical data allows characterizing collective patterns and indirectly investigating the nature of interactions between individuals. One can wonder whether a more systematic way exists of exploiting experimental information to develop better models of collective behavior. Some interesting works in which experiments on animal groups were used to test and calibrate reference models recently addressed this issue (69–73).

These works certainly represent an important step forward in the dialectics between theory and experiment. However, they are more similar to a model-fitting approach than to a true inference procedure in the statistical sense. For example, in Reference 70 the authors collected data on groups of surf scoters; they then considered an SPP model with attraction/alignment/repulsion forces and optimized the parameters of the model by minimizing the mean-squared difference between observed and predicted pair correlation function and angular density. In Reference 72, after an analysis of the dependence of individual accelerations on the distances, angles, and speeds of neighbors in groups of 2–3 fish, the authors looked at larger groups (10–30 individuals) using the force matching method (74): A model is postulated for the force acting on the individuals, and the optimal parameters are derived by minimizing the mean-squared difference between the observed force and the one of the model. A very similar analysis is pursued in Reference 71, in which a model is proposed for the dynamics of the individual fishes, and the functions appearing in the model are fit to data using an iterative expectation-maximization algorithm.

In many of the above cases, very small groups of individuals are considered, with the underlying thinking that interactions that regulate coordination between a few individuals remain the same in larger groups. This perspective is explicitly discussed in Reference 73, with the authors using an incremental approach: First, they proposed a model for a single fish and fit it to single-fish experiments, then they looked at a two-fish group and fit the pair-wise interaction parameters, after which they used this calibrated model for predicting behavior for larger groups and compared the modeling to the data, and so on for up to 30 fish.

Although the scalability of the interactions is in fact true in physics, this hypothesis is much less reasonable when dealing with behavior: Common experience tells us that individuals do not behave the same when alone, with only a few companions, and in crowds. For this reason, if one wants to pinpoint the essential mechanisms (i.e., the effective interactions) that determine collective behavior, it is important to focus on the large groups (i.e., hundreds rather than few tens of individuals) in which these interactions are known to be at play.

4.2. The Maximum Entropy Approach

Inferring interactions from large data sets (i.e., the inverse problem) has become a major research field in biophysics. Novel methods based on traditional Bayesian approaches (75) as well as statistical physics (76–86) have been pursued to make sense of the huge amount of data now

available on regulatory processes at the molecular scale. These methods also provide useful schemes for behavioral biology, where much less has been done.

One approach that allows the building of models directly from the data is the maximum entropy (ME) approach (76). The main philosophy of this method is to assume nothing a priori about the system and only use what we know from experiments. For example, consider the case in which we have measured a set of observables, $\{O(x)\}$, that are functions of the microscopic degrees of freedom x of the system. We can then ask, what is the least structured measure $[P(x)]$ that is consistent with the experiments? $P(x)$ can be formally computed by maximizing the entropy (i.e., imposing minimal structure) with the constraint that the expectation values for the observables $\{O\}$ given by P match the experimental ones, i.e.,

$$\langle O(x)\rangle_P = O_{exp}. \quad 3.$$

Of course, the model delivered by the ME method strictly depends on the observables we consider as inputs. The more observables with which the measure must be consistent, the more detailed the model is and the more faithful its overall predictions. In this respect, the spirit of the approach is to proceed in steps, providing at first the minimal amount of experimental information: If the resulting model is not adequate to describe the statistical properties of the system beyond the consistency enforced by construction, then additional information is needed. The aim is to arrive at a model that is both minimal and predictive (more on this in Section 5).

ME has been successfully applied to a variety of problems in which networks of interacting units produce nontrivial collective outcomes, such as neural populations (77, 78) and genetic (83) and protein (84) networks. What is mostly appealing for physicists is that because of its mathematical formulation, ME recasts an inference problem (retrieving interactions from data) in what is essentially a statistical physics problem. To better understand how this occurs and how we can apply ME to investigate collective animal behavior, let us now consider explicitly the case of flocking birds.

In the previous section, we discussed one of the main experimental findings characterizing flocking as a collective phenomenon: the presence of long-range correlations in the velocity fluctuations. A good starting point is therefore to use the velocity correlations $C_{ij} = \mathbf{v}_i \cdot \mathbf{v}_j$ as input observables for the ME approach (i.e., the observables O mentioned above) and compute the maximum entropy distribution $P(\{\mathbf{v}_i\})$ consistent with experimental measurements of such correlations. To simplify the problem even further, let us for the moment disregard the variability of the speed from bird to bird and focus on flight direction only, i.e., let us indicate with $\mathbf{v}_i$ the normalized individual velocity. The ME principle tells us that to find P we must maximize the entropy, enforcing a constraint on the correlations. Mathematically, this constrained maximization can be solved by introducing a generalized entropy

$$\mathcal{S}\left[P, \{J_{ij}\}\right] = S[P] - \sum_{ij} J_{ij}\left[\langle C_{ij}\rangle_P - C_{ij}^{exp}\right], \quad 4.$$

where $S[P] = -\int d_i P(\mathbf{v}_i) \ln P(\mathbf{v}_i)$ is the Shannon entropy related to P (87, 88), and the $\{J_{ij}\}$ are Lagrange parameters enforcing the constraints (89). The ME distribution can then be found by maximizing the entropy of Equation 4 in the space of all possible distributions and with respect to the parameters J_{ij}. In this way, we get

$$P(\{\mathbf{v}_i\}) = \frac{1}{Z\left(\{J_{ij}\}\right)} \exp\left[\frac{1}{2}\sum_{ij=1}^{N} J_{ij}\mathbf{v}_i \cdot \mathbf{v}_j\right], \quad 5.$$

where $Z(\{J_{ij}\})$ is the normalization, and the $\{J_{ij}\}$ are determined by the set of coupled equations

$$\left\langle \mathbf{v}_i \cdot \mathbf{v}_j \right\rangle_P = C_{ij}^{exp} \qquad 6.$$

(85). The ME distribution of Equation 5 is easily recognizable as a Boltzmann distribution, corresponding to a Hamiltonian $\mathcal{H} = -(1/2)\sum_{ij} J_{ij}\mathbf{v}_i \cdot \mathbf{v}_j$. This is a generalized Heisenberg model describing the flock as a set of spin variables (the flight directions), which tend to mutually align with interactions J_{ij}. The set of Equations 5 and 6 allows us to infer these effective interactions, provided that (*a*) we know how to compute the partition function $Z(\{J_{ij}\})$ for every possible value of the J_{ij}s and that (*b*) we have reliable measurements of the correlations. In general these are not trivial problems. On the one hand, computing pair-wise Hamiltonians on generic networks is not easy: In the case of neural populations, for example, where one deals with Ising rather than continuous spins (77), this must be done numerically. On the other hand, good experimental measurements of two-point observables typically require a huge amount of independent experiments.

How can we handle these two issues for flocks? Concerning Z, we can exploit the large polarization of flocks and perform a low-temperature expansion of Equation 5, what is called the spin-wave approximation in magnets (90). We can express the individual flight directions in terms of a longitudinal and a perpendicular component, i.e., $\mathbf{v}_i = v_i^L \mathbf{n} + \boldsymbol{\pi}_i$, where $\mathbf{n}$ is the direction of the motion of the group. Given that ordering is strong, the $\boldsymbol{\pi}_i$s are very small, and we can expand everything in the Hamiltonian, retaining only quadratic terms. The resulting Gaussian theory can be solved explicitly as a function of the $\{J_{ij}\}$ (85).

The question of measurements is more subtle. Not only do we have very few observations for the correlation (corresponding to the number of snapshots available for a single flocking event), but on top of that these observations correspond to different microscopic networks because the birds move (even if not much) and exchange positions. If two birds were close neighbors and strongly interacting at a given time, they might not be so at a later one. Another way of saying this is that both the correlations (C_{ij}) and the interactions (J_{ij}) are not stationary if defined in terms of the identities *i* and *j* of the birds. Because ME looks for a stationary distribution, this seems to be a serious problem. From a mathematical point of view, the inference Equation 6 cannot, in fact, be solved. One way out of this is to get rid of bird identities: We know that interactions and correlations depend on the distances between individuals rather than their identities. So instead of using a J_{ij} interaction matrix, we can parameterize it with a distance-dependent function. The simplest way to do that is to characterize the interaction with a scale and a range. We then consider in Equation 5 a matrix J_{ij}, which is equal to J if at least one of birds *i* and *j* is within the first n_c neighbors of the other, and is zero otherwise. Interestingly, this simplified distribution can be obtained directly from an ME procedure in which, instead of using the full C_{ij}^{exp} as experimental input, we use a much simpler observable, i.e., the average degree of correlation of an individual with its n_c interacting neighbors:

$$C^{int} = \frac{1}{N}\sum_i \frac{1}{n_c}\sum_{j \in n_c^i} \mathbf{v}_i \cdot \mathbf{v}_j. \qquad 7.$$

C^{int} is a scalar, is averaged over all individuals in a flock, and, contrary to the full correlation matrix, can be faithfully measured even in a single sample (snapshot). The ME model consistent with it is mathematically well defined (the distribution of Equation 5 with the parameterized J_{ij}), is very simple, and has only two unknown parameters, the scale J of the interaction and its range, n_c. To infer J, we use the ME equation enforcing $\langle C^{int} \rangle_P = C_{exp}^{int}$ (the analog of Equation 6), whereas to retrieve n_c, we use a maximum likelihood criterion and maximize the log likelihood (ln P) of the data (75) (see **Figure 7*a*,*b***).

Surprisingly, this very simple ME model with only two parameters provides excellent predictions. In **Figure 7*c***, for example, we can see that the correlations of the fluctuations of the flight

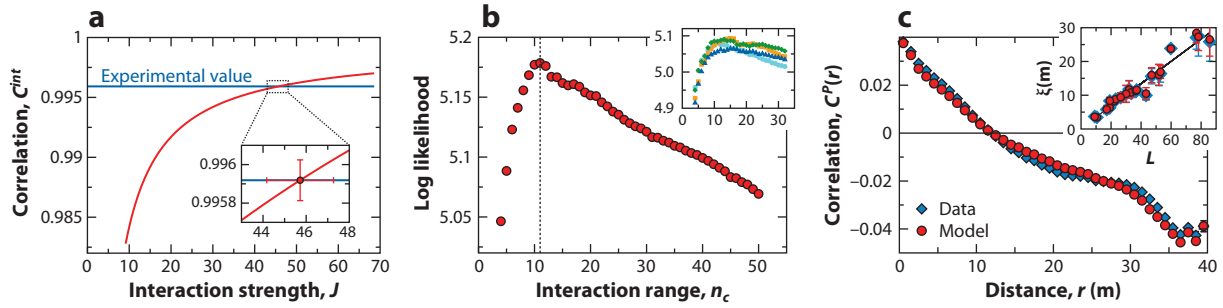

Figure 7

Maximum entropy approach. (*a*) Inference of the parameter J: The predicted value of the correlation C^{int} (*red line*) is equated to the experimental value (*blue line*). (*b*) Inference of the range n_c: The log likelihood of the data $\langle \ln(P) \rangle_{exp}$ at a given instant of time is maximized with respect to n_c. The dashed line indicates the position of the maximum. Inset: examples of the log likelihood at other times. (*c*) Correlation function of the fluctuations of the flight directions as predicted by the model and measured experimentally. Inset: correlation length as a function of the flock size. All figures correspond to the same snapshot of the same flock. Reprinted from Reference 75.

directions are well reproduced over all scales. This was not a priori evident because the input experimental information used (C^{int}) measures correlation only at short scales (a few meters). The same also holds for other multipoint correlations (85). From a statistical physics point of view this result tells us that in a system that is regulated by short-range interactions, the fundamental correlations are between birds/spins and their directly interacting neighbors; all more distant correlations are derivable from these. The inference procedure allows retrieval of the effective strength and range between individuals. Analysis of these parameters in 22 flocking events shows that J and n_c do not depend on size, indicating that there are no long-range effects. Looking at how n_c and its metric alternative definition, r_c, depend on density, one finds a picture analogous to **Figure 3**, confirming the topological nature of the interactions previously obtained with completely different structural investigations. The main conclusion is that flocks can be described in terms of mutual alignment pair-wise interactions, with the interaction network being short range and density invariant.

The application of the ME approach that we have described so far is only a first step. There are several developments that can be considered. For example, one can look at the full velocities (flight directions and speed) and investigate the mechanisms leading to the scale-free correlations in the speed (65). More ambitiously, one can try to take into account the dynamical nature of the network and look for a dynamical ME distribution instead of a static one (86). In all cases, the remarkable feature of this approach is that it provides a mathematically well-defined framework in which a model is built from the data according to a clear principle, entropy maximization. Experimental information is in this sense exploited to sublimation because for any piece of experimental information the minimal model consistent with that information is derived. As a result, if we get a model that works well, it is clear that there are no redundancies and only the effective necessary interactions are present.

5. WHAT IS A GOOD MODEL?

The maximum entropy model we presented in the previous section may seem crude to some bio-oriented readers. Of all the beautiful complexity of a real flock of birds, only a very schematized alignment interaction is described, and even this in a simplified way. Many ingredients have been ignored, such as attraction and repulsion between birds, the role of the environment, the presence of a roost, the aerodynamics of flight, and heterogeneity between individuals.

For the physicist, however, this very simplification is in fact a bonus. The maximum entropy model has just two parameters, and they are fixed by the sharp mathematical requirement to maximize the entropy consistently with the experimental observations. The model is simple and limited in its scope, so we can understand what the model does: If a prediction is experimentally verified, we can understand why this happens. If there is a problem, again we can understand why.

This dichotomy, in various degrees and with obvious exceptions, is common when comparing models developed by physicists with models developed by biologists. As a paradigmatic example, we can consider the Vicsek model of flocking (19, 22) and the very detailed model of flocking found in Reference 15. The first model, with only two parameters, has played a crucial role in understanding what regulates ordering in systems of active individuals. However, it would certainly offer poor quantitative predictions for real flocks of birds. The second model has twenty-four parameters, and numerical simulations produce synthetic flocks remarkably similar to natural ones. However, the space of the parameters is far too large to assess the role of the various ingredients included in the model.

We might say that physicists are not interested in reproducing reality if the price is to not understand what is behind it. That is often the problem of very realistic and detailed models. In short, the minimal models developed by physicists attempt to provide understanding rather than realism.

As far as this discussion goes, however, it seems that the right balance between detail and simplification and between realism and understanding is largely a matter of taste and of scientific background. In fact, it is not so. Within the framework of Bayesian statistics, there is a way to formally understand to what extent the physicist's feeling to "keep it simple" is reasonable. This is what we discuss in the below sections.

5.1. Model Fitting

Imagine we have collected experimental data (D) that we want to use to test a model (H). Here we do not discuss where H is from: It is an arbitrary model (out of common sense, a theory, etc.) that we had in mind before collecting the data. Let us call π the set of parameters of model H. What we need to do is to fit the model to the data, i.e., to find the particular set of parameters π^* that makes the model's prediction the closest to the experimental data D. To do that we can use Bayes formula on conditional probability (79),

$$P(\pi|D,H) = \frac{P(D|\pi,H)P(\pi|H)}{P(D|H)}, \qquad 8.$$

where

$$P(D|H) = \int d\pi\, P(D|\pi,H)P(\pi|H). \qquad 9.$$

The quantity at the left-hand side of Equation 8, $P(\pi|D,H)$, is called posterior probability: It is the probability of the parameters π, given the data D and the model H. This is the quantity we want to maximize with respect to π when we do model fitting: We must find the set of parameters π^* that maximizes the posterior probability. However, the quantity at the right-hand side, $P(D|\pi,H)$, is the probability of obtaining the data D, given the model H and the parameters π. Once expressed as a function of the parameters π, this quantity is known as the likelihood $\mathcal{L}(\pi|D,H) \equiv P(D|\pi,H)$. The other key quantity in the Bayes equation is $P(\pi|H)$, which is called prior probability of the parameters, given the model. This quantity expresses a possible bias of the scientist toward one

particular set of parameters with respect to another prior to receiving any experimental information. The prior probability is the dark side of the Bayesian approach (79). Finally, the denominator in Equation 8 is known as evidence, for reasons that are clearer below. The evidence does not contain π, so it is not crucial during model fitting.

If we have no prior knowledge about the parameters, we can assume that the prior probability is constant, $P(\pi|H) \sim 1$, so that Equation 8 tells us that maximizing the posterior probability is the same as maximizing the likelihood. This is the maximum likelihood method for model fitting in a nutshell. However, in general, the prior probability is not a constant. In the simplest and yet nontrivial case, the prior probability may simply take into account the fact that there may be some unphysical values of the parameters that we want to exclude so that $P(\pi|H) = 0$ in some regions and is constant in other regions of the parameter space. For example, if the parameter v is the velocity of a bird in a flocking model, we certainly have $v > 0$ and $v < v_{\max}$, an arbitrarily large velocity (that of light, for example). In this case, using the normalization condition $\int d\pi\ P(\pi|H) = 1$, we obtain $P(\pi|H) \sim 1/v(\pi)$, where $v(\pi)$ is the volume of the region a priori acceptable in the space of parameters. Imagine we have no data and we formulate model H. Its parameters will have a domain of a priori reasonability. The factor $v(\pi)$ is the volume of this domain.

5.2. Model Comparison

Now imagine that we have two different models, H_1 and H_2, and that we have fit both models to the same set of data D. We want to compare the two models and answer the following question: What model is the best one, given the data? Here, biologists and physicists typically give different answers. The biologist answers that the best model is the one that produces the best fit of the data. The physicist answers that the best model is the one with the lowest number of parameters, provided that the fit is not too poor. The first answer is very clear, whereas the second one seems quite vague. However, the second answer hints at some kind of trade-off, which is absent in the first answer. Let us see how the Bayesian framework answers this question. We want to calculate the probability of a certain model H, given the data D. The Bayes formula at the level of model comparison reads (75)

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}. \qquad 10.$$

Notice that $P(D|H)$ is the evidence defined in Equation 9. To compare two models we need to know which model has the higher probability $P(H|D)$, given the data. If we have no preference for one model with respect to the other (fair, unbiased comparison), then we have equal priors, $P(H_1) = P(H_2)$, and we obtain

$$\frac{P(H_1|D)}{P(H_2|D)} = \frac{P(D|H_1)}{P(D|H_2)}. \qquad 11.$$

We therefore see that the best model is the one with the highest evidence, given the data. This is the origin for why $P(D|H)$ is called "evidence."

Let us now make the very reasonable hypothesis that the likelihood $P(D|\pi, H)$ in the definition of the evidence (see Equation 9) has one single maximum, π^*, in the space of the parameters and that $P(D|\pi, H)$ is different from zero only in a certain portion of the parameter space around π^*, whose volume we call $v(\pi|D)$. Recalling that we can write $P(\pi|H) \sim 1/v(\pi)$, we can then approximate the integral in Equation 9 as

$$P(D|H) \sim P(D|\pi^*, H)\frac{\nu(\pi|D)}{\nu(\pi)} = \mathcal{L}(\pi^*|D, H)\frac{\nu(\pi|D)}{\nu(\pi)}. \quad 12.$$

This is an important formula. The first factor at the right-hand side is the likelihood, so the formula tells us that the model is good if the likelihood is large, and this is basically the same as saying that the model is good if the data fit is good. This is what the biologist suggests: Choose the model with the best data fit, i.e., with the largest likelihood. However, the second factor at the right-hand side also matters. This term $\nu(\pi|D)/\nu(\pi)$ is the ratio between the volume of the reasonable domain of the parameters after having knowledge of the data and the volume of the reasonable domain of the parameters before having knowledge of the data. Clearly, this factor is smaller than one: The experimental data will bring new knowledge, hence restricting the possible domain of the parameters. Let us assume that our model has more than one parameter (this is the safest assumption of all), and let us call k the number of parameters. Then we have

$$\frac{\nu(\pi|D)}{\nu(\pi)} = \prod_{\alpha=1}^{k}\frac{\nu(\pi_\alpha|D)}{\nu(\pi_\alpha)}. \quad 13.$$

Given that for each parameter π_α, the volume reduction factor is smaller than one, we finally understand the physicist's point: The total volume reduction factor in Equations 12 and 13 decreases exponentially with the number of parameters k of the model. Increasing the number of parameters, of course, gives a better fit, i.e., a larger likelihood $\mathcal{L}(\pi^*|D, H)$, but the cost of increasing k is a drastic decrease of the volume reduction factor $\nu(\pi|D)/\nu(\pi)$. If we assume that the volume reduction is the same for all parameters, and we define $r \equiv |\log(\nu(\pi|D)/\nu(\pi))| > 0$, we can write a simplified but very illuminating expression for the evidence of a model,

$$P(D|H) \sim \mathcal{L}(\pi^*|D, H)e^{-kr}. \quad 14.$$

This is the crucial trade-off in model fitting: The improvement in the fit due to adding one extra parameter to our model has to be significant to balance the sharp exponential decrease of the volume reduction factor. This is the Bayesian meaning of the physicist's answer. A highly detailed model with many parameters can give an excellent fit of the data, and yet the probability of this model being correct can be much smaller than the one of a simplified model with fewer parameters providing a very crude fit. The volume reduction factor $\nu(\pi|D)/\nu(\pi) \sim e^{-kr}$ is sometimes called Occam's factor (75). We believe it is a good name: An exponential razor is definitely sharp enough.

5.3. Predictive Power versus Understanding

This dialectic between highly detailed versus ultrasimplified models is a very typical trait of the somewhat complicated relationship between biology and physics. In our opinion, the core of the problem is that there is a sort of trade-off between predictive power and understanding. Let us make an example. Imagine that Harpo (a physicist) is severely ill and needs a life-saving drug. If he does not take the drug, he dies. However, the drug is nasty (as are all powerful drugs), so if he takes too much, he also dies. Harpo has to choose a model to make a simulation of his health condition to decide how much drug he should take. In this case, despite being a physicist, Harpo does not give a damn about understanding, but he is very concerned with prediction. Hence, he does not want the smart simple model that makes it clear why the drug works but rather the most detailed model, with a surplus of parameters. Because what Harpo really wants is to stay alive and to do that he needs a good fit for his health condition. In his particular condition, he does not care about why the drug works; he only wants to make it work right here, right now. A wrong model that gives an

excellent fit is great for Harpo. Basically, Harpo wants a simulation so detailed and realistic as to be just as good as another Harpo, which he can then use as a virtual guinea pig.

However, this trade-off between prediction and understanding only holds up to a point, as in the long run better understanding also brings better predictive power. Understanding needs a small and simplified arena to be developed and consolidated. In such a simplified arena, of course, prediction is typically quite poor. This fact, which is regularly criticized by some biologists, is really obvious: A simplified arena is not established for providing a detailed prediction but rather to develop a clear idea about the essential mechanisms that cause a certain phenomenon. Attacking the crude nature of this simplified arena is as good as criticizing a child's tricycle as an unsuitable means of transportation in city traffic. The tricycle is neither a real bike nor a real toy. It is a precious training tool. Once the understanding developed within the simplified arena is tested and consolidated, one can insert it into a more structured and detailed model whose prediction will be significantly better than the old one. The point is that a wrong model can make a reasonable prediction, or even quite a good one, but it cannot compete with an equally realistic model that is right. However, to reach the understanding necessary to develop the right model, crude and sometimes extreme simplifications are first needed.

6. CONCLUSIONS

To conclude this long marathon on flocking, let's summarize the understanding gathered so far on natural flocks and discuss how it connects to existing research in active matter. When looking at flocks as condensed matter systems, there are two main paradigms that come to mind: the paradigm of liquids and particles interacting with attraction-repulsion potentials and the paradigm of ferromagnetism, in which vectorial degrees of freedom (spins) try to align with neighbors. These two paradigms are in fact beautifully integrated in the modeling approaches that have been developed in the past ten years to treat active matter systems. In self-propelled particle models, particles move in space while obeying dynamical equations that update the individual velocity according to an alignment force with neighbors and to attraction-repulsion forces. Hydrodynamic theories that look at the large-scale, long-time properties of flocking fluids describe the hybridization of these two aspects through mixed modes that couple density and orientational fluctuations, resulting in new propagation laws and correlation exponents.

These two paradigms have also proven extremely helpful when looking at data from real flocks. As briefly illustrated in this review, the concepts and the methodology used to characterize flocks directly come from the analogy with physical systems. Structure in space was investigated using generalizations of the structural observables used in liquids, whereas properties of the velocity fields were addressed using fluctuation correlations and scaling arguments. From this perspective, the ME approach provides a clear and rigorous way to bridge the gap between real experimental data and statistical physics models of the kind we have in mind when we think about flocks. The findings and the understanding that this approach brought us in part confirm our intuitions and are consistent with theoretical results in the active matter literature and in part reveal novel features that demand further investigation.

Flocks of birds can be seen as very sparse liquid-like systems in which particles/birds move and diffuse, there is almost no structure in absolute distances beyond a neat hard-core repulsion, and strong angular anisotropies are present. Mutual alignment with neighbors is certainly a crucial ingredient driving coordination between the individuals and is able to account for some of the statistical properties of the velocity fields.

One of the first and most surprising results coming from data analysis is that the interaction between individuals in a flock is topological, i.e., density invariant. As we discussed, this kind of

interaction grants more robustness in cohesion and is therefore biologically more efficient than a metric one. Topological interactions were not fully addressed in models of flocking prior to these findings. In fact, some models had topological ingredients (for example, see 22), but the role of invariance with respect to density was neither realized nor investigated. Interestingly, the results of Reference 48 triggered a novel interest in this problem. SPP models of particles interacting in a purely topological way were studied numerically (25), revealing how the topological nature of the interaction suppresses the coupling between local density and local order present in standard metric SPP models and making the order-disorder transition continuous rather than discontinuous (see **Figure 8***a*,*b*). Continuous theories were also derived with a kinetic approach (91), confirming numerical results and showing that the structure of the hydrodynamic equations for the ordered phase is the same as that found in Reference 26.

When comparing results from real flocks with the predictions of the hydrodynamic theory (27, 31), one finds some qualitative agreements and several differences. For example, when looking at diffusion of individuals in the flock, one finds that birds move superdiffusively and anisotropically, as predicted in Reference 27, but with a different diffusion exponent. The velocity correlation functions decay with a power law but with an exponent that is significantly smaller than the one predicted in Reference 27. There are several possible origins for these discrepancies. First, hydrodynamic theories have been originally formulated for fluids of active particles, whereas real flocks are finite, with sharply defined borders. In other words, a flock is more like a drop of fluid than a bulk fluid (as first discussed in Reference 22). We have argued that the coming of information through the boundary can explain the small value of the decay exponent of the correlations. Therefore, the most appropriate comparison of experimental results would in fact be with a hydrodynamic theory of active drops. Some recent progress in this respect has been achieved in a Stokesian context in Reference 92, where drops of active gel are studied numerically, and in Reference 93, where a theory for drops of apolar active fluid is presented. Another point to be considered is that in real flocks, rearrangements of individuals are very slow; as a consequence, the scales over which hydrodynamic effects become truly visible might be beyond the ones present in current experiments. Finally, recent results on turns (68) suggest that for birds there are conservation

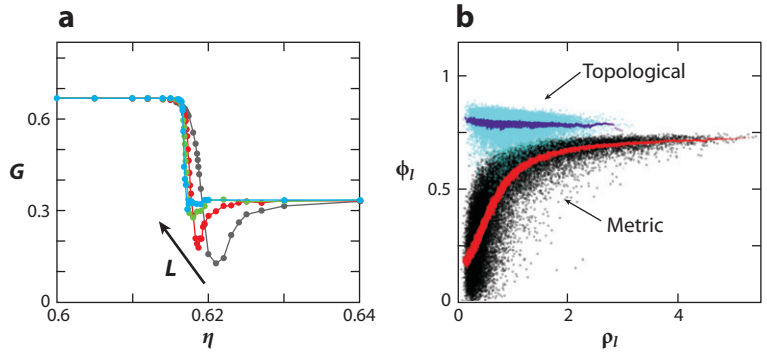

Figure 8

Two-dimensional Vicsek model of self-propelled particles with topological interactions. (*a*) Binder cumulant G versus noise intensity η for different system sizes (*colors*): The dip of the cumulant increases with system size, indicating that the transition is continuous. The arrow indicates the direction of increasing size L. (*b*) Local order parameter ϕ_l (average polarization in a box of size $l = 16$) versus local density ρ_l. In the metric original version of the model, these two quantities are coupled, whereas polarization does not depend on density in the topological case. Reprinted with permission from Reference 25.

laws, which are not considered in existing hydrodynamic theories and are crucial to quantitatively describing transport of information in flocks. Adding these novel ingredients opens new stimulating possibilities for theoretical developments in active matter.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

It is a pleasure to thank all present and past members of the Collective Behaviour in Biological Systems group at ISC-CNR and our collaborators W. Bialek, F. Ginelli, T. Grigera, T. Mora, G. Parisi, and A. Walczak for the research performed together, the many discussions, and the shared enthusiasm in the past years. We thank S. Ramaswamy for several comments and suggestions. This work was supported by grants FP6-NEST12682 STARFLAG, IIT-Seed Artswarm, ERC-StG n. 257126, and AFOSR-Z809101.

LITERATURE CITED

1. Krause J, Ruxton GD. 2002. *Living in Groups*. Oxford: Oxford Univ. Press
2. Giardina I. 2008. *HFSP J.* 2:205–19
3. Sumpter DJT. 2010. *Collective Animal Behavior*. Princeton: Princeton Univ. Press
4. Camazine S, Deneubourg J-L, Franks NR, Sneyd J, Theraulaz G, Bonabeau E. 2001. *Self-Organization in Biological Systems*. Princeton: Princeton University Press
5. Couzin ID, Krause J. 2003. *Adv. Stud. Behav.* 32:1–75
6. Parrish JK, Edelstein-Keshet L. 1999. *Science* 284:99–101
7. Parrish JK, Hamner WH, eds. 1997. *Animal Groups in Three Dimensions*. Cambridge: Cambridge Univ. Press
8. Aoki I. 1982. *Bull. Jpn. Soc. Sci. Fish.* 48:1081–88
9. Huth A, Wissel C. 1992. *J. Theor. Biol.* 156:365–85
10. Reynolds CW. 1987. *Comput. Graph.* 21:25–33
11. Couzin ID, Krause J, James R, Ruxton GD, Franks NR. 2002. *J. Theor. Biol.* 218:1–11
12. Couzin ID, Krause J, Franks NR, Levin SA. 2005. *Nature* 433:513–16
13. Hemelrijk CK, Kunz H. 2005. *Behav. Ecol.* 16:178–87
14. Hildenbrandt H, Carere C, Hemelrijk C. 2010. *Behav. Ecol.* 21:1349–59
15. Hemelrijk CK, Hildenbrandt H. 2011. *PLoS ONE* 6:e22479
16. Vicsek T, Zafeiris A. 2012. *Phys. Rep.* 517:71–140
17. Ramaswamy S. 2010. *Annu. Rev. Condens. Matter Phys.* 1:323–45
18. Marchetti MC, Joanny JF, Ramaswamy S, Liverpool TP, Prost J, et al. 2013. *Rev. Mod. Phys.* 85:1143–89
19. Vicsek T, Czirók A, Ben-Jacob E, Cohen I, Shochet O. 1995. *Phys. Rev. Lett.* 75:1226–29
20. Toner J, Tu Y, Ulm M. 1998. *Phys. Rev. Lett.* 80:4819–22
21. Grégoire G, Chaté H. 2004. *Phys. Rev. Lett.* 92:025702
22. Grégoire G, Chaté H, Tu Y. 2003. *Physica D* 181:157–70
23. D'Orsogna MR, Chuang YL, Bertozzi AL, Chayes LS. 2006. *Phys. Rev. Lett.* 96:104302
24. Aldana A, Dossetti V, Huepe C, Kenkre VM, Larralde H. 2007. *Phys. Rev. Lett.* 98:095702
25. Ginelli F, Chaté H. 2010. *Phys. Rev. Lett.* 105:168103
26. Toner J, Tu Y. 1995. *Phys. Rev. Lett.* 75:4326–29
27. Toner J, Tu Y. 1998. *Phys. Rev. E* 58:4828–58
28. Simha RA, Ramaswamy S. 2002. *Phys. Rev. Lett.* 89:058101
29. Liverpool TB, Marchetti MC. 2003. *Phys. Rev. Lett.* 90:138102

30. Kruse K, Joanny J-F, Jülicher F, Prost J, Sekimoto K. 2004. *Phys. Rev. Lett.* 92:078101
31. Toner J, Tu Y, Ramaswamy S. 2005. *Ann. Phys.* 318:170–244
32. Baskaran A, Marchetti MC. 2008. *Phys. Rev. E* 77:011920
33. Bertin E, Droz M, Gregoire G. 2006. *Phys. Rev. E* 74:022101
34. Ihle T. 2011. *Phys. Rev. E* 83:030901
35. Kudrolli A, Lumay G, Volfson D, Tsimring LS. 2008. *Phys. Rev. Lett.* 100:058001
36. Deseigne J, Dauchot O, Chaté H. 2010. *Phys. Rev. Lett.* 105:098001
37. Narayan V, Menon N, Ramaswamy S. 2007. *Science* 317:105–8
38. Sokolov A, Aranson IS, Goldstein RE, Kessler JO. 2007. *Phys. Rev. Lett.* 98:158102
39. Makris NC, Ratilal P, Symonds DT, Jagannathan S, Lee S, Nero RW. 2006. *Science* 311:660–63
40. Luthi B, Tsinober A, Kinzelbach W. 2005. *J. Fluid Mech.* 528:87–118
41. Ouellette NT, Xu H, Bodenschatz E. 2006. *Exp. Fluids* 40:301–13
42. Cullen JM, Shaw E, Baldwin HA. 1965. *Anim. Behav.* 13:534–43
43. Partridge BL, Pitcher TJ, Cullen JM, Wilson J. 1980. *Behav. Ecol. Sociobiol.* 6:277–88
44. Davis MJ. 1980. *Anim. Behav.* 28:668–73
45. Pomeroy H, Heppner F. 1992. *Auk* 109:256–67
46. Major PF, Dill LM. 1978. *Behav. Ecol. Sociobiol.* 4:111–22
47. Hartley R, Zisserman A. 2003. *Multiple View Geometry in Computer Vision*. Cambridge: Cambridge University Press
48. Ballerini M, Cabibbo N, Candelier R, Cavagna A, Cisbani E, et al. 2008. *Proc. Natl. Acad. Sci. USA* 105:1232–37
49. Ballerini M, Cabibbo N, Candelier R, Cavagna A, Cisbani E, et al. 2008. *Anim. Behav.* 76:201–15
50. Cavagna A, Giardina I, Orlandi A, Parisi G, Procaccini A, et al. 2008. *Anim. Behav.* 76:217–36
51. Cavagna A, Giardina I, Orlandi A, Parisi G, Procaccini A. 2008. *Anim. Behav.* 76:237–48
52. Cavagna A, Cimarelli A, Giardina I, Parisi G, Santagati R, et al. 2010. *Proc. Natl. Acad. Sci. USA* 107:11865–70
53. Attanasi A, Cavagna A, Del Castello L, Giardina I, Jelic A, et al. 2013. arXiv:1305.1495 [q-bio.QM]
54. Cavagna A, Cimarelli A, Giardina I, Orlandi A, Parisi G, et al. 2008. *Math. Biosci.* 214:32–37
55. Cavagna A, Cimarelli A, Giardina I, Parisi G, Santagati R, et al. 2010. *Math. Models Methods Appl. Sci.* 20:1491–1510
56. Camperi M, Cavagna A, Giardina I, Parisi G, Silvestri E. 2012. *Interface Focus* 2:715–25
57. Caraco T, Martindale S, Pulliam HR. 1980. *Nature* 285:400–1
58. Cresswell W. 1994. *Anim. Behav.* 47:433–42
59. Procaccini A, Orlandi A, Cavagna A, Giardina I, Zoratto F, et al. 2011. *Anim. Behav.* 82:759–65
60. Emmerton JD. 1993. In *Vision, Brain, and Behavior in Birds*, ed. HP Zeigler, HJ Bischof. Cambridge, MA: MIT Press
61. Nieder A. 2005. *Nat. Rev. Neurosci.* 6:177–90
62. Inada Y, Kawachi KJ. 2002. *J. Theor. Biol.* 214:371–87
63. Young GF, Scardovi L, Cavagna A, Giardina I, Leonard NE. 2013. *PLoS Comp. Biol.* 9:e1002894
64. Parisi G. 1998. *Statistical Field Theory*. Boulder, CO: Westview Press
65. Bialek W, Cavagna A, Giardina I, Mora T, Pohl O, et al. 2013. arXiv:1307.5563 [physics.bio-ph]
66. Ginelli F. 2013. *Phys. Rev. Lett.* 110:168107
67. Cavagna A, Duarte Queiros SM, Giardina I, Stefanini F, Viale M. 2013. *Proc. Biol. Sci.* 280:20122484
68. Attanasi A, Cavagna A, Del Castello L, Giardina I, Grigera T, et al. 2013. arXiv:1303.7097 [cond-mat.stat-mech]
69. Buhl J, Sumpter DJT, Couzin ID, Hale JJ, Despland E, et al. 2006. *Science* 312:1402–6
70. Lukeman R, Li YX, Edelstein-Keshet L. 2010. *Proc. Natl. Acad. Sci. USA* 107:12576–80
71. Herbert-Read JE, Perna A, Mann RP, Schaerf TM, Sumpter DJT, Ward AJW. 2011. *Proc. Natl. Acad. Sci. USA* 108:18726–31
72. Katz Y, Tunstrom K, Ioannou CC, Huepe C, Couzin ID. 2011. *Proc. Natl. Acad. Sci. USA* 108:18720–25
73. Gautrais J, Ginelli F, Fournier R, Blanco S, Soria M, et al. 2012. *PLoS Comp. Biol.* 8:e1002678
74. Eriksson A, Nilsson Jacobi M, Nyström J, Tunstrøm K. 2010. *Behav. Ecol.* 21:1106–11

75. MacKay DJC. 2003. *Information Theory, Inference and Learning Algorithms*. Cambridge: Cambridge Univ. Press
76. Jaynes ET. 1957. *Phys. Rev.* 106:620–30
77. Schneidman E, Berry MJ II, Segev R, Bialek W. 2006. *Nature* 440:1007–12
78. Tang A, Jackson D, Hobbs J, Chen W, Prieto A, et al. 2008. *J. Neurosci.* 28:505–18
79. Yu S, Huang D, Singer W, Nikolić D. 2008. *Cereb. Cortex* 18:2891–901
80. Halabi N, Rivoire O, Leibler S, Ranganathan R. 2009. *Cell* 138:774–86
81. Cocco S, Leibler S, Monasson R. 2009. *Proc. Natl. Acad. Sci. USA* 106:14058–62
82. Weigt M, White RA, Szurmant H, Hoch JA, Hwa T. 2009. *Proc. Natl. Acad. Sci. USA* 106:67–72
83. Lezon TR, Banavar JR, Cieplak M, Maritan A, Federoff NV. 2006. *Proc. Natl. Acad. Sci. USA* 103:19033–38
84. Mora T, Walczak AM, Bialek W, Callan CG. 2010. *Proc. Natl. Acad. Sci. USA* 107:5405–10
85. Bialek W, Cavagna A, Giardina I, Mora T, Silvestri E, et al. 2012. *Proc. Natl. Acad. Sci. USA* 109:4786–91
86. Cavagna A, Giardina I, Ginelli I, Mora T, Piovani D, et al. 2013. arXiv: 1310.3810 [cond-mat.stat-mech]
87. Shannon CE. 1949. *Bell Sys. Tech. J.* 27:379–423, 623–656. Reprinted in *The Mathematical Theory of Communication*, ed. CE Shannon, W Weaver. Urbana, IL: Univ. Ill. Press
88. Cover TM, Thomas JA. 1991. *Elements of Information Theory*. New York: Wiley
89. Bender CM, Orszag SA. 1978. *Advanced Mathematical Methods for Scientists and Engineers*. New York: McGraw–Hill
90. Dyson FJ. 1956. *Phys. Rev.* 102:1217–30
91. Peshkov A, Ngo S, Bertin E, Chateé H, Ginelli F. 2012. *Phys. Rev. Lett.* 109:098101
92. Tjhung E, Marenduzzo D, Cates ME. 2012. *Proc. Natl. Acad. Sci. USA* 109:12381–86
93. Joanny JF, Ramaswamy S. 2012. *J. Fluid Mech.* 705:46–57

Annual Review of Condensed Matter Physics

Volume 5, 2014

Contents

Whatever Happened to Solid State Physics?
John J. Hopfield 1

Noncentrosymmetric Superconductors
Sungkit Yip 15

Challenges and Opportunities for Applications of Unconventional Superconductors
Alex Gurevich 35

Correlated Quantum Phenomena in the Strong Spin-Orbit Regime
William Witczak-Krempa, Gang Chen, Yong Baek Kim, and Leon Balents 57

Dirac Fermions in Solids: From High-T_c Cuprates and Graphene to Topological Insulators and Weyl Semimetals
Oskar Vafek and Ashvin Vishwanath 83

A Quantum Critical Point Lying Beneath the Superconducting Dome in Iron Pnictides
T. Shibauchi, A. Carrington, and Y. Matsuda 113

Hypercomplex Liquid Crystals
Zvonimir Dogic, Prerna Sharma, and Mark J. Zakhary 137

Exciton Condensation in Bilayer Quantum Hall Systems
J.P. Eisenstein 159

Bird Flocks as Condensed Matter
Andrea Cavagna and Irene Giardina 183

Crossover from Bardeen-Cooper-Schrieffer to Bose-Einstein Condensation and the Unitary Fermi Gas
Mohit Randeria and Edward Taylor 209

Crackling Noise in Disordered Materials
Ekhard K.H. Salje and Karin A. Dahmen 233

Growing Length Scales and Their Relation to Timescales in Glass-Forming Liquids
Smarajit Karmakar, Chandan Dasgupta, and Srikanth Sastry 255

Multicarrier Interactions in Semiconductor Nanocrystals in Relation to the Phenomena of Auger Recombination and Carrier Multiplication
Victor I. Klimov 285

Polycrystal Plasticity: Comparison Between Grain-Scale Observations of Deformation and Simulations
Reeju Pokharel, Jonathan Lind, Anand K. Kanjarla, Ricardo A. Lebensohn, Shiu Fai Li, Peter Kenesei, Robert M. Suter, and Anthony D. Rollett 317

Molecular Beam Epitaxy of Ultra-High-Quality AlGaAs/GaAs Heterostructures: Enabling Physics in Low-Dimensional Electronic Systems
Michael J. Manfra 347

Simulations of Dislocation Structure and Response
Richard LeSar 375

Errata

An online log of corrections to *Annual Review of Condensed Matter Physics* articles may be found at http://www.annualreviews.org/errata/conmatphys