Empirical questions for collective-behaviour modelling

NICHOLAS T OUELLETTE

Department of Mechanical Engineering and Materials Science, Yale University, New Haven,
CT 06520, USA
E-mail: nicholas.ouellette@yale.edu

Abstract. The collective behaviour of groups of social animals has been an active topic of study across many disciplines, and has a long history of modelling. Classical models have been successful in capturing the large-scale patterns formed by animal aggregations, but fare less well in accounting for details, particularly for groups that do not display net motion. Inspired by recent measurements of swarming insects, which are not well described by the classical modelling paradigm, I pose a set of questions that must be answered by any collective-behaviour model. By explicitly stating the choices made in response to each of these questions, models can be more easily categorized and compared, and their expected range of validity can be clarified.

1. Introduction

One of the most remarkable sights in nature is the coordinated motion of huge groups of social animals. From flocks of birds to schools of fish to herds of ungulates, this collective motion is widespread across many different species, and examples of collective behaviour can be found at every size scale in biology. Acting together is thought to be able to grant a range of potential advantages for both individuals and populations [1], such as enhancing the efficiency of mating [2] or foraging [3], more effectively avoiding predators [4], or navigating complex and changeable terrain [5].

Collective behaviour has therefore attracted the attention of scientists from many disciplines, from biology to physics to engineering. Each of these communities wishes to understand collective behaviour for a different reason – to determine its evolutionary underpinnings, for example, or to exploit the benefits of collectivity for artificially engineered systems. In all cases, however, the development of accurate models is of central importance, both as a check on our understanding and as a starting point for adapting collective behaviour to man-made systems.

It is not surprising, then, that a wide range of models have been developed to capture collective motion. Models have been successful at reproducing qualitative features of animal aggregations, such as the high degree of ordering in bird flocks [6] or the range of group morphologies observed in different systems [7]. Simply reproducing the shape or ordering of an aggregation, however, is not proof that the model actually captures the true biology [1], or even the most salient features of the real system. To make progress, models must be benchmarked against data from real animal groups. Such data have historically been rare, but the past decade has seen tremendous progress in the acquisition of quantitative and detailed empirical results for a range of animal systems. Some of these datasets – particularly those for which the animals mostly move with velocities that are highly ordered in space – have largely validated assumptions made by common models [8–11]. But models have been less successful with aggregations that do not show net motion, such as insect swarms [12].

Motivated both by modelling successes and particularly failures, the aim of this paper is to give a brief review of the current modelling landscape in the context of recent experiments in the laboratory on mating swarms of flying insects. The dynamics of these swarms are not captured well by the current social-force-based models [12], and recent results have suggested that one must consider the biological goals of the individual animals more carefully to generate more accurate models [13]. These considerations lead to the formulation of a set of biologically motivated questions, some of which have been largely overlooked, whose answers define any collective-behaviour model. By clearly specifying the choices made in response to these questions, models can be more easily compared, and their range of expected validity more clearly specified.

Section 2 contains a brief summary of the experiments with mating swarms of *Chironomus riparius* midges and our key results. Section 3 describes the current approaches to modelling collective behaviour, and poses a set of defining questions for models. Finally, §4 has a brief discussion of situations when biologically accurate modelling is needed and suggestions for future research.

2. Empirical measurements of insect swarms

Unlike highly ordered animal aggregations such as bird flocks or fish schools, mating swarms of flying insects present a particular challenge for modelling as there is no simple average state of the swarm that can be used to benchmark a model. Swarms do not display net motion, in that their centre of mass is fixed; but unlike other stationary aggregations such as toroidal mills of fish, insect swarms also show no ordered rotational motion inside the group. Instead, each individual insect moves erratically and seemingly at random [14], so much so that it has been suggested that swarms are not truly examples of collective behaviour [2]. Swarms, however, show a high degree of spatial cohesion, even if the motion of the individuals is complex, and thus are likely to be a collective system [15], albeit one that is different from flocks and schools. Understanding the ways in which swarming insects are similar to and different from more highly ordered groups can therefore shed light on the aspects of collective-behaviour models that are likely to be universal, and which may be more species-specific.

2.1 *Experimental measurements*

We have recently undertaken a laboratory study of swarms of the non-biting midge *Chironomus riparius*. Swarming is an integral part of their life cycle, and these midges will swarm and breed under laboratory conditions [16–19]. Adult midges swarm as part of their mating ritual [20]. At dawn and dusk, males congregate in groups of anywhere from 10 to 10^4, typically near stream edges. Females fly through the swarm, where they are met in midair by males. Coupled pairs may drop to the ground for copulation [14] or may mate on the wing [2]. The females subsequently deposit fertilized egg masses in nearby water; males return to the swarm to mate again.

A self-sustaining breeding colony of *C. riparius* is maintained in our laboratory. The midges spend their entire life cycle in a cubical enclosure measuring 91 cm on a side. This enclosure contains nine independent development tanks, each containing aerated, dechlorinated fresh water and a cellulose substrate into which the larvae can burrow. The enclosure is illuminated from above with a diurnal light source, providing 16 h of light and 8 h of darkness per day. When this light turns on and off (corresponding to 'dawn' and 'dusk'), male adult midges will spontaneously form mating swarms. It has been observed that swarms in our enclosure range from a few individuals up to about 100 [21,22]. Females occasionally enter the swarms and mate, after which they deposit fertilized egg masses in the development tanks. The mating swarms can be positioned in the enclosure by means of a 'swarm marker' that helps to nucleate the swarm [22]. In the wild, swarms can form over any visual feature; in our experiments, a piece of black felt is used.

To make quantitative measurements of the midges, a fully automated, three-dimensional optical particle tracking algorithm is used. We illuminate the midge swarms with near-infrared light; the midges cannot see this light, and so it does not affect their behaviour. We then film the swarms at a rate of 100 pictures/s with three Point Grey Flea3 cameras arranged around the enclosure. By calibrating the camera system using a standard method [23], the redundant information recorded by the cameras can be used to locate the midges in three-dimensional space. We link the instantaneous positions into trajectories using a predictive multiframe algorithm [24]. Velocities and accelerations can then be computed by differentiating these trajectories [21,25].

2.2 *Summary of results*

Detailed descriptions of our measurements of the swarm dynamics observed are recorded in other works [12,13,21,22]; here, they are described only briefly. Aside from very small swarms, with fewer than 10 individuals [22], swarms with robust, repeatable statistical properties are observed. The number density of the swarms is approximately independent of the total number of participating individuals; thus, the swarm radius grows approximately as the 1/3 power of the number of individuals [21]. Inside the swarms, the midges do not arrange themselves randomly [21]; rather, the distribution of midges is closer to that expected for a gas of soft spheres [12]. The overall swarm has no net linear or angular momentum [21], and neither acceleration nor velocity statistics show clear evidence for force-like interactions between individuals [12]. Indeed, at the average level, the midges are qualitatively similar to particles in an ideal gas, albeit a rarified one that is too dilute to be described well as a continuum system [12], with a nearly Maxwell–Boltzmann

distribution of velocities, an exponential distribution of free paths, and statistically ballistic motion [12,21]. Nevertheless, there is evidence, both indirect and direct, of interactions. The midges are tightly bound to the swarm, and behave as if the swarm creates an effective harmonic trap with a stiffness that decreases with size [21]. More subtly, time–frequency analysis of midge trajectories reveals pairwise interactions (not necessarily with nearest neighbours) characterized by high-frequency periodic oscillations in the relative distance between pairs [13]. These results together pose a challenge for current models: individuals in the swarm do show aspects of collective behaviour and interaction, but not of a form that is simple to describe. In particular, the implications of collective dynamics are difficult to discern in the average behaviour of the swarm, although they are clearly present.

3. Modelling collective behaviour

High-resolution, in-depth empirical measurements such as those described previously have become available for animal groups only recently, and experimental data are still limited. For much of its history of study, modelers of collective animal behaviour have therefore been forced to settle for qualitative information such as the overall shape of an aggregation or to make assumptions about the underlying interactions between individuals. Current models have been successful at qualitatively reproducing various aggregation morphologies observed in nature such as polarized flocking or toroidal milling [1,7]. Simple models can also quantitatively account for the average behaviour of animal groups that primarily exhibit unidirectional motion [8–11]. The same models are largely unsuccessful, however, in capturing details or fluctuations away from the mean-field behaviour of the group [26], or in describing aggregations such as insect swarms that do not show net motion [12].

With the ever-increasing availability of empirical data, the time is now ripe for reevaluating many of the assumptions commonly made by models in light of some of these failures. After first briefly reviewing the ideas that underpin models of animal behaviour, a set of questions are presented that must be answered to define a model. By clearly and specifically addressing the questions posed here, different models can be more straightforwardly compared – particularly in terms of their expected range of applicability.

3.1 *Modelling paradigms*

Although the literature on modelling collective animal motion is vast, most proposed models can be sorted into three categories: continuum, agent-based, or rule-based discrete models [1,27–29]. In analogy with the two common formulations of continuum mechanics, continuum models are sometimes known as Eulerian models, while agent-based models are sometimes called Lagrangian models.

In a continuum model, the individual animals that make up the aggregation are abstracted away, and only a continuum population density field ρ is explicitly considered. In this way, the system can be modelled by partial differential equations (PDEs), allowing the application of powerful mathematical tools [30]. Models in one, two, and three dimensions have all been studied, although in more than one dimension analytical progress is

very challenging. Most models have a similar basic form. The evolution of ρ is given by an advection–diffusion equation of the form

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\mathbf{v}(\rho)\rho) = \nabla \cdot (D(\rho)\nabla \rho),\tag{1}$$

where $\mathbf{v}(\rho)$ is a (continuum) velocity field and $D(\rho)$ is a diffusion coefficient [27,31]. Both $\mathbf{v}$ and D may or may not depend on the density, and the functional form of this dependence models the interactions between the individuals that make up the aggregation. If $\mathbf{v}(\rho)$ is given by a spatial convolution of the density field with some appropriate kernel, non-local interactions can be modelled [27,28,32,33], at the cost of turning the PDE into an integrodifferential equation. $\mathbf{v}(\rho)$ may also contain multiple terms to capture competing tendencies toward aggregation or repulsion between individuals. Continuum models have been successful in reproducing the qualitative patterns observed in real animal groups [34], and are valuable tools for studying the stability of groups with different assumed interaction potentials. The continuum assumption, however, is very strong, and may be unjustified for real animal groups, particularly those that are small or dilute [12]. Additionally, a continuum model necessarily treats all individual animals as identical, making them unsuitable for capturing heterogeneities, either in space or in time, in the population.

In contrast to the continuum case, agent-based models consist of explicit equations of motion for each individual in the aggregation. Instead of a single PDE system, then, an agent-based model is a set of coupled ordinary differential equations (ODEs). This approach has strengths and weaknesses when compared with a continuum model. Agent-based models can be very flexible: each kind of interaction (such as attraction to or repulsion from other individuals or coupling to the background environment) can be recorded explicitly for each individual organism. Multiple types of individuals, such as different species [35], genders [2], or 'leaders' that possess special information [36,37], are also simple to incorporate, as are stochastic effects that model the imperfect response of individuals to their environment. Agent-based models are also more appropriate for small or dilute aggregations. But, as the number of equations is typically large, and the equations are coupled and may contain nonlinear and/or stochastic terms, analytical progress is tremendously difficult [33], and agent-based models must be characterized almost entirely by numerical simulation.

Rule-based models, such as those based on cellular automata, are similar to agent-based models in that they consider individual animals separately. They typically, however, are not based on differential equations, instead specifying only interaction rules with no explicit dynamics [1].

In this study, the focus is on agent-based models. Such models are, in a sense, the most fundamental way to describe an animal aggregation. A continuum model is necessarily an abstraction of what is in reality a discrete system; and, using formal methods of coarse-graining from statistical mechanics, a continuum model can always be derived from an agent-based model [30,38–40]. A simple rule-based model can likewise be considered to be a discretization of an underlying agent-based model. In addition, in an agent-based model the full dynamics of each individual can be separately described, complete with population-level heterogeneity or non-stationarity. Agent-based models are therefore more flexible and can capture more types of behaviours than a coarse-grained model.

Thus, the focus here is on the basic questions that must be answered to define an agent-based model.

3.2 *Fundamental questions for models*

The space of possible agent-based models is tremendously large, because an agent-based model is basically a set of coupled ODEs. It is not surprising, then, that a great number of models have been proposed to capture collective behaviour, by scientists working in a range of fields. The exact form of these models can be quite different, and is often implicitly predicated on the kind of assumptions typical in different disciplines: models proposed by physicists, for example, often treat the agents as indistinguishable particles and focus instead on the interactions, as is common in statistical mechanics, while those devised by biologists tend to include more heterogeneity in the population. Due to such fundamental differences in approach, models originating in different communities can be challenging to compare, because they are so different from each other.

In this study, the aim is to focus on the problem of building agent-based models of collective behaviour by formulating a set of specific, biologically relevant questions. Any model of a collective animal system must answer the questions, and the choices that are made serve to define the model.

3.2.1 *How do individuals behave without interactions?*
The general consensus in the community is that the macroscopic collective behaviour of an animal group arises spontaneously from the behaviour of the individuals and their interactions, in much the same way as the thermodynamic state of a material emerges from the motion and coupling of its constituent atoms or molecules. If the behaviour of the animals is dominated by their interactions, as might be expected in very dense groups, then the way they behave in the absence of interactions is likely to be unimportant. Many models therefore, devote little attention to the non-interacting dynamics of the agents; for example, it is common to assume that in the absence of interactions each individual moves in a straight line at a constant speed [6,7].

Real animals, however, do not move in such a simple fashion – and if the aggregation is dilute or weakly coupled, as is the case in, e.g. insect swarms [12], the independent dynamics of the individuals may become relevant. And even when animals are interacting, they may respond to others by modulating their speed rather than simply their direction, as has been observed, e.g., in schooling fish [10] or flocking birds [41]. Such effects can be incorporated into an agent-based model by including an explicit equation of motion for the velocity of each individual in addition to the position [42]; these velocity equations may also then be coupled between different individuals. Although these additions add to the complexity of the model, they may sometimes be necessary for accurately describing real animals.

3.2.2 *Which individuals interact?*
We expect that interactions are the key driver of the emergence of collective behaviour in an animal group. Thus, the core of any collective behaviour model will be the nature of these interactions. But in addition to defining exactly how to represent these interactions, it is necessary to choose which individuals interact. The most agnostic answer to this question would be to assume that each

individual may interact with any or all of the others in the group. This situation, however, is unlikely for large groups, where individuals may be separated by large distances, or for dense groups, where the presence of distant animals may be screened by local neighbours. It is therefore extremely common to restrict the range of interactions to be with those other individuals that are nearby, effectively replacing the interaction network in the aggregation by the proximity or communication network [43,44].

But, even if we make this potentially restrictive choice, we must decide how to define the interaction range. Two choices are commonly made: one can use metric distance, so that a given animal interacts with all neighbours within a prescribed physical region, or topological distance, where an animal interacts with a fixed number of others, regardless of the distance [45]. Models based on metric distance implicitly assume that more distant neighbours are less important to a given animal. Such an assumption is likely to be reasonable for organisms such as bacteria with only limited sensing ability, because in that case information transfer is highly localized. But for more intelligent animals, a topologically defined local neighbourhood may be more reasonable. Models with metric interactions are less stable to density fluctuations than those with topological interactions [46,47], because transient local expansion can abruptly move two nearby individuals outside their interaction range in a metric model. Given that large density fluctuations are common in many active systems [48], defining interaction partners via a topological distance will help ensure more realistic cohesion of the aggregation.

Interaction models based on either metric or topological distance, however, make the assumption that a given individual in the aggregation responds more strongly to nearer neighbours than to others that are more distant. This choice is reasonable for animals where vision is the primary sensing mechanism (because vision is strongly screened in dense groups) and where the primary goal is simply not to collide with neighbours. But in many other situations, this assumption may fail. For example, many animals use additional senses to locate other individuals or to navigate their environment, including hearing, smell, or electromagnetic sensing [49]. Information of these types is often available over longer ranges, and may lead to more distant coupling. In midge swarms, sound, which has a longer effective range in the swarms than vision, plays a key role in the interaction dynamics [50,51] – and evidence shows that the dominant interindividual interactions are not necessarily between nearest neighbours, but rather occur at longer distances [13]. A similar neglect of very local interactions in favour of longer-distance 'interactions' is found in humans moving toward goals in crowded environments [52]. And the simple assumption that the local neighbourhood dominates the interactions of an individual also neglects any social hierarchy in the group, even though pre-existing social structure may strongly determine which individuals preferentially interact [43,53].

3.2.3 *What is the form of the interactions?*   After deciding which individuals interact with each other, we must then posit a form for these interactions; namely, we must decide on the rules that govern the collective behaviour. Many choices for such rules have been suggested, but the most common choice is to balance three effects [6,7,45]: an attraction to distant individuals, to help keep the aggregation cohesive; a repulsion from nearby neighbours, to avoid a collapse of the aggregation to a point; and a tendency to align one's direction of motion with intermediate-range neighbours, to promote a net polarization of the aggregation. When combined with self-propulsion, such a model can produce

aggregations that qualitatively behave like real animal groups, and that can show macro-scopic behaviour ranging from flocking to milling to swarming [7]. Such a model is also appealing because these tendencies can be interpreted as social 'forces,' making their representation by ODEs straightforward.

This kind of model has been successfully fit to animal groups that display strongly ordered motion, including marching locusts [8], floating ducks [9], and schooling fish [10,11]. But in systems like our midge swarms that are stationary, effective social forces have not been clearly identified [12]. Instead, interactions in the swarms are more subtle, and are difficult to identify from the kinematics of single individuals alone [13]. And even for animal groups for which effective-force models are good descriptors, the actual biology is quite different.

It is likely that, in many cases, attraction to, repulsion from, and particularly align-ment with neighbours are emergent epiphenomena rather than true low-level interactions. Nevertheless, these behavioural rules may be useful modelling tools, as long as it is understood what kinds of more fundamental interactions lead to their emergence – and what kinds do not. To accomplish this task, models should be formulated that explicitly account for the biological goals of each individual, such as moving to a target, exploring space efficiently, or minimizing energy expenditures, rather than being built at an intermediate level.

3.2.4 *Are the interaction rules constant throughout the population?*   The simplest agent-based model of a collective system would consist of a set of identical ODEs, one for each individual in the population. If the number of agents is very large and the system is in the 'thermodynamic' limit, then this ansatz is reasonable, because when averages are taken over a large system, individual variation is expected to be smoothed away. This is therefore the approach taken by most models that originate from statistical physics [6,29,45,48]. The question of course remains as to how large is 'very large' for an animal aggregation; for insect swarms, at least, recent results suggest that it may be surprisingly small (only about 10 individuals) [22].

Treating the agents as identical particles, however, is inappropriate if the differences between individuals are not stochastic, but instead have important biological conse-quences. If a population of animals has an intrinsic hierarchy of social dominance [53] or if individuals have different levels of confidence [37], for example, the rules governing interaction between individuals may be highly non-uniform in the population and would depend on the identity of the individuals interacting [43]. In aggregations of foraging or migrating animals, some individuals may have distinct information as to the goal of the population, and can act as leaders [36]. Some animal groups may contain individuals of more than one species, which can lead to heterogeneities in composition or behaviour within the aggregation [35]. In all these cases, a model that assumed that all the agents were identical up to statistical fluctuations would likely perform poorly.

3.2.5 *Are the interaction rules constant in time?*   Most agent-based collective-behaviour models assume that the rules for any given individual are static in time, even if they may vary between different individuals in the population. This assumption is again reasonable in the context of statistical mechanics, or for animal groups where every individual is pursuing simple, static goals. But in many cases, enforcing static interaction

rules is biologically questionable. In the mating swarms of insects we have studied, for example, the default behaviour of a male midge is to explore the swarm volume searching for females. When the male finds a female, however, its behaviour qualitatively changes from exploration to chasing – and its effective interaction rules change as well, as it focusses only on the female and not on other males. The behaviour of a male even changes qualitatively when it interacts with another male to determine its gender [13]. An accurate model for this situation cannot, therefore, use fixed interaction rules, but rather must include the potential for switching, in either a stochastic or a deterministic way, between different behavioural modes. More broadly, accurate models for many situations may require multistate agents, potentially with memory, with additional criteria for switching between states.

4. Discussion and conclusions

Ultimately, the questions posed in this study are empirical, and can be answered by careful (if sometimes difficult) data acquisition and processing. They can also potentially be used as a guide to model development: if the answers are known to some but not all of them, then possible answers to the remaining questions can be systematically checked by comparing the model output to empirical data.

A deeper question, however, still remains: how much does it matter that a model is completely correct? Or, to phrase it differently, is it a problem if a model correctly predicts the mean-field behaviour even with unrealistic assumptions? The answer to this question is somewhat philosophical, and depends on the reason the model is being developed. From the standpoint of a physicist or an applied mathematician who wishes to understand the minimal requirements for the emergence of macroscopic structure in interacting non-equilibrium systems, a fully accurate description of the underlying biology is likely to be unnecessary [45]. Indeed, physics is replete with examples of useful, and even predictive, 'toy' models of complex systems. But in other cases, realistic models are potentially important. For example, if one wants to argue that a model fully explains an actual biological system, then it must be rooted in the correct biology, rather than assumptions at the epiphenomological level. Realism is also probably important when using the ideas of collective behaviour to design artificial, engineered systems. It is reasonable to expect that collective behaviour has evolved to be robust against even large perturbations, because animals routinely operate in noisy, fluctuating, uncontrolled environments. Simple models, however, are not guaranteed to posses this level of stability [54], and robustness to all kinds of perturbations can be difficult to test, particularly for large agent-based models.

The questions posed in this study suggest two promising lines of future research. In a wide range of very different systems and potentially for very different reasons, quite similar self-organized collective states emerge. This observation suggests that collective behaviour may be insensitive to some of the questions that have been posed. Identifying which of these questions are the most influential for animal groups should thus be studied, and the result will have implications for potential universality in describing collective animal behaviour. At the same time, new methods need to be developed for quantitatively validating models against empirical data. As many different underlying assumptions can generate similar emergent states [1], model validation requires testing the assumptions

that go into a model rather than its output. Developing better methods to accomplish this task will be tremendously useful in deepening our understanding of how collective behaviour emerges.

Acknowledgements

References

[1] J K Parrish and L Edelstein-Keshet, *Science* **284**, 99 (1999)
[2] J A Downes, *Annu. Rev. Entomol.* **14**, 271 (1969)
[3] D E Jackson, M Holcombe and F L W Ratnieks, *Nature* **432**, 907 (2004)
[4] C C Ioannou, V Guttal and I D Couzin, *Science* **337**, 1212 (2012)
[5] N J Mlot, C A Tovey and D L Hu, *Proc. Natl. Acad. Sci. USA* **108**, 7669 (2011)
[6] T Vicsek, A Czirók, E Ben-Jacob, I Cohen and O Shochet, *Phys. Rev. Lett.* **75**, 1226 (1995)
[7] I D Couzin, J Krause, R James, G D Ruxton and N R Franks, *J. Theor. Biol.* **218**, 1 (2002)
[8] J Buhl, D J T Sumpter, I D Couzin, J J Hale, E Despland, E R Miller and S J Simpson, *Science* **312**, 1402 (2006)
[9] R Lukeman, Y-X Li and L Edelstein-Keshet, *Proc. Natl. Acad. Sci. USA* **107**, 12576 (2010)
[10] Y Katz, K Tunstrøm, C C Ioannou, C Huepe and I D Couzin, *Proc. Natl. Acad. Sci. USA* **108**, 18720 (2011)
[11] J E Herbert-Read, A Perna, R P Mann, T M Schaerf, D J T Sumpter and A J W Ward, *Proc. Natl. Acad. Sci. USA* **108**, 18726 (2011)
[12] J G Puckett, D H Kelley and N T Ouellette, *Sci. Rep.* **4**, 4766 (2014)
[13] J G Puckett, R Ni and N T Ouellette (2014) (submitted)
[14] A Okubo and H C Chiang, *Res. Popul. Ecol.* **16**, 1 (1974)
[15] A Attanasi *et al*, *PLoS Comput. Biol.* **10**, e1003697 (2014)
[16] P F Credland, *Freshwater Biol.* **3**, 45 (1973)
[17] A E R Downe and V G Caspary, *Can. Entomol.* **105**, 165 (1973)
[18] A E R Downe, *Can. Entomol.* **105**, 291 (1973)
[19] C P McCahon and D Pascoe, *Chemosphere* **12**, 2471 (1988)
[20] R M Neems, J Lazarus and A J Mclachlan, *Behav. Ecol.* **3**, 285 (1991)
[21] D H Kelley and N T Ouellette, *Sci. Rep.* **3**, 1073 (2013)
[22] J G Puckett and N T Ouellette, *J. R. Soc. Interface* **11**, 20140710 (2014)
[23] R Y Tsai, *IEEE J. Robotic. Autom.* **RA-3**, 323 (1987)
[24] N T Ouellette, H Xu and E Bodenschatz, *Exp. Fluids* **40**, 301 (2006)
[25] N Mordant, A M Crawford and E Bodenschatz, *Physica D* **193**, 245 (2004)
[26] A Cavagna *et al*, arXiv:1403.1202 (2014)
[27] C M Topaz, A L Bertozzi and M A Lewis, *Bull. Math. Biol.* **68**, 1601 (2006)
[28] R Eftimie, G de Vries, M A Lewis and F Lutscher, *Bull. Math. Biol.* **69**, 1537 (2007)
[29] I Giardina, *HFSP J.* **2**, 205 (2008)
[30] G Flierl, D Grünbaum, S Levin and D Olson, *J. Theor. Biol.* **196**, 397 (1999)
[31] A Okubo, *Adv. Biophys.* **22**, 1 (1986)
[32] A Mogilner and L Edelstein-Keshet, *J. Math. Biol.* **38**, 534 (1999)
[33] C M Topaz and A L Bertozzi, *SIAM J. Appl. Math.* **65**, 152 (2004)
[34] R Eftimie, G de Vries and M A Lewis, *Proc. Natl. Acad. Sci. USA* **104**, 6974 (2007)

[35] J W Jolles, A J King, A Manica and A Thornton, *Anim. Behav.* **85**, 743 (2013)
[36] I D Couzin, J Krause, N R Franks and S A Levin, *Nature* **433**, 513 (2005)
[37] A J King, *Behav. Proc.* **84**, 671 (2010)
[38] J Toner and Y Tu, *Phys. Rev. Lett.* **75**, 4326 (1995)
[39] J Toner and Y Tu, *Phys. Rev. E* **58**, 4828 (1998)
[40] J Toner, Y Tu and S Ramaswamy, *Ann. Phys.* **318**, 170 (2005)
[41] W Bialek, A Cavagna, I Giardina, T Mora, O Pohl, E Silvestri, M Viale and A M Walczak, *Proc. Natl. Acad. Sci. USA* (2014)
[42] M R D'Orsogna, Y L Chuang, A L Bertozzi and L S Chayes, *Phys. Rev. Lett.* **96**, 104302 (2006)
[43] N W F Bode, A J Wood and D W Franks, *Anim. Behav.* **82**, 29 (2011)
[44] N W F Bode, A J Wood and D W Franks, *Behav. Ecol. Sociobiol.* **65**, 117 (2011)
[45] T Vicsek and A Zafeiris, *Phys. Rep.* **517**, 71 (2012)
[46] M Ballerini, N Cabibbo, R Candelier, A Cavagna, E Cisbani, I Giardina, V Lecomte, A Orlandi, G Parisi, A Procaccini, M Viale and V Zdravkovic, *Proc. Natl. Acad. Sci. USA* **105**, 1232 (2008)
[47] A Peshkov, S Ngo, E Bertin, H Chaté and F Ginelli, *Phys. Rev. Lett.* **109**, 098101 (2012)
[48] S Ramaswamy, *Annu. Rev. Condens. Matter Phys.* **1**, 323 (2010)
[49] V Guttal and I D Couzin, *Proc. Natl. Acad. Sci. USA* **107**, 16172 (2010)
[50] M V Federova and R D Zhantiev, *Entomol. Rev.* **89**, 896 (2009)
[51] L J Cator, B J Arthur, L C Harringtona and R R Hoy, *Science* **323**, 1077 (2009)
[52] S J Guy, S Curtis, M C Lin and D Manocha, *Phys. Rev. E* **85**, 016110 (2012)
[53] M Nagy, Z Ákos, D Biro and T Vicsek, *Nature* **464**, 890 (2010)
[54] N Khurana and N T Ouellette, *New J. Phys.* **15**, 095015 (2013)