J. theor. Biol. (1999) **196,** 397–454
Article No. jtbi.1998.0842, available online at http://www.idealibrary.com on IDEAL

From Individuals to Aggregations: the Interplay between Behavior and Physics

G. Flierl*†, D. Grünbaum‡, S. Levin§ and D. Olson¶

* *Department of Earth, Atmospheric and Planetary Sciences, M.I.T., Cambridge, MA* 02139, *U.S.A.*, ‡ *Department of Zoology, University of Washington, Box* 351800, *Seattle, WA* 98195, *U.S.A.* § *Ecology and Evolutionary Biology, Princeton University, Princeton, NJ* 08544-1003, *U.S.A. and* ¶ *Rosenstiel School of Marine and Atmospheric Science, University of Miami*, 4600 *Rickenbacker Cswy, Miami, FL* 33149, *U.S.A.*

(*Received on* 20 *May* 1997, *Accepted in revised form on* 18 *September* 1998)

This paper analyses the processes by which organisms form groups and how social forces interact with environmental variability and transport. For aquatic organisms, the latter is especially important—will sheared or turbulent flows disrupt organism groups? To analyse such problems, we use individual-based models to study the environmental and social forces leading to grouping. The models are then embedded in turbulent flow fields to gain an understanding of the interplay between the forces acting on the individuals and the transport induced by the fluid motion. Instead of disruption of groups, we find that flows often enhance grouping by increasing the encounter rate among groups and thereby promoting merger into larger groups; the effect breaks down for strong flows.

We discuss the transformation of individual-based models into continuum models for the density of organisms. A number of subtle difficulties arise in this process; however, we find that a direct comparison between the individual model and the continuum model is quite favorable. Finally, we examine the dynamics of group statistics and give an example of building an equation for the spatial and temporal variations of the group-size distribution from individual-based simulations.

These studies lay the groundwork for incorporating the effects of grouping into models of the large scale distributions of organisms as well as for examining the evolutionary consequences of group formation.

© 1999 Academic Press

1. Introduction

The problem of animal aggregation is a central one in both ecological and evolutionary theory. Why do animals group, what proximate cues do they use, and what are the ecological consequences? Patterns are manifest and observable only at the level of the population, but are mediated at the level of the individual, the focal point for selection. Our thesis is that it is essential to explore the linkages between these levels, and in particular to understand how changes in individuals' responses to their environments translate into changes in observable patterns.

† Author to whom all correspondence should be addressed.

0022–5193/99/004397 + 58 $30.00/0
© 1999 Academic Press

Because individual-based modelling provides a powerful methodology for addressing these issues, it will be one of the principal techniques employed in this paper. In Section 2, we will introduce aggregation models within a setting that allows consideration both of physical and of biological forces. Primary attention will be to aggregation in homogeneous environments, illustrating the patterns of patchiness that can arise autonomously. In reality, however, such self-organization occurs within environments that also impose patterns, at least at some scales. For marine organisms, subject to the ever-changing fluid environment, we must understand the degree to which patterns are imposed by the fluid dynamics, the biology, or some combination of both. In Section 3, therefore, we will introduce explicit consideration of turbulent flow regimes and describe the interaction between self-organization and fluid transport/mixing across scales. Individual-based models allow us to relate forces at multiple scales and to examine the signatures of different behaviors.

While much insight can be drawn from simulations based on following limited numbers in a population, it is not possible to treat large populations in this manner. Instead some level of approximation must be made to reduce the problem to a set of field equations, e.g. an Eulerian description, where the variables represent statistics over the configuration. It remains to translate from individualistic rules to macroscopic descriptions of collective phenomena and to grid-based or continuum descriptions of fluxes of individuals through real space and through aspect space (Levin & Segel, 1985). It is the transformation between these two views of the problem that holds the greatest potential for understanding. In Sections 4 and 5, we demonstrate how individual-based, Lagrangian models lead to macroscopic equations for statistical properties such as the density (4) or group size (5). Such equations, which we call "Eulerian models" since they deal with these properties at fixed spatial or aspect space points, are much more amenable to analysis. We discuss several examples of the dynamics of the population density $\rho(\mathbf{x}, t)$ in Section 4, focusing on the spontaneous development of patches and the shape of an isolated group. Section 5 explores how individual-based rules govern the formation and maintenance of groups of different sizes and how fission and fusion rates may be determined from individual rules. These give dynamical equations for the group-size distribution, which may be solved for equilibrium characteristics.

1.1. BACKGROUND

Aggregation can result from density-independent responses to local cues or from the density-dependent responses of individuals to other individuals, which we shall call social behavior. In practice, both mechanisms are likely to be at work, but at different spatial scales. In many cases, for example the broad-scale distributions of Antarctic krill, the patterns may be best understood in terms of the passive or active responses of individuals to local physical or chemical conditions; yet, in other situations, as in the fine-scale distributions of the same species, individuals are responding to the positions or movements of their neighbors. Data analysis, especially the examination of correlations in observed patterns, confirms this view. Spectral analysis of patterns in the distribution of temperature, fluorescence and krill in the Southern Ocean suggests that krill distributions, as well as those of phytoplankton, are well explained by the fluid dynamics on broader scales. At least for krill, however, another (density-dependent) explanation is needed at the finer scales (Levin *et al*., 1989). This is consistent with data such as those of Mackas (1984), whose transects display increasing patchiness as one moves from temperature to phytoplankton to zooplankton. The argument that different forces dominate at different scales may also apply to the distributions of insects, birds, and terrestrial ungulates. For example, herding animals exist in aggregations determined at large scales by regimes that are defined in terms of soil type or soil moisture content, at mesoscales by vegetation patterns, and at yet smaller scales by responses to each other. However, patterns at large scale are not independent of small-scale structure, and vice versa: the nonlinearity inherent in social behavior couples the scales together.

While the spatial structure of the physical environment and temporal scales associated with its variability are important to the biology of populations, the response of the population must be approached as a combined bio-physical problem. Not only do individuals respond passively to environmental forces, but they also respond actively both to physical and biological cues, with the latter often dominating, at least at small scales. The cues organisms use in making decisions may vary a great deal, and the size of the neighborhood censused can span a wide range. At one extreme, such as monkey troops, each individual in a group may recognize all the other individual members of the group and be able to assess the total group size; in such cases, group size itself may become an important variable, closely linked to individual fitness. At the opposite extreme, as in the aggregations of marine invertebrates or terrestrial ungulates, groups are large and much less clearly delineated, changing continuously in size and composition. In such cases, local densities become more directly relevant than the total group size as cues for individual decisions; groups are largely epiphenomena, with the action focused at much smaller scales, and group size distributions typically will show much higher variance.

Modelling provides an approach to exploring the ways in which the dynamics of individuals lead to groups, the interactions of groups, and the effects of environmental variability and transport. The manner in which groups develop and move will be dependent on whether individual motion involves apparently random orientation, as in swarming, or orientation of individuals relative to each other, as in schooling. In this paper, we use individual-based models to explore some of the macroscopic consequences that emerge from differences in individual rules.

Individual-based models have a long history; indeed, some of the most successful of them have been produced as animations, or cartoons, for broad general audiences (Reynolds, 1987; Parrish & Hamner, 1997). To a large extent, the animator proceeds as would the scientist: different mechanisms are hypothesized and compared with observations; models are then modified to include patterns of movement that seem to produce more realistic simulations. Alternative models are considered, and accepted or rejected on the basis of how well observed phenomena can be imitated. The scientific approach has very different objectives: the goal is not just to produce realistic simulations but to understand and explain the local mechanisms that control patterns. As such, the wealth of behavioral data available can be incorporated. Secondly, not only visualizations, but also appropriate spatial and temporal statistics, need to be applied to evaluate critically the success of particular models under wide ranges of conditions. Finally, and perhaps most importantly, analytical methods can be used to translate the hypothesized local mechanisms into macroscopic dynamics.

The starting point for individual-based models is a specification of the state of an individual in terms of location, genotype and phenotype, ontogenetic status (age, size, maturity), physiological status (nutrient status, hunger, hormonal state), or behavioral status (dominance, motivation). Environmental information must be superimposed on this, ranging from physical and chemical features to the status of neighbors, who can affect access to resources as well as vulnerability to natural enemies and potential for mating. From this information, one determines responses such as growth, mortality, reproduction, and movement (cf. Diekmann *et al*., 1990). The result of this formulation is a description of the changes in states for all members of a population.

As the time and space scales increase, one naturally turns to models of statistical properties, preferably derived from the individual-based model itself and the results of simulations. As an example of important problems that can be addressed once a solid understanding of the grouping dynamics is obtained, consider the interrelationships between behavior and selection. In general, there is a trade-off between the fitness of an individual, conditioned on whether or not it joins a group of a particular size, and the fitness of the individuals in the group, as modified by the potential joiner's action. When the size of the group is small and individual members recognize each other, the decisions can

be controlled to a large extent by those already belonging to the group. For groups with many more individuals, however, strangers are less easily distinguished, and the effects of a new member would in any case be diminished. In this situation, control shifts to the individual contemplating joining a group. This kind of large, transient, loosely-knit group appears predominant in marine situations (with the conspicuous exception of marine mammals).

In general, an individual's ability to assess its environment and its ability to act successfully on that information is key, both from an ecological and from an evolutionary perspective. Information can be gathered in isolation by assessing environmental conditions or, indirectly, by observing the decisions of neighbors. Some combination of these undoubtedly would often be best from the viewpoint of the group, allowing group foraging for resources or detection of predator threats; from the viewpoint of an individual, however, the cost–benefit calculations may appear quite different, leading to fascinating evolutionary puzzles. To what extent do the observed behaviors reflect a collective and cooperative optimum, and to what extent do the individual's selfish interests force a suboptimal situation for the group, in terms for example of resource acquisition? Furthermore, how does decision-making under uncertainty evolve? Individuals can deal with uncertainty by averaging, spatially or temporally; by sharing information; or by expending resources on gaining more detailed information. Do the costs of the latter justify the information gain? The framework for analyzing collective behavior and its dependence upon individual decisions developed in this paper provides the foundation for approaching such questions by evaluating evolutionarily stable strategies within an individual-based context. Applications to evolutionary problems will be addressed in future work.

2. The Role of Biology in Inducing Pattern

We shall be discussing the dual problems of describing social animal behaviors in quantitatively rigorous ways and of generating higher level descriptions for movements and dynamics of social populations. In this section, we begin with a review of some individual-based models of social aggregation. The individual-based aggregation models described here will be used in later sections, which discuss two approaches to deriving population-level descriptions.

2.1. DESCRIBING ANIMAL MOVEMENTS AT THE LEVEL OF THE INDIVIDUAL

We begin by describing social behavior at the level of the individual animal, i.e. following animals in the Lagrangian framework in terms of their position, velocity, internal states, etc. (see Grünbaum & Okubo, 1994, for a review). The acceleration of an individual may include both deterministic $\mathbf{a}_i$ and stochastic $\delta\mathbf{V}_i$ components. Examples of deterministic forces are:

(1) locomotory forces, such as those resulting in preferred swimming speeds;
(2) social forces, such as attraction or repulsion between individuals;
(3) arrayal forces, equalizing speeds and directions of neighboring animals; and
(4) environmental effects, such as chemical gradients, which could lead to directional biases in movement.

Stochastic accelerations, $\delta\mathbf{V}_i$ may come from both behavioral and environmental variability. This component of the acceleration may include some of the effects discussed in the context of deterministic forces; e.g. the grouping term may evoke a range of responses, or accelerations, with some probability distribution. The resulting second order system for the position $\mathbf{x}_i$ and velocity $\mathbf{v}_i$ of the i-th individual can be written

$$\delta\mathbf{x}_i = \mathbf{v}_i\delta t + \delta\mathbf{X}_i \tag{1a}$$

$$\delta\mathbf{v}_i = \mathbf{a}_i\delta t + \delta\mathbf{V}_i \tag{1b}$$

where we have kept stochasticity in the position equation for our later discussions. We shall use non-dimensional equations with the characteristic length scale and velocity scale being those associated with the behavior, e.g. the perception distance and preferred swimming speed. Some marine organisms, such as copepods, have two distinct movement patterns—(1) slow cruising, with fairly small accelerations and (2) sudden "hops" (for example as escape responses) in which the velocities increase transiently by a

large factor. One might represent the former variability in the $\delta\mathbf{V}$ term and the latter in the $\delta\mathbf{X}$ term. The papers discussed in the rest of this section do not include the "hopping" behavior and have $\delta\mathbf{X}_i = 0$; for a discussion of "great leap" models, see Mollison (1977), Mollison & Levin (1995). For N individuals, we have an order $2N$ system of stochastic differential equations, with some parts of the forces describing the outcomes of many-body social interactions. As in the finite step model, one can devise lattice approaches (e.g. Turchin, 1989), in which the velocities take on a discrete set of values, and convert the integral operators used below into matrices.

Lagrangian models must contain specific assumptions about several poorly-understood aspects of social behavior. Differences in these assumptions, too subtle to be easily resolved by current empirical knowledge, may nonetheless have profound effects on the dynamics of the resulting aggregation models. An example is the modulation of responses to neighbors by separation distance and relative orientation, and limits (which may vary with environmental conditions and the degree of local crowding) to the distances at which individuals are able to detect and respond to neighbors. A limit to direct interactions establishes a characteristic aggregation length scale that may have a fundamental role in the fragmentation and coalescence of aggregations. For example, groups much larger than this length may be easily fragmented, which would then strongly influence the population-level distribution of individuals into groups.

Another crucial, but little-understood, aspect of social behavior is the way in which aggregating animals balance the influences of multiple neighbors. In many Lagrangian aggregation models, reactions to neighbors are independently computed and summed. However, this is likely to be an oversimplification of social behavior in many species. Several studies have focused on different influence-sharing schemes (Aoki, 1982; Huth & Wissel, 1990, 1992; Warburton & Lazarus, 1991). For example, Huth & Wissel compared a "decision" algorithm, in which only a single neighbor influences movement at any one time, to an "averaging" method, in which an individual orients relative to the arithmetic mean of properties of the neighbors.

2.1.1. *Example*

To make the discussions that follow more concrete, we shall consider a simple example in which the only variability is from random forces ($\delta\mathbf{X} = 0$). The acceleration is chosen to be a relaxation back towards some preferred velocity $\mathbf{V}$ which depends on the organism's environment (chemical gradients, closeness and direction of neighbors, etc.):

$$\mathbf{a} = \alpha(\mathbf{V} - \mathbf{v}) \tag{2}$$

and the stochastic variability is given by

$$\delta\mathbf{V} = \beta\delta\mathbf{e} \tag{3}$$

where $\delta\mathbf{e}$ is the standard Gaussian increment with zero mean and $\langle\delta\mathbf{e}_i\delta\mathbf{e}_j\rangle = \delta_{ij}\delta t$. The relaxation rate α (with units T^{-1}) and the strength of the random accelerations β (with units $L\ T^{-3/2}$) will generally be functions of the environmental conditions.

We can now distinguish several cases (although an organism may have a behavior which is some combination): the response may be non-directional ($\mathbf{V} = 0$) or directional ($\mathbf{V} \neq 0$) and it may be social (α, β, $\mathbf{V}$ depend on positions or velocities of other organisms) or involve taxis (parameters depend on a field that is independent of the distribution of the organisms, e.g. chemical gradients).

2.2. AUTONOMOUS BEHAVIOR: SIMPLE DIFFUSION

In the case when there is no preferred velocity $\mathbf{V} = 0$ and the random accelerations are constant ($\beta = \text{const.}$), the organisms disperse via a random walk (or, more strictly, random flight) (Fig. 1). The effective diffusivity, shown below to be

$$\kappa = \frac{\beta^2}{2\alpha^2}, \tag{4}$$

can be used quite accurately to predict the mean square displacement of the organisms from their initial positions. The value of κ (the ratio of the dimensional diffusivity to the product of the characteristic sensing distance and the preferred movement speed) is 0.045. The pattern, not surprisingly, is the familiar one associated with diffusion: mean square displacement increases

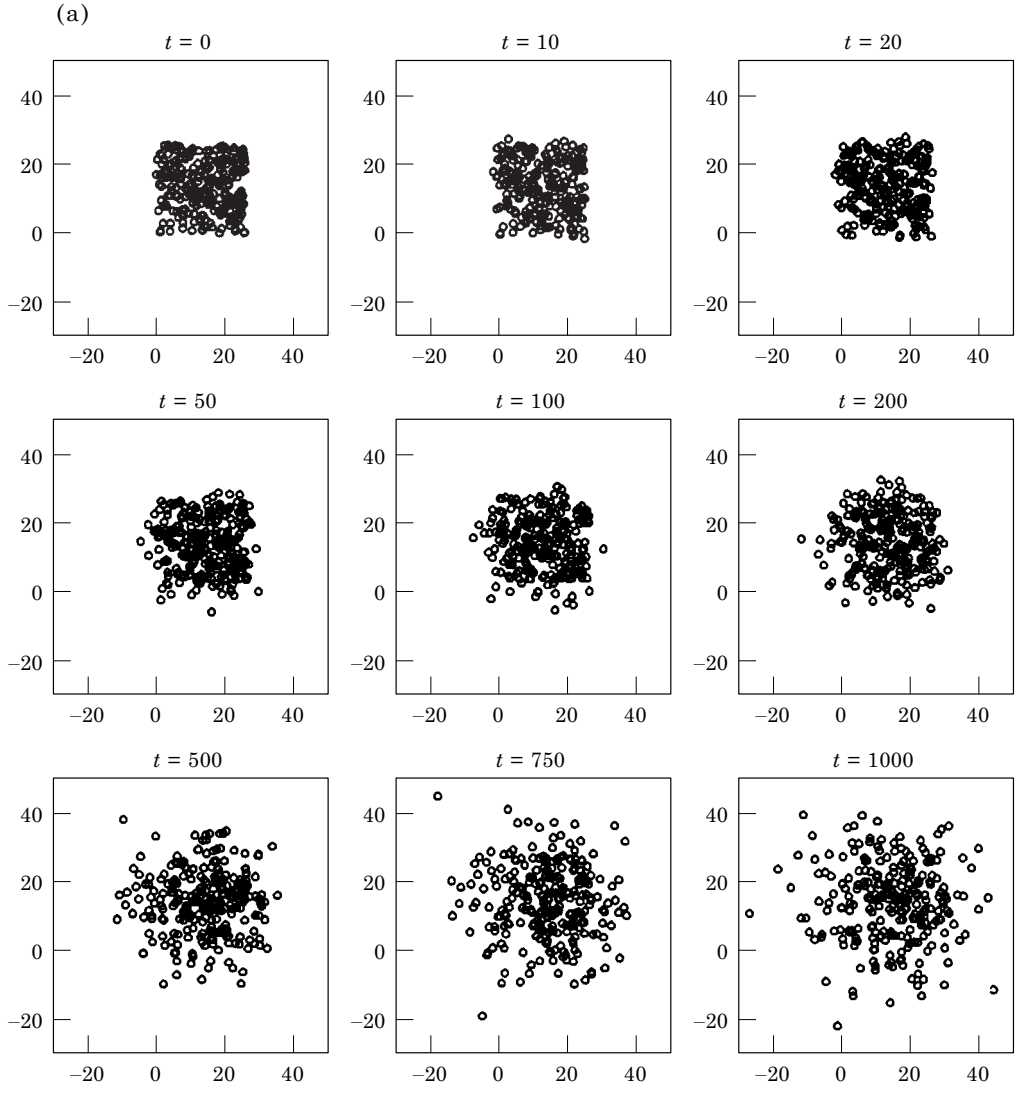


FIG. 1(a)—(*Caption on opposite page*).

linearly with time at a rate (in two dimensions) of 4κ.

2.3. DIRECTIONAL RESPONSE TO ENVIRONMENTAL CUES: TAXIS

The simplest example beyond pure diffusion is directional taxis, in which the preferred velocity depends upon the gradient of a fixed environmental cue. Gradient responses of this kind, like hill-climbing algorithms in combinatorial optimization, are the most direct way organisms can explore a resource spectrum. Figure 2 shows the movement for a sinusoidally distributed resource so that $\mathbf{V} = \cos(2\pi x/10)\mathbf{x}$ with parameters $\alpha = 1.0$ and $\beta = 1.5$ in a doubly-periodic domain of length 10. If we compute the numbers in unit width bins in x or in y, we can see the clustering around the peaks of the environmental signal. The figure also shows the results from a continuum model described later (Section 4)—the two compare very well for these parameters. In either case, individuals climb the gradient successfully, achieving a distribution centered around the peaks.

2.4. NON-DIRECTIONAL RESPONSE TO ENVIRONMENTAL CUES: KINESIS

Responding to a gradient may be the most efficient way to explore a resource spectrum in the absence of information sharing between individuals, but it involves a higher-order calculation that may be difficult to make,

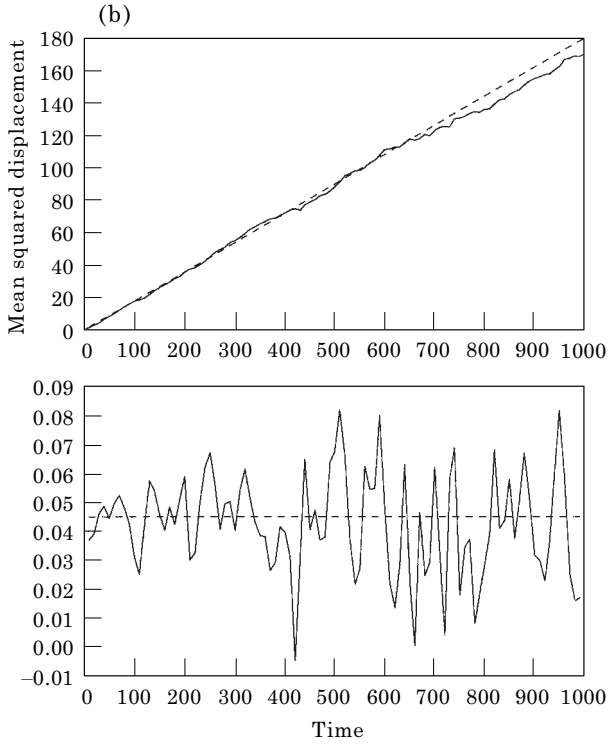


FIG. 1(b)

FIG. 1. (a) Random flight with $\alpha = 1$, $\beta = 0.3$, $\delta t = 0.125$. Particle positions at various times. (b) Mean squared displacement from the particles' starting positions: (– – –) shows the predicted spread for a diffusion model with $\kappa = \beta^2/2\alpha^2 = 0.045$; lower plot shows diffusivity estimated by the rate of change of the mean squared displacement.

especially if the scale of variation is very large compared with the size of an individual (although bacteria, for example overcome this problem by adaptive probing: moving in a random direction to gain, by sensing temporal derivatives, information about gradients, Berg & Brown, 1972). A simpler, if less efficient scheme, kinesis, requires only information about the local properties. Organisms, rather than detecting the gradient in the cue, respond simply by increasing or decreasing their random acceleration as the local concentration of the chemical decreases or increases (cf. Davis *et al.*, 1991 and also the discussion on "repulsive" and "attractive" random walks in Okubo, 1986). In our example, Fig. 3, the parameters are $\mathbf{V} = 0$, $\alpha = 1$, $\beta = [1 - 0.8 \sin(2\pi x/10)] \times 1.5$. Again, the evolution shows clustering in the region of the highest environmental signal. The continuum model reproduces the general shape of the distribution, but is sharper; as we shall see, the continuum models assume that the mean free path is small compared with the scale of the gradient, and this assumption is not sufficiently accurate for this example.

2.5. DIRECTIONAL SOCIAL BEHAVIOR

Relying solely on information gathered by solitary individuals provides direct access to fine-scale characteristics, but limited data about the broad-scale. One way to gain larger scale information is by some form of social behavior. Social behavior involves cues associated with the characteristics of neighboring organisms: their closeness, their velocities, their orientation, etc.

2.5.1. *Preferred velocity function of neighbors' positions*

In the simplest example, animals try to move towards higher concentrations of their neighbors. We might define the preferred velocity, $\mathbf{V}_i$, as

$$\mathbf{V}_i = \Sigma \mathbf{W}_1(\mathbf{x}_j - \mathbf{x}_i) \tag{5}$$

with

$$\mathbf{W}_1(\mathbf{z}) = \begin{cases} \mathscr{V}\mathbf{z}(1 - \mathbf{z}\cdot\mathbf{z}) & |\mathbf{z}| < 1 \\ 0 & |\mathbf{z}| > 1 \end{cases}. \tag{6}$$

Figure 4 shows examples of the evolution, indicating the formation of groups when the grouping tendency is strong enough compared with the random accelerations.

To quantify the degree to which distributions such as those in Fig. 4 are grouped, we can calculate a form of Lloyd's (1967) patchiness index—the ratio of the number of neighbors in the sensing domain compared with the number expected in a Poisson distribution with the same organism density ρ:

$$p = \left[\frac{1}{N}\sum_i \sum_j \mathscr{H}(1 - |\mathbf{x}_j - \mathbf{x}_i|)\right] \Bigg/ \left[\pi\rho + 1 - \frac{\pi}{\text{Area}}\right] \tag{7}$$

where $\mathscr{H}$ is the Heaviside step function and Area is the domain area. Figure 5 shows the patchiness for a case with the social forces turned off

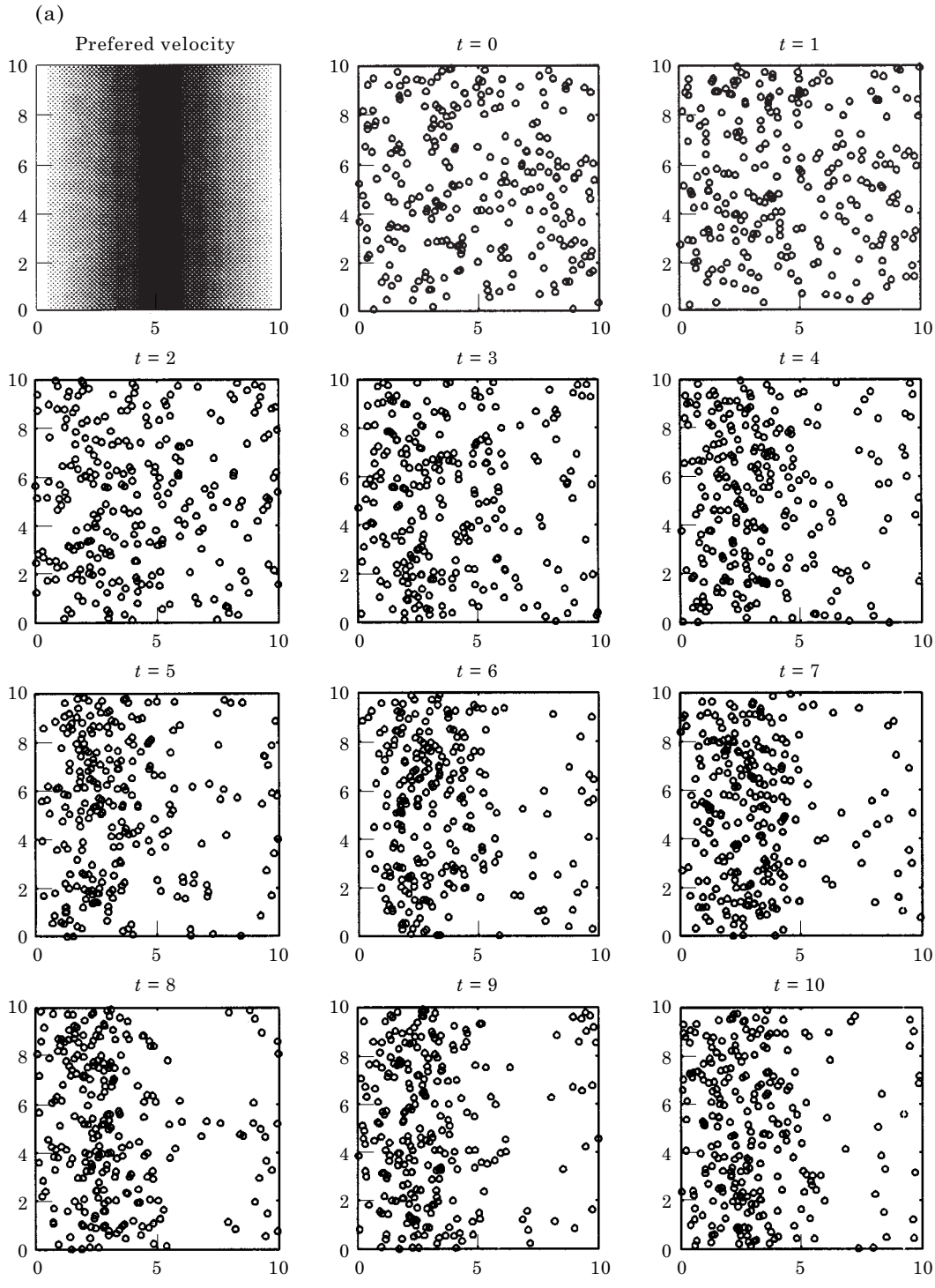


FIG. 2(a)—(*Caption on opposite page*).

($\mathscr{V} = 0$) and for the case above. The value of p reaches more than 10, indicating significant patchiness. Another measure can be obtained by identifying individual groups by collecting all individuals that are within one sensing range ($|\mathbf{x}_j - \mathbf{x}_i| < 1$) of another member of the group. We can use these counts to determine the number of groups (or, inversely, the average number in a group) or the percentage of the population in groups of particular sizes. These

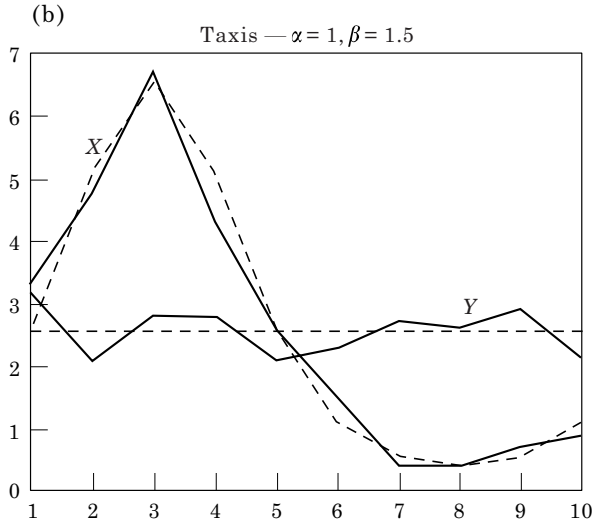


FIG. 2(b)

FIG. 2. (a) Taxis with a sinusoidal distribution of resource. The preferred velocity is proportional to the gradient $[\cos(2\pi x/10)\mathbf{x}]$ as shown in the first frame. The random motion is fairly strong ($\kappa = 1.125$). (b) Final density of organisms, averaged in x or y: (– – –) show the prediction from a continuum model (Section 4).

measures again show that the distribution is significantly non-random.

For the weighting function $\mathbf{W}_1$, the groups are very compact; if we modify the weighting function to be repulsive at short separations,

$$\mathbf{W}_2(\mathbf{z}) = \begin{cases} 3{,}3\mathscr{V}\mathbf{z}(1 - \mathbf{z}\cdot\mathbf{z})(\mathbf{z}\cdot\mathbf{z} - \frac{1}{4}) & |\mathbf{z}| < 1 \\ 0 & |\mathbf{z}| > 1 \end{cases} \tag{8}$$

we find somewhat more spread out groups (Fig. 6). We show the statistics from five different realizations; they show the same trends and suggest that the standard deviations of these properties around the time-varying mean are small.

We can estimate how the tendency for group formation depends on the parameters ($\mathscr{V}$, α, β, and the organism density ρ) as follows: the characteristic convergent velocity is order $\mathscr{V}\rho$, so that the time to accumulate (to move over a sensing distance) is order $1/\mathscr{V}\rho$. The random motions, on the other hand, try to disperse groups, with a characteristic time on the order of κ^{-1}. If this time is short compared with the accumulation time ($\kappa > \mathscr{V}\rho$), the groups will not form. Figure 7 shows the patchiness as a function of time for different κ values; for the mean density and $\mathscr{V}$ used here, groups cannot form for κ greater than 0.05. When we reduce $\mathscr{V}$ by half, the diffusivity necessary to disperse the patches is also reduced by half (to 0.02) as implied by the inequality above. Decreasing the density ρ by a factor of 2 leads to less clear results, since small patches are still formed and then persist; however, the average number in a group remains smaller than five unless $\kappa < 0.02$. Thus we can conclude that groups will form when $\kappa/\mathscr{V}\rho$ is smaller than some constant (about 0.5–0.6 for this weighting function).

2.5.2. *Preferred velocity function of neighbors' velocities*

In both these cases [eqns (6), (8)], however, groups, once formed, tend to persist. Over longer periods the slowly wandering groups encounter each other and merge so that the organisms become more and more patchy. Apparently, the inward accelerations (which increase as the number of individuals in the group) are enough to overcome breakup of a group by random motions once the group has become sufficiently large. In contrast, directional arrayal forces (aligning with neighbors' velocities) tend to lead to more dynamic grouping behavior as shown in Fig. 8. Here the preferred velocity is given by

$$\mathbf{V}_i = \mathscr{V}\tilde{\mathbf{V}}/|\tilde{\mathbf{V}}| \tag{9a}$$

with

$$\tilde{\mathbf{V}} = \Sigma\mathbf{v}_j(1 - |x_j - \mathbf{x}_i|^2) + 0.4\mathbf{W}_2(x_j - \mathbf{x}_i)/\mathscr{V} \tag{9b}$$

representing tendencies to align with and to move towards neighbors. Now, the patchiness index and the average group number fluctuate, indicating the formation and break-up of groups. Examples of such events are shown in Fig. 9.

Finally, we show examples using the schooling model of Grünbaum (1998a). This model is intentionally simplistic; our aim is to include only the barest essentials of schooling behavior (Grünbaum & Okubo, 1994). Each organism travels at the same constant speed. The directions are adjusted each time step by a turn to the left or right, depending on four factors. First, there is a random deflection through a

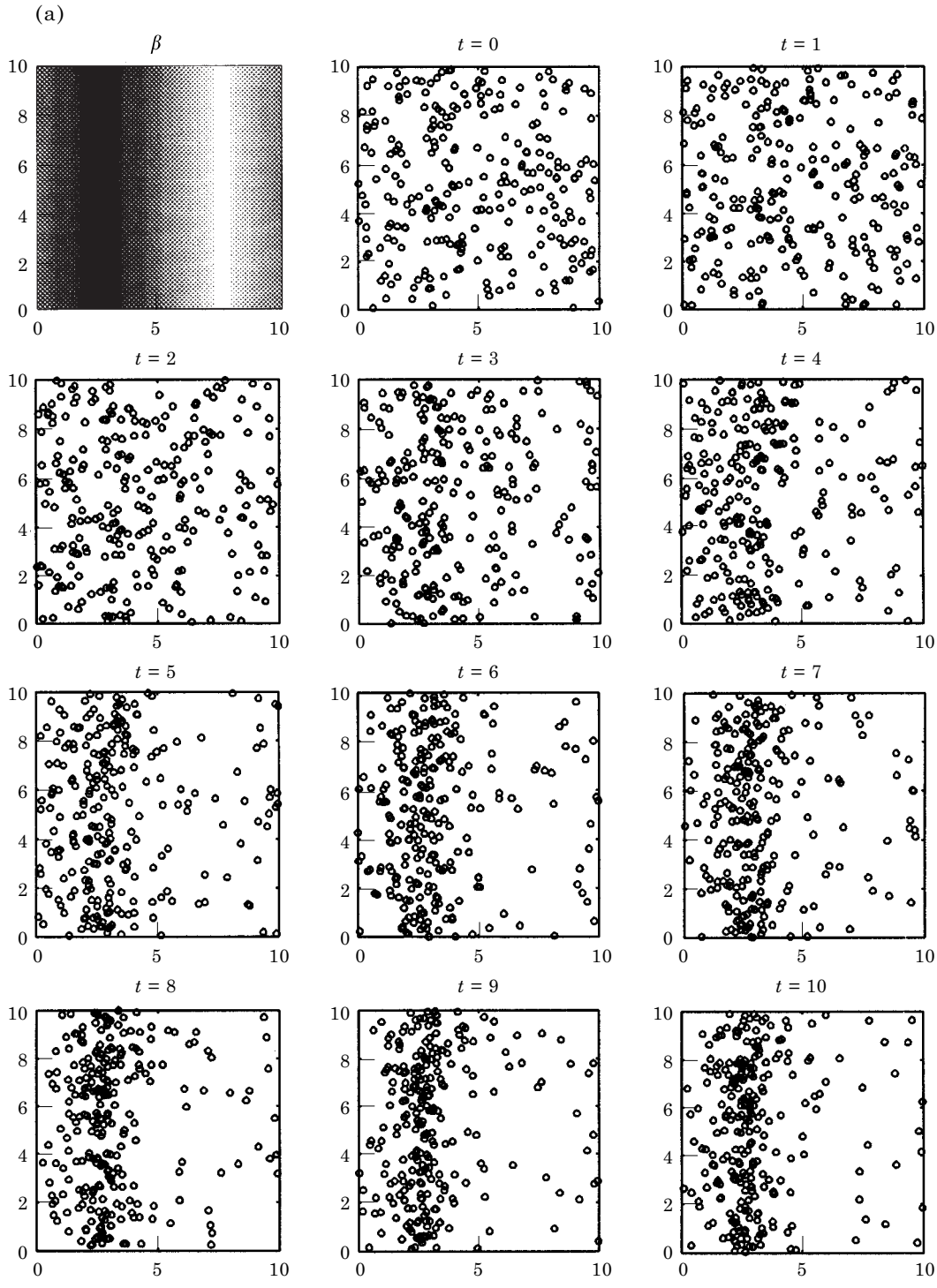


FIG. 3(a)—(*Caption on opposite page*).

fixed angle. We represent the social response by a second contribution to the turn, which is towards the right or left depending on which side has more neighbors within the sensing radius. A third contribution, also representing social behavior, causes a turn away from the side which has more nearby neighbors (within 0.3 sensing distances). Thus we build in a decreasing probability of detection or responsiveness to neighbors at large separation distances and a

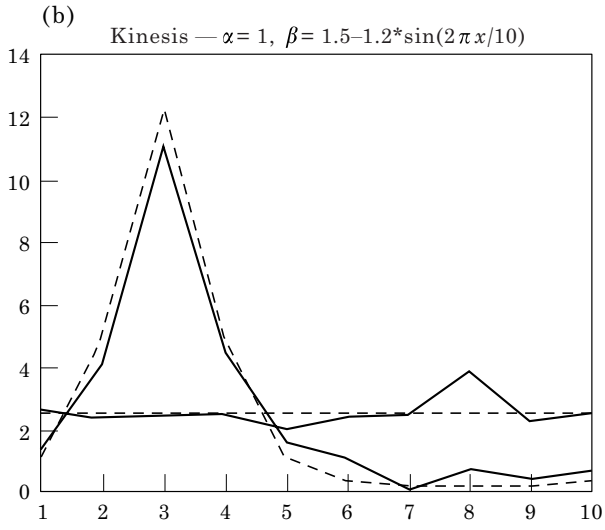


FIG. 3(b)

FIG. 3. (a) Kinesis with a sinusoidal distribution of resource (shown in the first frame). $\beta = [1.5 - 1.2\sin(2\pi x/10)]$. (b) Densities from the experiment and from the continuum theory.

switch from attractive to repulsive interactions as neighbors get too close. Finally, we include a tendency to align with neighbors (Inagaki *et al*., 1976; Aoki, 1982; Matsuda & Sannomiya, 1980, 1985; Huth & Wissel, 1990, 1992; Warburton & Lazarus, 1991; Grünbaum, 1994, 1997; Gueron *et al*., 1996): this small fourth contribution represents turning towards the direction of travel of a randomly selected neighbor. Figure 10 shows examples of the development of schools in this model. For these simulations, we considered larger populations with 512 to 2048 individuals, but a larger domain, corresponding to densities of 0.078 to 0.312 (as opposed to 0.39 for the other experiments). Note again the frequent mergers and splittings; we infer that models where the velocity decreases substantially as groups form are unlikely to split in comparison with cases where the organisms maintain their speed.

2.6. NON-DIRECTIONAL SOCIAL BEHAVIOR

The previous example involved a kind of taxis, with the gradient representing the variation in local density of organisms. Again, clustering can also occur by kinesis, with the random accelerations related to the local density. In this vein, Fig. 11 shows an example where the movements are reduced as the number of neighbors grows: $\mathbf{V}_i = 0$ and

$$\beta = \beta_0 \exp\left(-\sum_{j \neq i} W_b(|\mathbf{x}_j - \mathbf{x}_i|) \right) \qquad (10)$$

with

$$W_b(z) = \begin{cases} w_0(1 - z^2) & z < 1 \\ 0 & z > 1 \end{cases}.$$

Again grouping can occur for suitable choices of α, β_0, and w_0.

In these examples, clustering is the only measure of success considered; more generally, it is a surrogate for more basic information. For example, when resource detection and utilization are important, one can introduce direct measures of success in performing tasks (Grünbaum, 1998b). We do not discuss this topic further in this paper, but it is a rich area for examination.

We shall return to these examples from the continuum point of view (Section 4) and in the presence of turbulent flow (Section 3).

3. The Role of Physics in Inducing Pattern

A starting point for discussion of patterns is the determination of what needs to be explained. As discussed in the earlier sections, many features of biological distributions simply mimic patterns seen in driving physical variables, from soil features to oceanic thermal fronts; therefore, we must first understand the patterns imposed by the physics. For example, Langmuir circulations (Leibovich, 1993) concentrate zooplankton in the surface layer of lakes, estuaries, and the oceans in rows tens to hundreds of meters in size. The "squirts" and "jets" found in eastern boundary current systems (Abbott & Zion, 1987) cause zooplankton concentrations on the scale of tens to hundreds of kilometers. The ocean's ubiquitous mesoscale eddies (Robinson, 1983) also exert profound concentrating effects over the scales of 50 to hundreds of kilometers (Lobel & Robinson, 1988). And the larger ring structures that are generated by western boundary currents like the Gulf Stream (Ring Group, 1981) provide persistent environments separated from the rest of the sea, which allow entire plankton communities to interact over broad

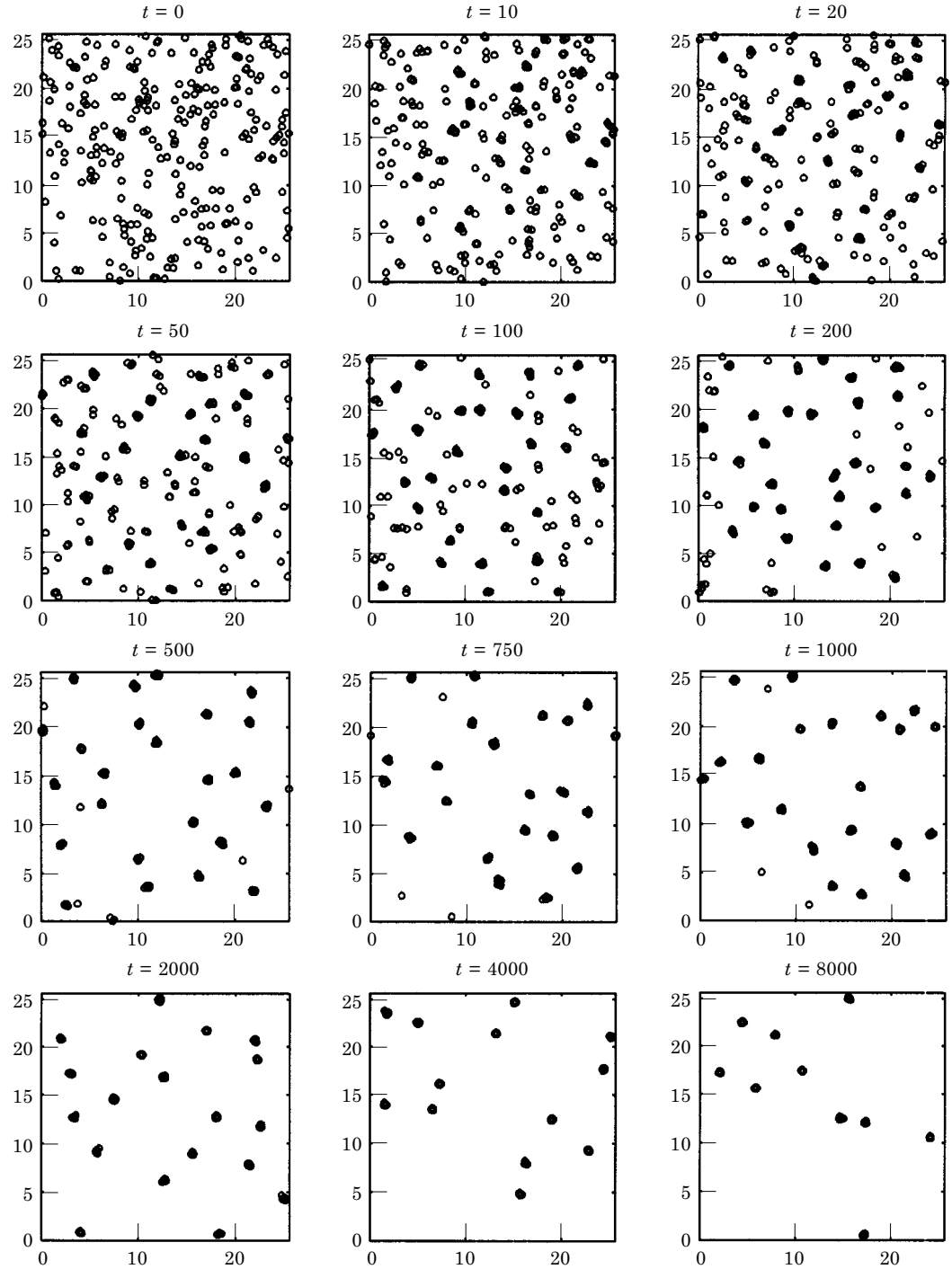


FIG. 4. Social taxis with a preferred velocity directed towards neighbors [eqn (5)]. The parameters are $\alpha = 1$, $\beta = 0.2$, $\mathscr{V} = 0.2$, $\kappa = 0.02$, and $\delta t = 0.125$. The domain is doubly-periodic, with 25.6 units on a side. There are 256 organisms in the domain ($\rho = 0.39$).

time scales as long as years (Mann & Lazier, 1991).

Furthermore, nonlinearity and threshold behavior, persistent features of the dynamics of populations, mean that even biological patterns that are driven entirely by the physical environment may show features that are distinct from those in the driving variables; this complicates the picture fundamentally. For example a large literature exists on critical patch

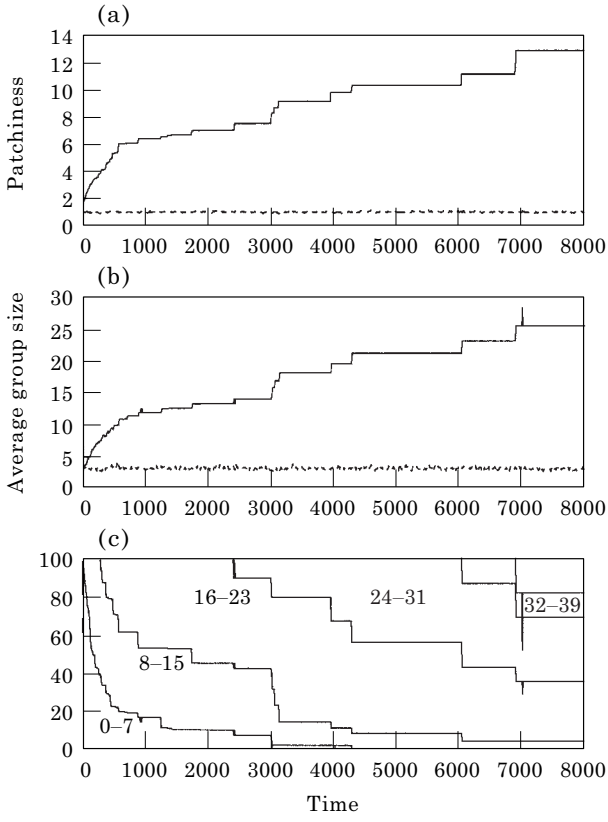


FIG. 5. (a) Lloyd's patchiness index [eqn (7)] for doubly-periodic domain showing (– – –) the case without any social forces and (—) the index calculated from the numerical experiment shown in Fig. 4; (b) average number in a group (which must have two or more members); (c) percentage of the population in groups of different sizes.

sizes in plankton (Kierstead & Slobodkin, 1953; Platt & Denman, 1975; Okubo, 1980), demonstrating that patchiness will not develop at all in finite regions where intrinsic growth rates are below some threshold value (which varies inversely with the size of the region). Beyond that threshold, distributions respond nonlinearly to variability in the nutrient fields. Nonlinearity in population dynamics, moreover has been shown to drive patchiness in homogeneous regions (Segel & Levin, 1976; Okubo, 1974; Levin & Segel, 1976; Segel & Jackson, 1972) or to exaggerate the effects of underlying heterogeneity through quasiperiodic and chaotic responses (Hastings & Powell, 1991; Pascual-Dunlap, 1995).

A first step in the development of individual-based models is to specify only individual responses to environmental cues, rather than to other individuals, and to characterize the patterns that result. This approach removes the nonlinearities stemming from the responses of individuals to each other and the feedbacks that can result. Such an approach has been used by Yamazaki & Osborne (1988) and Olson & Podesta (1988) to describe the responses of individuals to thermal cues, and by Hofmann (1993) to describe the movement of the eggs and larvae of krill in response to physical features of the environment. In this approach, typically, one adopts a Lagrangian description of individual movement, for example writing a force equation for determining the acceleration of each individual. Because individuals are moving independently of each other, it is relatively easy to pass to a description of collective behavior. The situation is made much more complicated, as we shall see in the next section, when individuals are actually responding to each other (Sakai, 1973; Okubo, 1980; Grünbaum, 1992; Gueron & Levin, 1993; Gueron *et al*., 1996), rather than simply to features of the physical environment.

3.1. RESPONSE TO A CHANGING ENVIRONMENT

A classic form of animal aggregation is fish schooling. One question on the macroscale is how fish behavior leads to aggregations that then socially form schools. Here the effect of simple non-directional responses to local conditions (kinesis) is considered as a source for migratory aggregations in a large pelagic, the Atlantic Bluefin tuna (*Thunnus thynnus*). These fish spawn in spring in the Gulf of Mexico and then migrate northwards to New England waters to feed through the summer and early fall. As a null hypothesis, this may be treated as a response to local thermal conditions, i.e. that the fish change their movement behavior depending on whether they are at a comfortable temperature or not. The model assumes the tuna have no sense of conditions other than at their current position, and that they sense ambient temperature relative to some preferred temperature, similar to Neill's (1979) klinokinetic model of behavioral thermoregulation. The individual fish adjusts its probability of changing direction according to the difference between ambient (T) and preferred (T_0) temperatures. If the difference is very small, the tuna tends to continue moving in the same

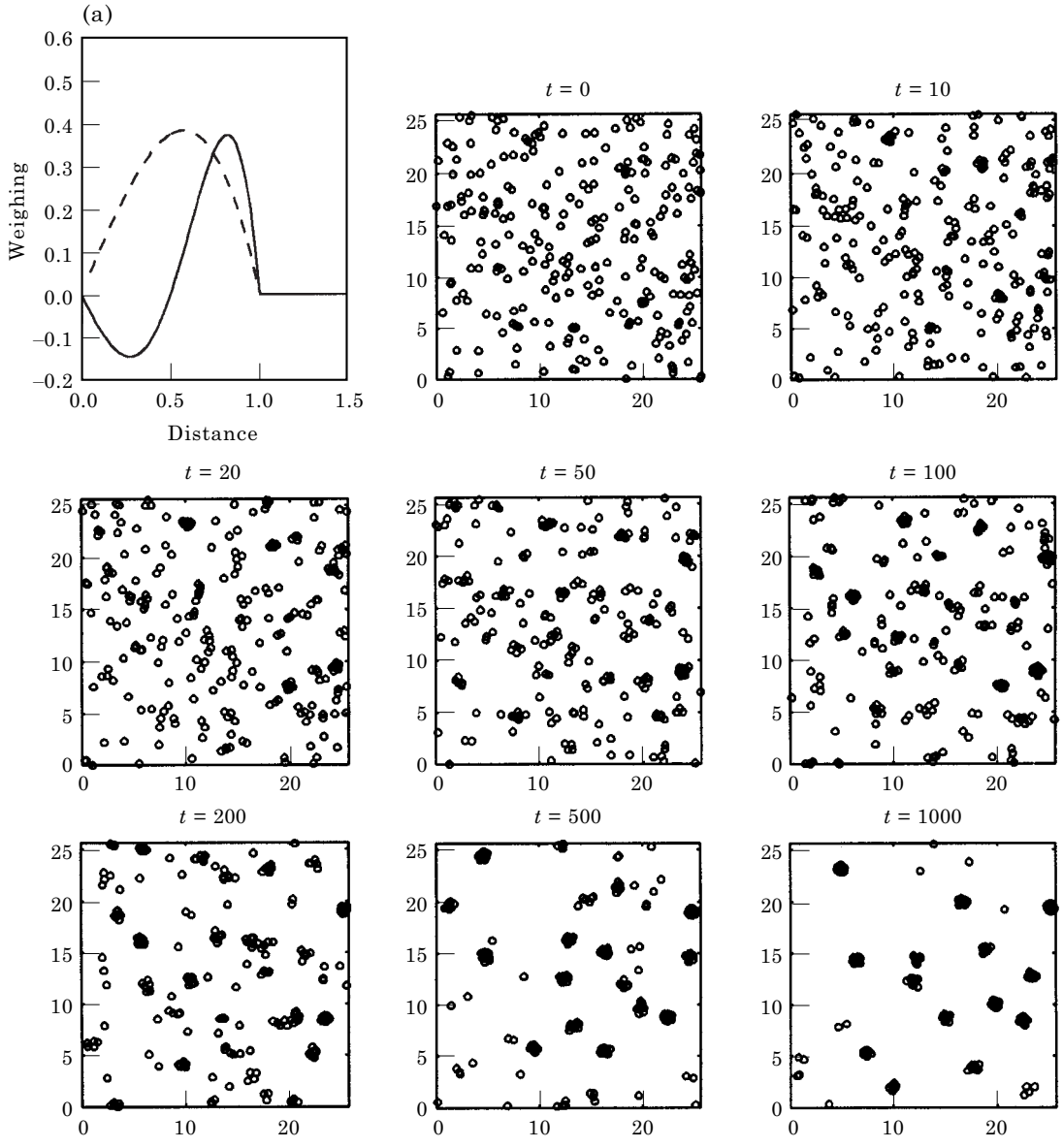


FIG. 6(a)—(*Caption on opposite page*).

direction as the previous time step. As the difference increases, the fish maintains less of its previous velocity and begins to move in a more random manner. If the fish is far outside its comfort range, movement becomes a random walk. In terms of eqns (2)–(3), we choose

$$\alpha = 1 - 0.75 \exp[-(T - T_0)^2/2\sigma^2]$$

and

$$\beta = 1 - 0.9 \exp[-(T - T_0)/2\sigma^2].$$

The first experiment was run on a Cartesian grid with the temperature initially set to optimum at the origin and decreasing as a cosine function of Y. A seasonal temperature cycle is simulated by then raising the temperature as a $\cos(t)$ (Olson *et al*., 1996; Humston *et al*., 1998). As the temperature rises at the origin, the fish migrate out along the Y-axis to maintain their preferred thermal domain. This takes them 1000 km out the Y-axis until the temperature begins to cool and the return migration starts.

The first simulation (Fig. 12) considers the annual migration from south to north and then back again to the Gulf of Mexico. The temperature changes as a sinusoidal function of

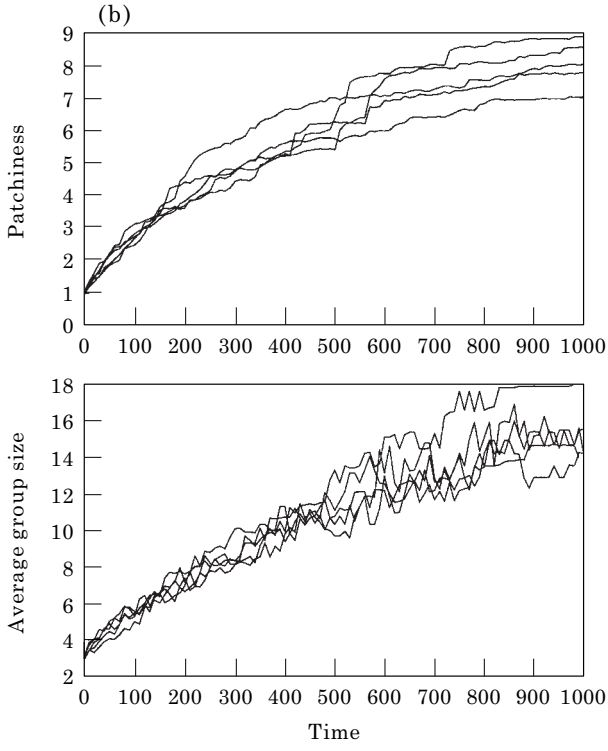


FIG. 6(b)

FIG. 6. (a) Social taxis with a repulsion from near neighbors [eqn (8)] as shown in first frame: (—) $\mathbf{W}_2$, (– – –) $\mathbf{W}_1$. (b) Patchiness index and average number in a group for five different realizations.

time, causing the preferred isotherm to migrate northwards with the onset of summer and then back southwards in the fall and winter. The fish are initialized at $y = 0$ at $t = 0$ years. The figure shows the second year of simulation. At $t = 1$ year most of the tuna have returned to the origin, although there is a more or less even distribution of "lost" fish distributed over the entire domain. As the origin begins to warm in spring, the mode in the distribution moves northward ($t = 1.25$ years). The distribution of fish is not symmetric, but has a sharp leading edge and a long tail of fish that are behind the preferred isotherm position. In the summer, the bluefin reach their preferred thermal regime ($t = 1.5$ years) and then return southward in the fall. While there are undoubtedly other factors such as water clarity that influence the annual migration of this species in the western North Atlantic, the model shows that temperature alone with simple kinesis leads to a realistic aggregated group migration. At finer spatial scales the simple kinesis model also reproduces the observed distribution of bluefin schools on surface thermal fronts in the Gulf of Maine. Aerial spotting of surface schools [Fig. 13(a)] shows them to be aggregated in temperatures just above 18°C. The simulation with this isotherm gives similar patterns [Fig. 13(b)]. Again, other factors are involved in determining these patterns. In fact tuna have to move through colder waters to show up in places like the warm region inside the hook of Cape Cod in Fig. 13(a). The simple kinesis model does, however, produce a reasonable description of bluefin schools along thermal fronts, and therefore is a good starting place for a more complete behavioral model.

3.2. GROUPING IN TURBULENT FLOWS

In the previous section, it was shown that common responses to environmental cues can result in aggregation. This is consistent with observations of a wide range of organisms, including birds, ungulates, and a variety of terrestrial and marine invertebrates. Physical forces such as turbulence, however, can also serve to disrupt aggregation patterns established

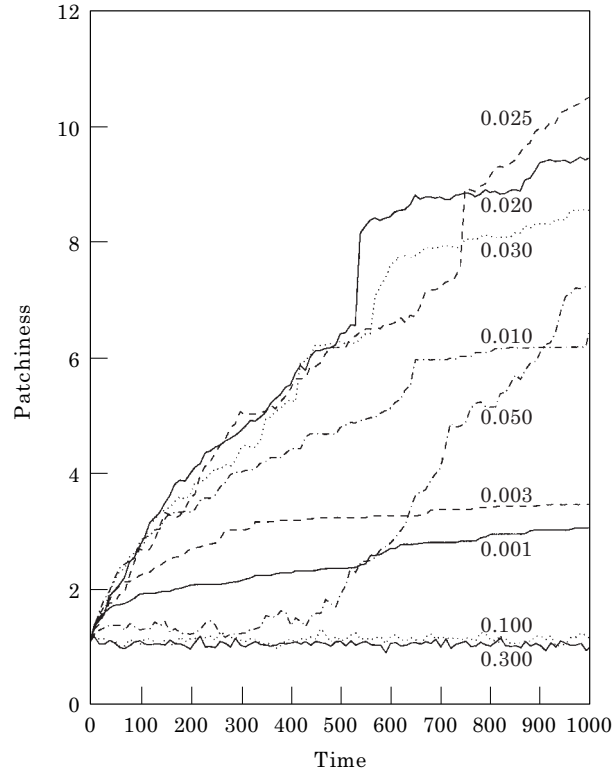


FIG. 7. Patchiness index for different diffusivities κ in (4).

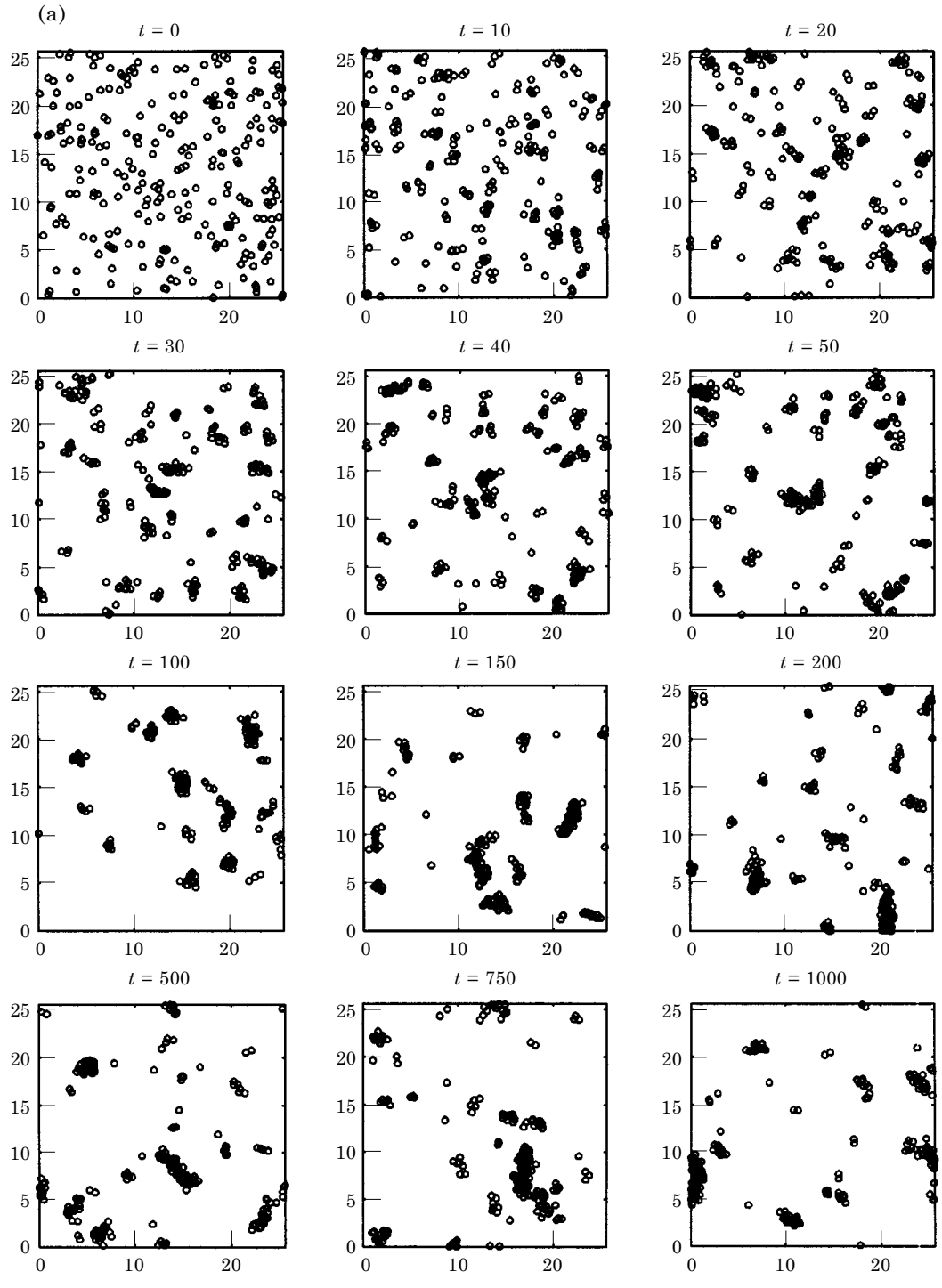


FIG. 8(a)—(*Caption on opposite page*).

in response to other factors. In the ocean, species with natural grouping tendencies must fight against the dispersive effects of turbulence. To what degree can they succeed? Does the turbulence affect the grouping in the same way as an increased diffusivity would, or are the effects more complex?

To answer these questions, we have coupled an aggregation model to a two-dimensional fluid flow model (three dimensions being

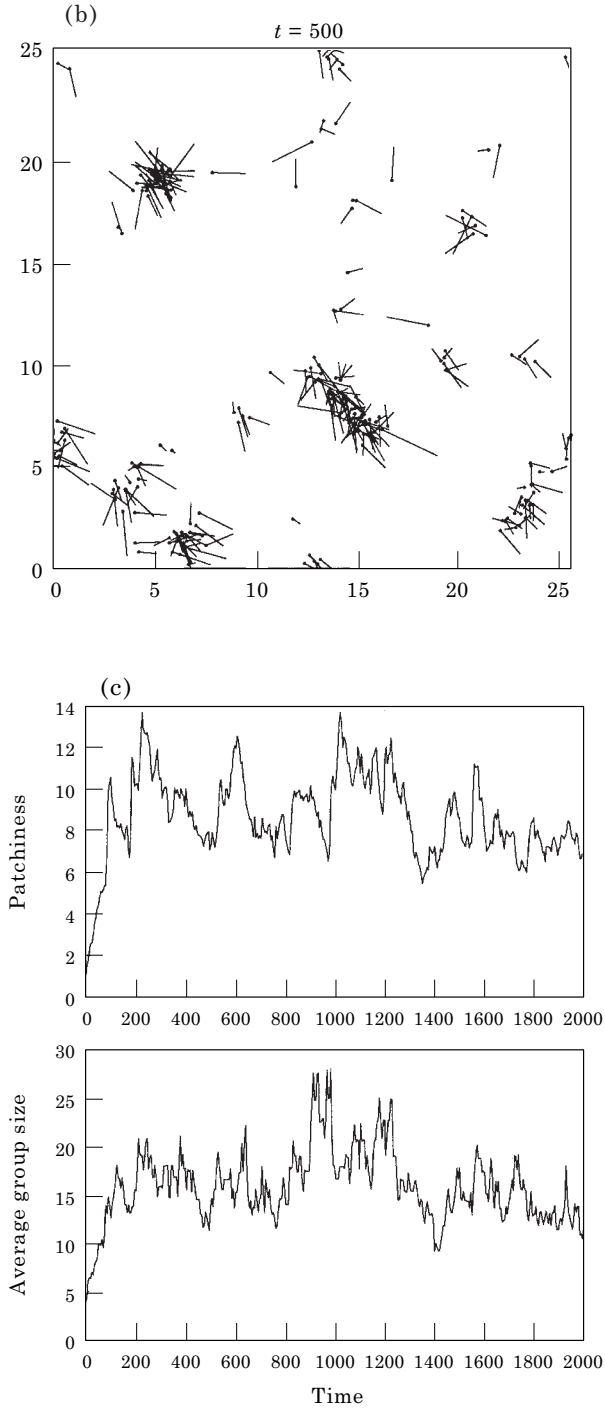


FIG. 8(b) and (c)

FIG. 8. (a) Social taxis with tendency to align velocities as well as attraction/repulsion [eqn (9)]. (b) Velocity pattern, showing the tendency for alignment. (c) Patchiness index and average number in a group.

computationally too expensive!). To enforce incompressibility of the fluid, the velocities are written in terms of a streamfunction ψ (which has its level lines instantaneously parallel to the flow): $\mathbf{u} = (-\partial\psi/\partial y, \partial\psi/\partial x)$. In this case, the Navier–Stokes equations can be reduced to a single PDE for the evolution of the vorticity $\nabla^2\psi$ (twice the local rotation rate of fluid parcels):

$$\frac{\partial}{\partial t}\nabla^2\psi + \mathbf{u}\cdot\nabla(\nabla^2\psi) = \text{Forcing} + \nu_{-1}\nabla^{-2}\psi + \text{Filter}. \quad (11)$$

The flow is forced randomly (Maltrud & Vallis, 1991) and dissipated at both large and small scales to keep the energy concentrated in eddies that are well resolved but do not fill the domain (Sundermeyer, 1995). The flow is doubly periodic [$\psi(x + m\mathscr{L}, y + n\mathscr{L}) = \psi(x, y)$ for integer m, n and domain size $\mathscr{L}$]. A high wavenumber filter is used to prevent buildup of mean square vorticity in the small scales while the low wavenumber filter prevents energy accumulation in the largest scales. The flow was begun with a specified spectrum and allowed to evolve until the forcing and dissipation came into balance. The energy continues to fluctuate to a small degree (Fig. 14). Maps of the streamfunction for the period after equilibration are also shown.

We have explored the individual-based approach in which the flow is seeded with organisms that are advected by the flow as well as moving under their own volition

$$\frac{\partial}{\partial t}\mathbf{x} = \mathbf{u}(\mathbf{x}, t) + \mathbf{v}(\mathbf{x}, t). \quad (12)$$

3.3. PASSIVE TRACER DISPERSION

To understand how behavior modifies the distributions, we must begin with the passive case ($\alpha = \beta = 0$). Figure 15(a,b) show the locations and trajectories of a set of "floats" in the turbulent flow. Figure 15(c) shows the evolution of the mean squared displacement, $D = |\mathbf{x}_i(t) - \mathbf{x}_i(0)|^2$, with time. Note that the displacement grows more or less linearly; however, if we examine the Taylor (1921) estimate of the diffusivity

$$\kappa = \frac{1}{4}\frac{\partial}{\partial t}D, \quad (13)$$

we find that κ varies between 0.5 and 2, with an average value of about 1. Thus the turbulent diffusion is rather strong.

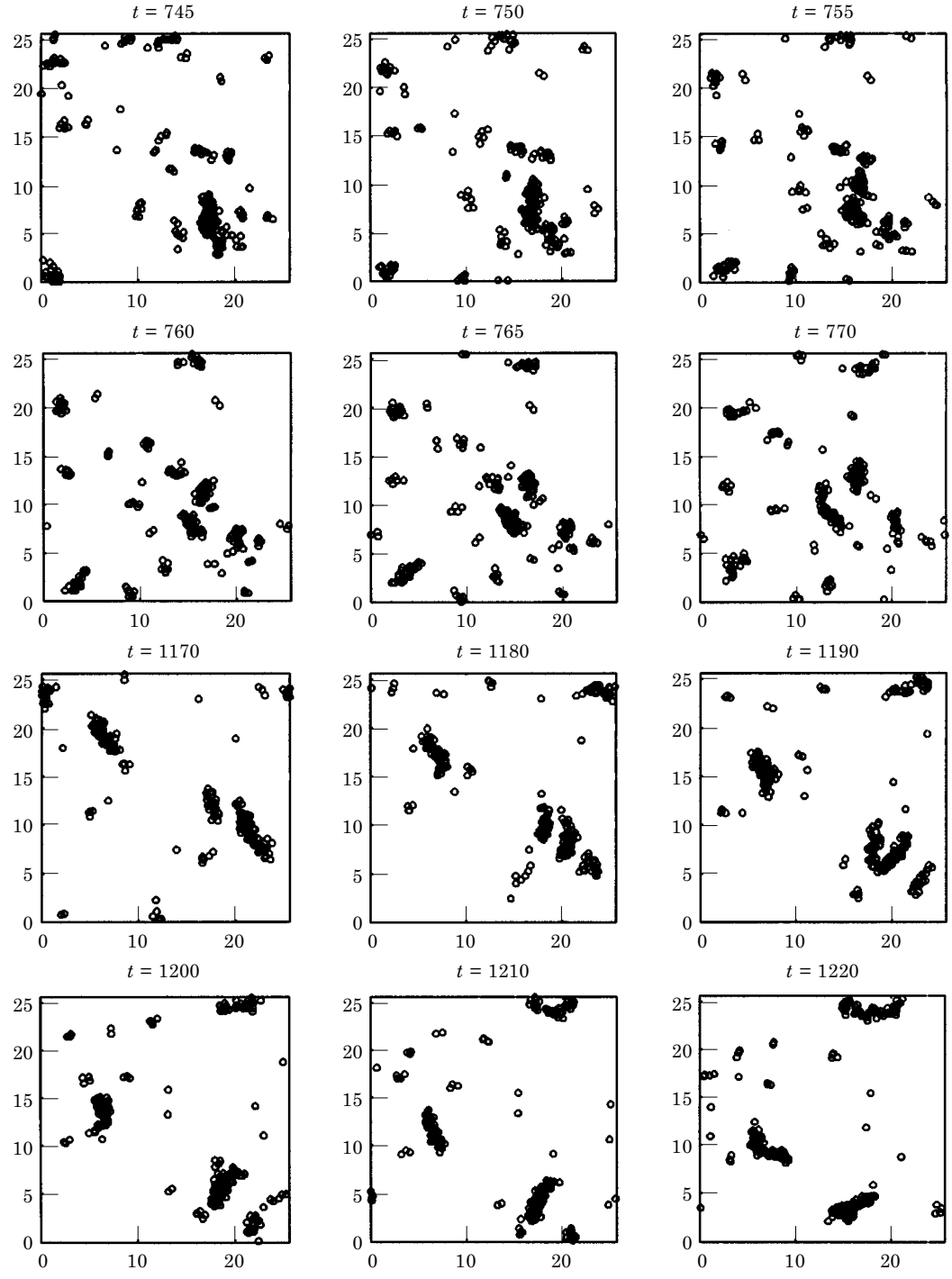


FIG. 9. Group breakup (top six panels) and merger (bottom six) events.

3.4. FROZEN FIELD

Against this background, we consider the simplest form of grouping, kinesis, given by eqns (2), (3), (6) with $\mathbf{V} = 0$. We choose an exponential dependence of β upon ψ giving positive values and about a two order of magnitude range: $\beta = 2.5 \exp(\psi - 2)$. When the flow is steady (ψ held fixed at its initial value), the organisms gradually migrate to the areas where β (or the diffusivity) is small—the lows in the streamfunction field (Fig. 16). The flow

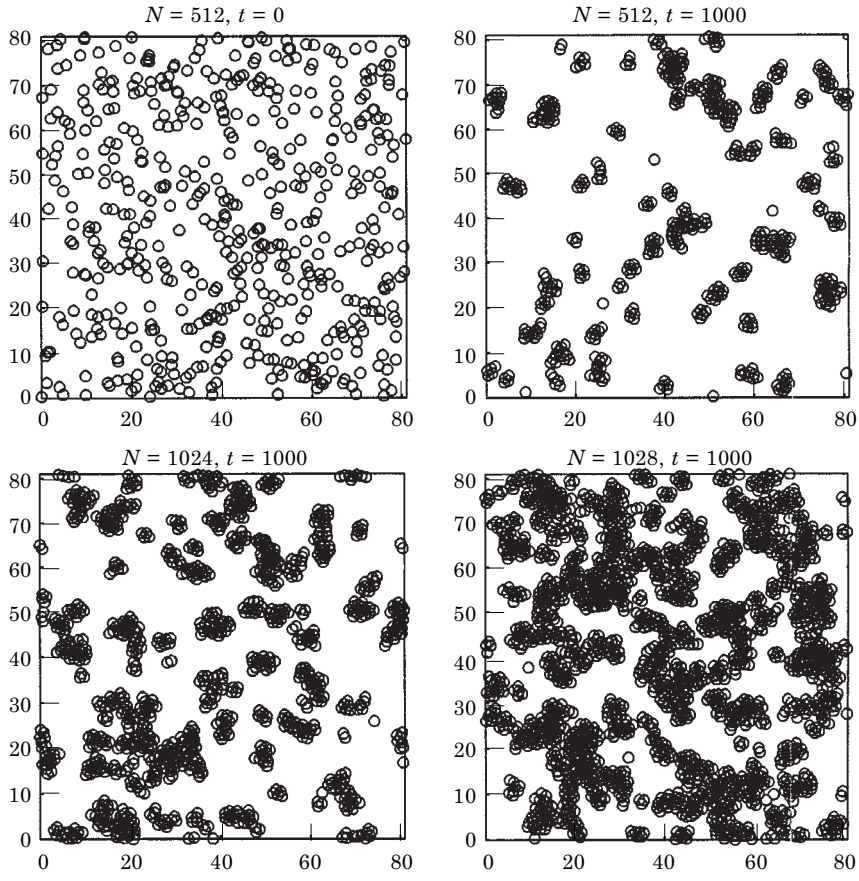


FIG. 10. Schooling behavior, showing one initial state and final states for three different densities. The schools are not fixed, but merge and split frequently (Section 5).

cannot concentrate the particles, since it is non-divergent; rather it advects and, via shear dispersion, mixes them along the streamlines. Kinesis results in cross-streamline movement and the high particle concentrations in the favorable regions.

3.5. EVOLVING FLOW FIELD

If we now allow the flow to evolve, we see similar migration, though somewhat less successful (Fig. 17). Whereas the tuna example above had single time and space scales, the flow here has full frequency and wavenumber spectra. The value of ψ now changes with time following a particle, so that favorable (low β) regions may become unfavorable. If this happens too rapidly, the kinesis will not be effective, and groups will not form. For the parameters we have used, however, the organisms can overcome the dispersion and the changing conditions to collect in favorable regions.

3.6. SOCIAL BEHAVIOR: TURBULENT DIFFUSIVITY AND TURBULENT FLOWS

Turbulence is often regarded as causing a much-enhanced "eddy diffusivity," which is nearly unity in the experiments [Fig. 15(c)]. If that were the case, we could predict grouping behavior by adding back randomness of suitable strength to the model without flows. The β value that would give an equivalent growth in $|\mathbf{x}_i(t) - \mathbf{x}_i(0)|^2$ is about 1.4, far beyond the range in which groups can survive. Eddy diffusion, however, turns out to be a poor approximation for turbulent flow.

When we add more physically realistic turbulence, we find that the groups can still form (Fig. 18); although the physical dispersion (as measured by the eddy diffusivity) is stronger than the grouping tendency, the turbulence does not inhibit the formation of groups. We shall see that the result reappears below when we use the continuum model and calculate the stability—

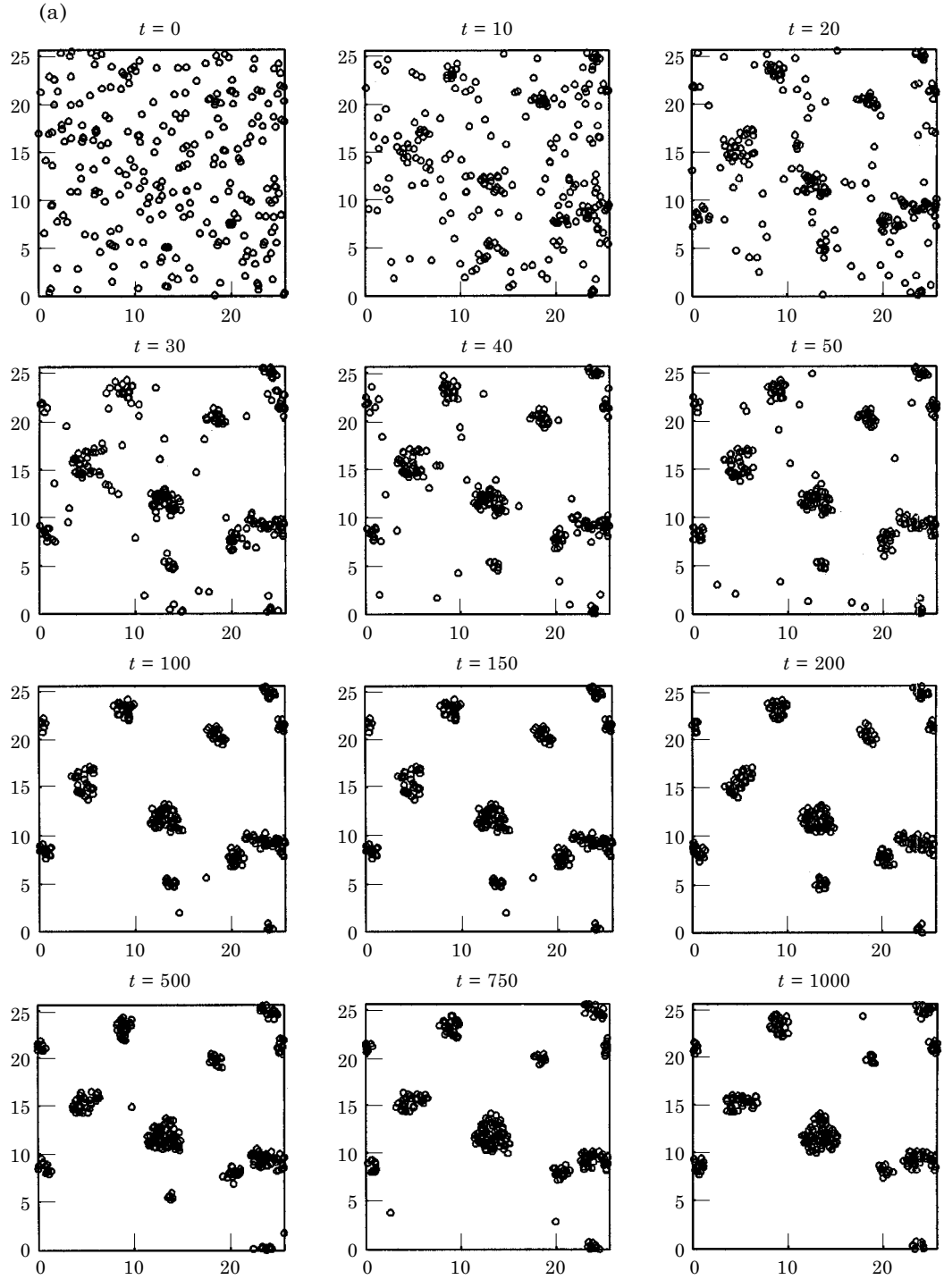


FIG. 11(a)—(*Caption on opposite page*).

the turbulent diffusivity is strong enough to stabilize the uniform state, yet groups still form in the actual turbulent flow. We attribute this to the presence of temporary "refuges," where the shear and strain fields are not too large and the groups can condense.

However, the turbulence has a much more profound effect once groups form: it increases the encounter rate between groups (or groups and individuals) so that mergers occur much more frequently. The patchiness index [eqn (7)] increases rapidly. The average group size appears

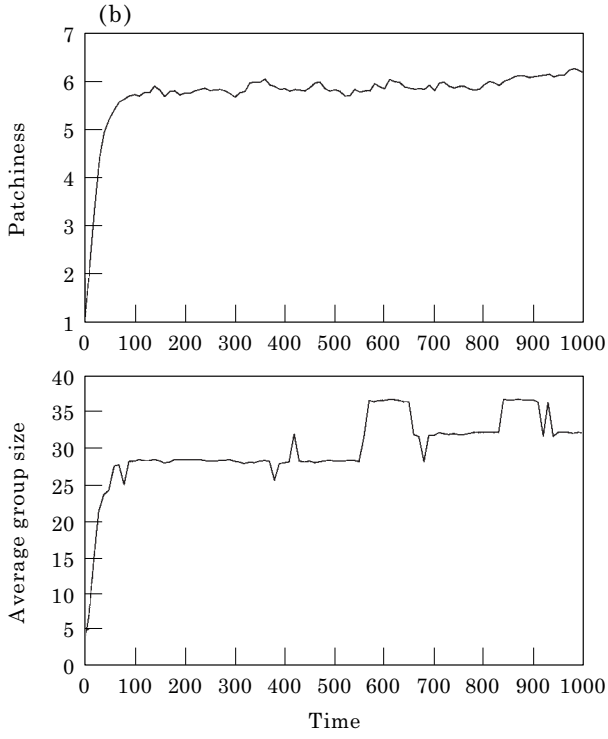


FIG. 11(b)

FIG. 11 (a). Social kinesis with the random motion decreasing as the number of neighbors increases [eqn (10)]. The pattern is changing slowly after $t = 100$, with small groups being absorbed occasionally into larger groups. (b) Patchiness indices for non-directional social behavior.

initially similar to the run without turbulence, but the group size distribution becomes quite different: at the end of the run shown, one group has 77% of the individuals, one group 9% and the rest of the groups consist of single individuals.

The interplay between grouping and turbulence can be studied either by beginning with a flow regime and considering stronger and stronger grouping tendencies until patches form, or by beginning with a grouping model and increasing the strength of the turbulence until groups disintegrate. When there are no grouping forces (Fig. 19), the organisms do not aggregate and the patchiness index remains small. The eye does pick out larger scale aggregations occasionally; calculating group sizes where individuals are within a distance of 1.5 of each other (50% larger than the sensing radius) does reveal these groupings, but they are transient, essentially statistical fluctuations. We now consider varying the grouping strength by altering the parameters such that groups form in the same configurations (in the case without flow), but in a shorter or longer time. To double the grouping rate in (1)–(3), we require exactly the same $\delta \mathbf{x}_i$ values but with time steps δt that are half as large. Therefore we need $\mathbf{v}_i$ to be twice as large. This requires $\mathscr{V}$ and α to be doubled also and β to be multiplied by a factor of $2^{3/2}$ (since $\delta \mathbf{e}$ is reduced by a factor of $1/\sqrt{2}$, or, alternatively, κ is doubled). Similar arguments apply for other speed-up or slow-down factors. As we increase

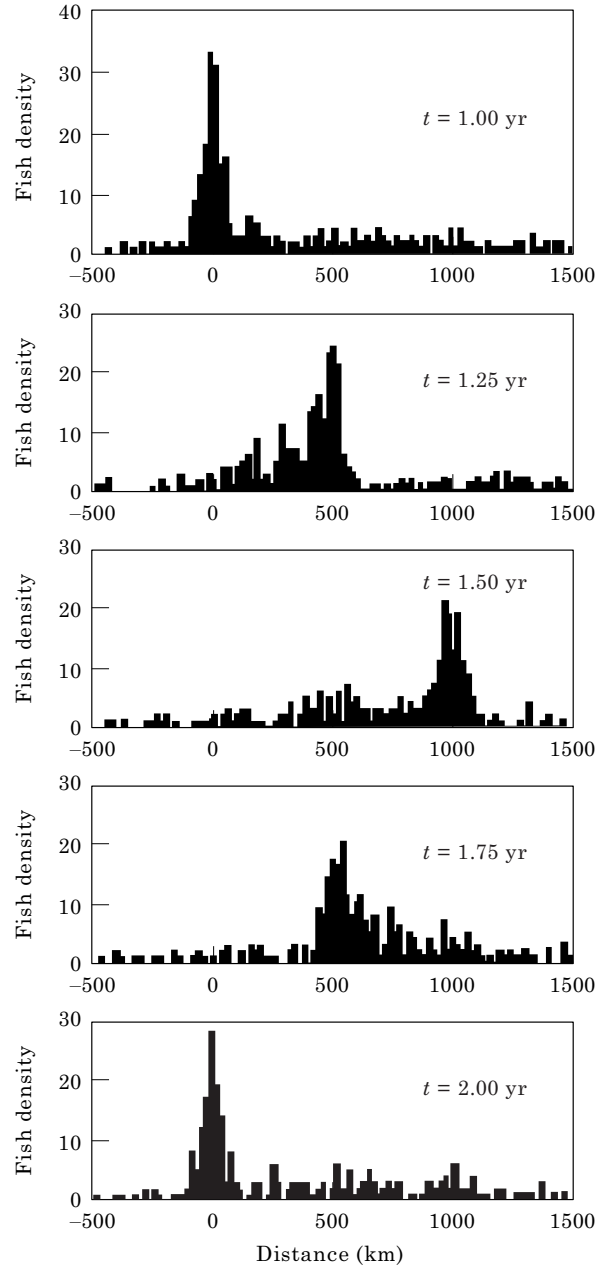


FIG. 12. Simulation of tuna migration.

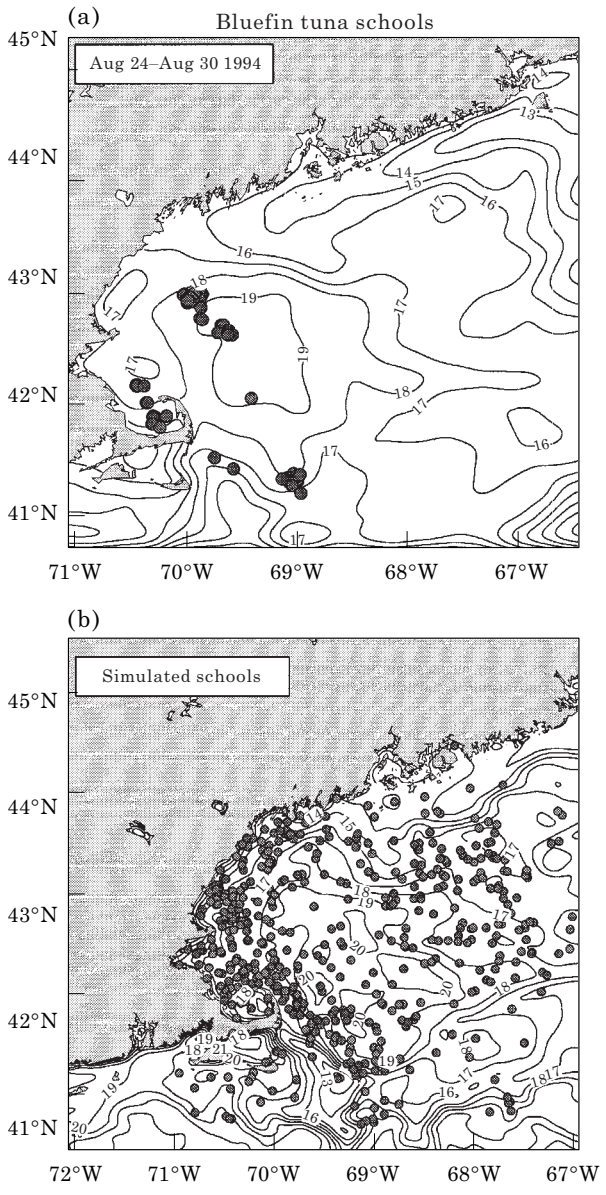


FIG. 13. (a) Observed distribution of bluefin schools. (b) Simulation using kinesis model.

the grouping strength, we find that groups form for a strength of 70% of the original parameters (or, alternatively, a flow that is 43% stronger), but are transient and small (Fig. 20). For stronger grouping tendencies relative to the flow, we see distinct grouping and group mergers such as those shown in Fig. 19. In Fig. 21, we compare the patchiness vs. time for different grouping rates with and without turbulent flow.

The final figure in 21(b) shows that the simultaneous rescaling of time and the grouping strength parameters gives consistent results. These experiments show that strong turbulence can tear apart groups and prevent patchiness, but weak flow actually enhances grouping; such flows do not cause disintegration, but rather increase the rates of contact of individuals and groups, hence speeding group formation and merger. This seems analogous to the problem of the optimal design of spider webs (Craig, 1987): web characteristics that lead to slight oscillations enhance prey capture by increasing contact. Similarly, Rothschild & Osborn (1988) showed that turbulence can increase encounter rates of predators with their prey. Not surprisingly, however, strong turbulence added to the grouping model destroys the fine-scale structures.

The perception distance still plays an important role in the group formation process, of course. The density must become high enough locally to initiate group-formation; after that, the turbulence simply brings new members close enough to be brought in or brings groups within the range where they can merge. The late-stage groups can be quite widely separated, but remain compact. Of course, large scale structure can be generated by other means; studies of taxis with light or depth as a cue suggest that upward swimming and convergent flow (Stommel, 1949; Olson & Backus, 1985, Franks, 1992; Epstein, 1996), can generate patterns similar to observed distributions.

4. From Microscopic to Macroscopic

We shall now examine two approaches to generating higher level descriptions for movements and dynamics of social populations. The first attempts to generate spatially explicit partial-differential equation approximations to the density distribution resulting from the social behavior, i.e. models of the spatial dynamics of groups. The second approach (in the next section) attempts to describe directly the "population dynamics" of groups, i.e. the group-size distribution resulting from size-specific group characteristics that determine fusion and fission rates. Usually this description is non-spatial and statistical, typically yielding a set of coupled, nonlinear ordinary differential equations for the number of groups in each size bin. Group-size distributions may subsequently

be approximated by partial differential equations. To obtain a spatially explicit model, local estimates of group-size distributions, together with size-specific velocity and taxis, may, with appropriate assumptions, be combined into population-level advection–diffusion equations. These implicitly incorporate the effects of individual social behaviors in the form

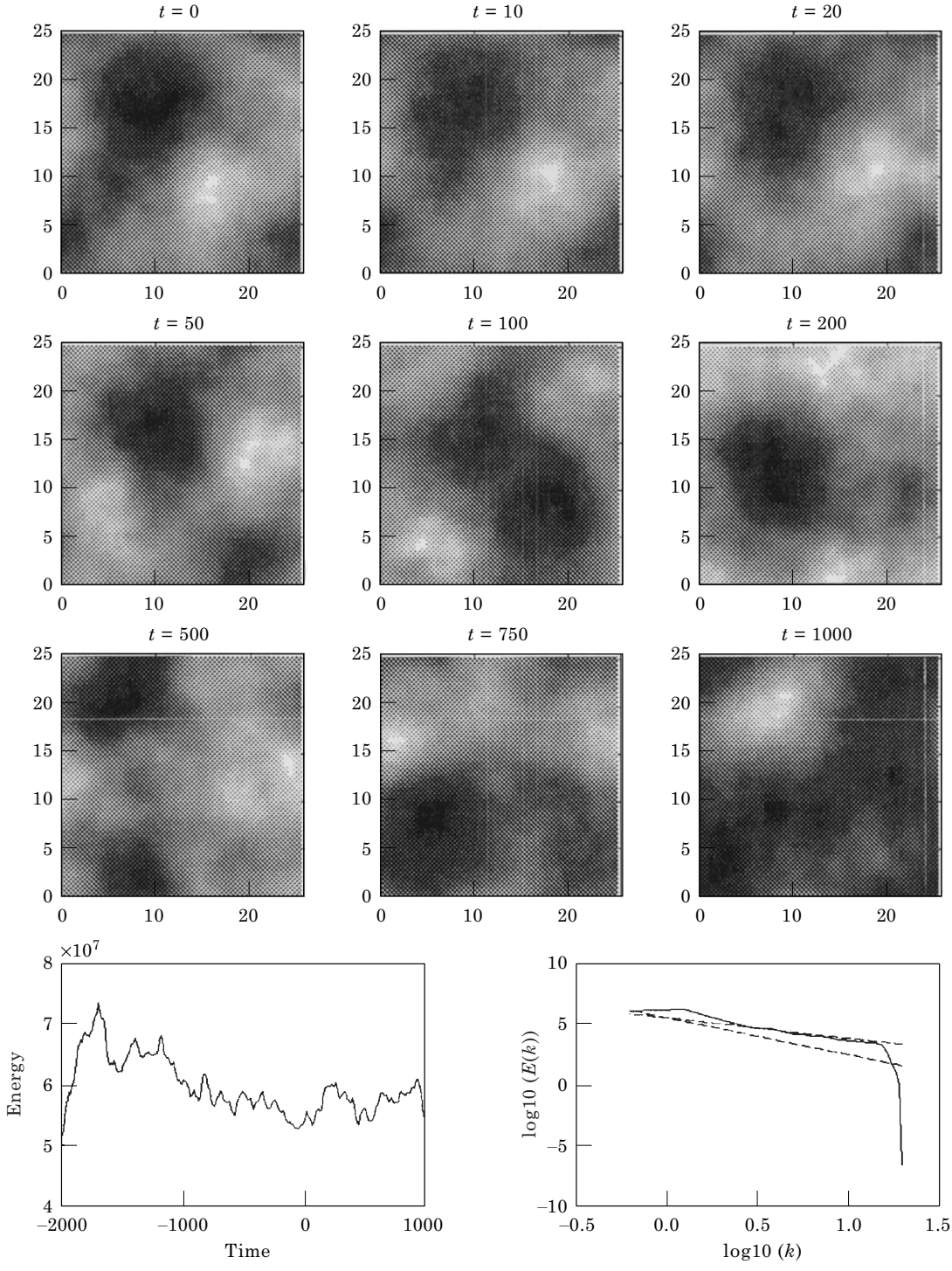


FIG. 14. Streamfunction greyscale maps and energy time history for the turbulent flow (the last 1000 time units are used for the ψ plots and later calculations). Highs (white) in ψ correspond to centers of clockwise circulation, while lows (black) are centers of counterclockwise swirl. Also shown is the final wavenumber spectrum (– – –) are $k^{-5/3}$ and k^{-3}.

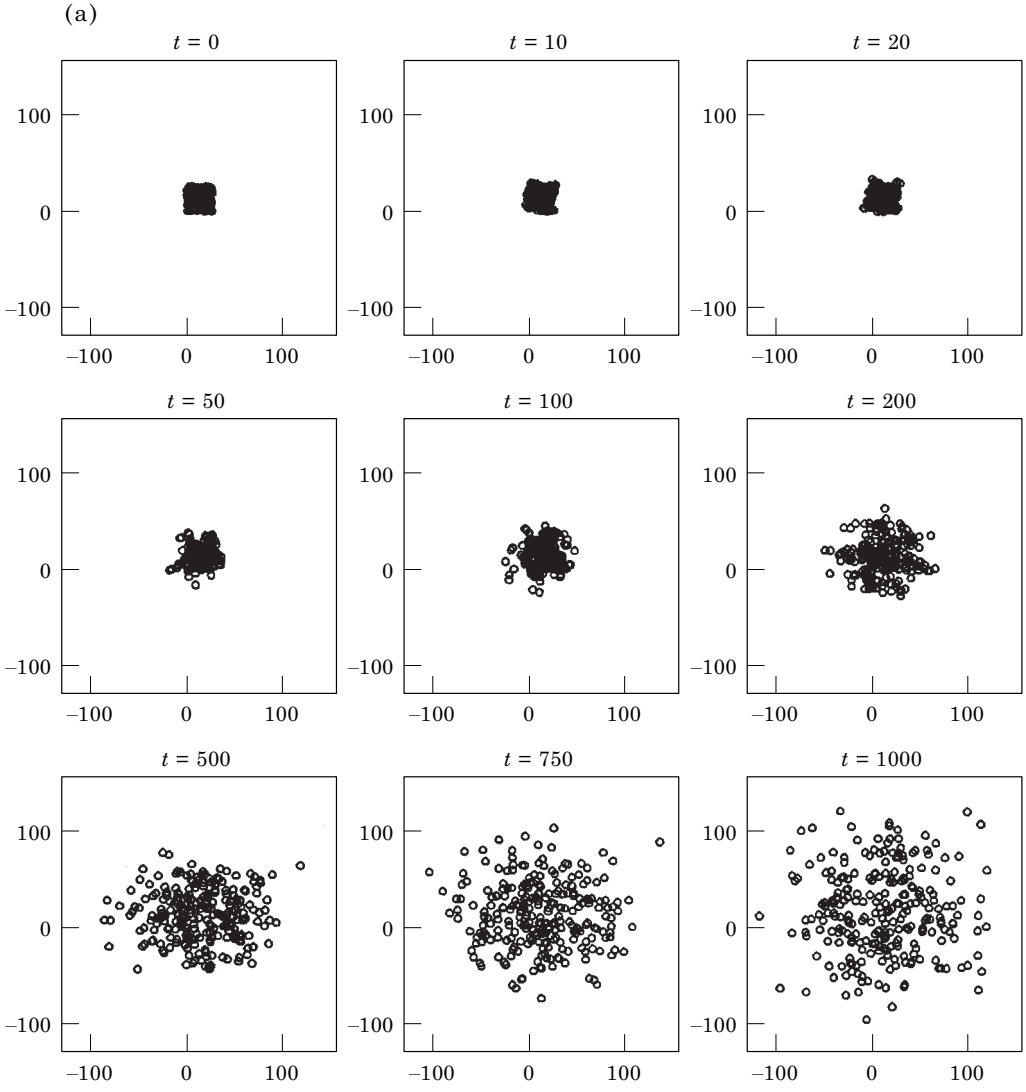


FIG. 15(a)—(*Caption on opposite page*).

of density-dependent advection and diffusion coefficients. In this context, "local" implies length scales much larger than that of groups, but small compared with distances over which group densities change significantly. These approximate, higher level descriptions for social animals form a complementary set of tools, which may by themselves or in combination help determine both the role of social behaviors in creating observed patterns of animal distribution and some of the pressures driving the evolution of specific social interactions.

In general, the spatial density of a population of socially interacting organisms, ρ, will be changed by convergence or divergence of the flux, $\mathbf{J}$:

$$\frac{\partial}{\partial t}\rho = -\nabla \cdot \mathbf{J}. \tag{14}$$

In the more standard population-level descriptions of asocial behaviors such as chemotaxis, the flux in eqn (14) is usually derived from a Fokker–Planck equation and takes the form of an advection–diffusion equation with coefficients that are functions of t and x. If instead we are principally interested in socially interacting populations, we anticipate that the flux represented by $\mathbf{J}$ will be a functional of ρ as well. A

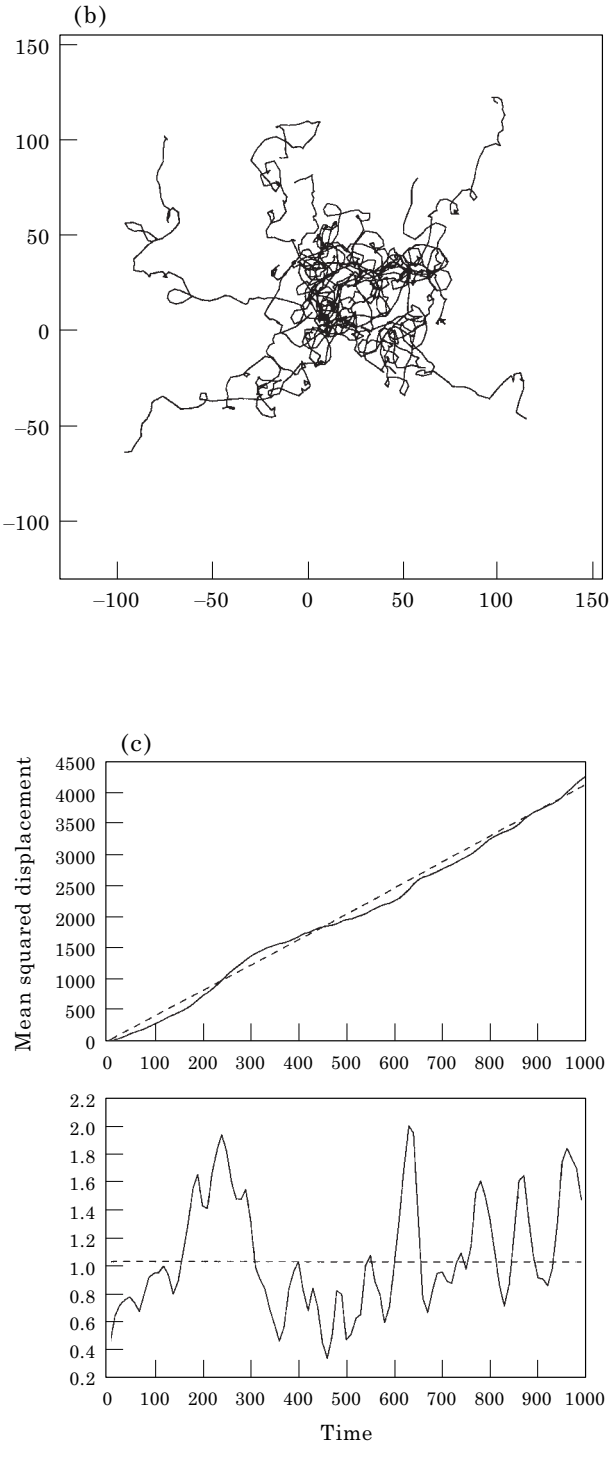


FIG. 15(b) and (c)

FIG. 15. (a) Positions of floats in turbulent flow. In this case, the flow itself has a periodicity of 25.6 units in both directions, but we seed it only in one region; i.e. the absolute displacement of each float is calculated. (b) Trajectories of 16 of the particles. (c) Growth of the mean squared displacement with time: (– – –) shows the spread predicted by a diffusion model with $\kappa = 1.03$; the lower panel compares the two terms in (13).

number of studies investigate population distributions resulting from nonlinear diffusion, advection, and/or reaction terms (Mimura & Yamaguti, 1982; Dal Passo & de Mottoni, 1984; Grindrod, 1991). However, the nonlinearities in these models are not associated with particular individual-level behaviors that give rise to the population-level movements. In this section, our primary focus is to derive explicitly population-level fluxes in eqn (14), beginning with specific social responses between neighbors at the individual level.

4.1. STATISTICS AND SOCIAL BEHAVIOR

Carrying out this explicit derivation immediately presents the difficulty that in order to calculate the flux, **J**, we are forced to represent the probability distributions of forces such as $F_{i,g}$ (social attraction and repulsion) in terms of functionals of macroscopic variables such as ρ. However, the expected value of something like the grouping force, which is generally a nonlinear function of the set of vector distances to neighbors [e.g. eqns (4)–(6)], cannot be expressed as a functional only of the density. Instead, it will also depend upon the detailed shape of the probability distribution for the positions. In other words, the spatial occurrences of individuals is underdetermined by the density ρ in the case of social interactions between neighbors.

Thus, density is by itself generally not sufficient in the case of social behaviors and even its interpretation is open to question. To yield an equation describing density flux, the higher order statistics must be somehow implicit in the density distribution, a requirement not present for density-independent behaviors and one not addressed in most models to date. Otherwise, the model must be formulated to include these statistics as part of the calculation. We are aware of three basic strategies for overcoming the under-determination of the micro-scale individual distributions while retaining only ρ:

(1) assumption of a specific relationship between the spatial density (or zeroth moment), ρ, and higher joint probability distributions of neighbors. For example, assuming individuals occur as Poisson points specifies all

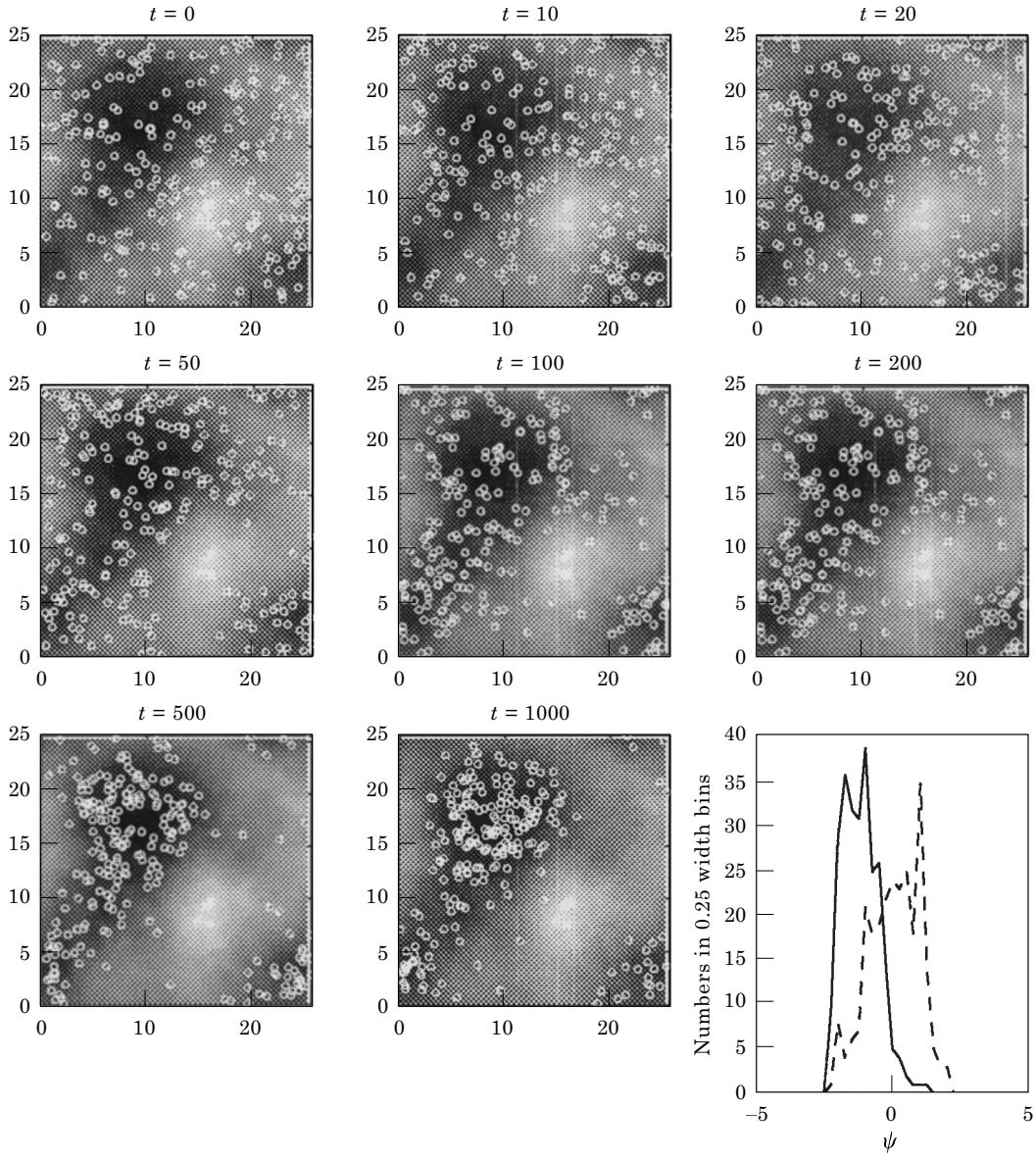


FIG. 16. Kinesis in frozen field. The random motion is small in regions where ψ is negative. The final panel shows a histogram of the number of particles between various levels of ψ with (– – –) being the initial distribution.

statistics of individuals occurring in a neighborhood (Grünbaum, 1994). We discuss this approach in detail in Section 4.2 and the following development. Alternatively, characteristic relationships between moments could be "measured" from individual-based simulations;

(2) an "internal energy" approach (Cohen & Murray, 1981; Murray, 1990), in which a given population density is associated with a characteristic internal energy whose gradient determines population flux. The internal energy could be derived analytically from individual behaviors, in which case some assumptions such as those in (1) would be required. Again, the energy function could be "measured" from individual-based simulations, in which case the higher statistics of spatial distributions would be implicitly contained in the estimated energies;

(3) assumption that group boundaries essentially represent discontinuities of population density and behavior (analogous to surface tension), and that the dynamics of these

boundaries can be represented directly without explicit computation of spatial distributions within them (Gueron & Levin, 1993). In this case, the problem is reduced to the interactions of neighbors in the vicinity of group boundaries. Within this restricted spatial domain, however, the same options as in (2) are presented.

All these approaches emphasize the fact that the definition and use of density distributions for interacting individuals (i.e. ones following density-dependent rules) involve some subtle issues that do not arise in models of density-independent behavior such as chemotaxis. Mathematically, if we have many realizations of the animal distribution—or, in the case of a statistically stationary distribution, multiple sample of the same realization over time—we can define the expected number of animals $\bar{v}_i$ in a given volume $\mathscr{V}_i$ at time t. This is related to the density distribution at a point by

$$\bar{v}_i = E\{v_i\} = \int_{v_i} d\mathbf{x}\rho(\mathbf{x}, t)$$

(Papoulis, 1984), where v_i is the number in volume $\mathscr{V}_i$ in a particular realization. The rate of

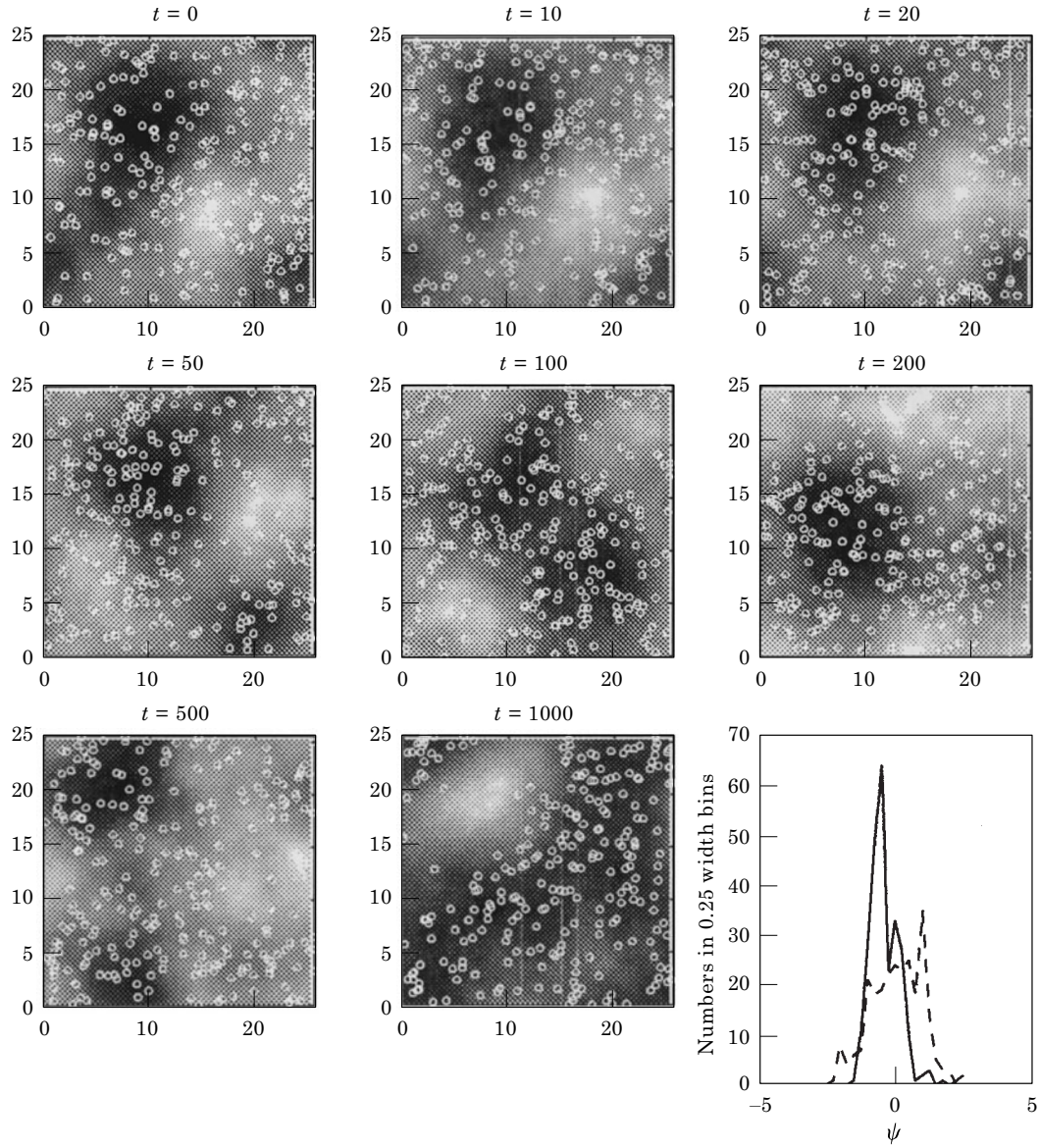


FIG. 17. Kinesis in turbulent flow.

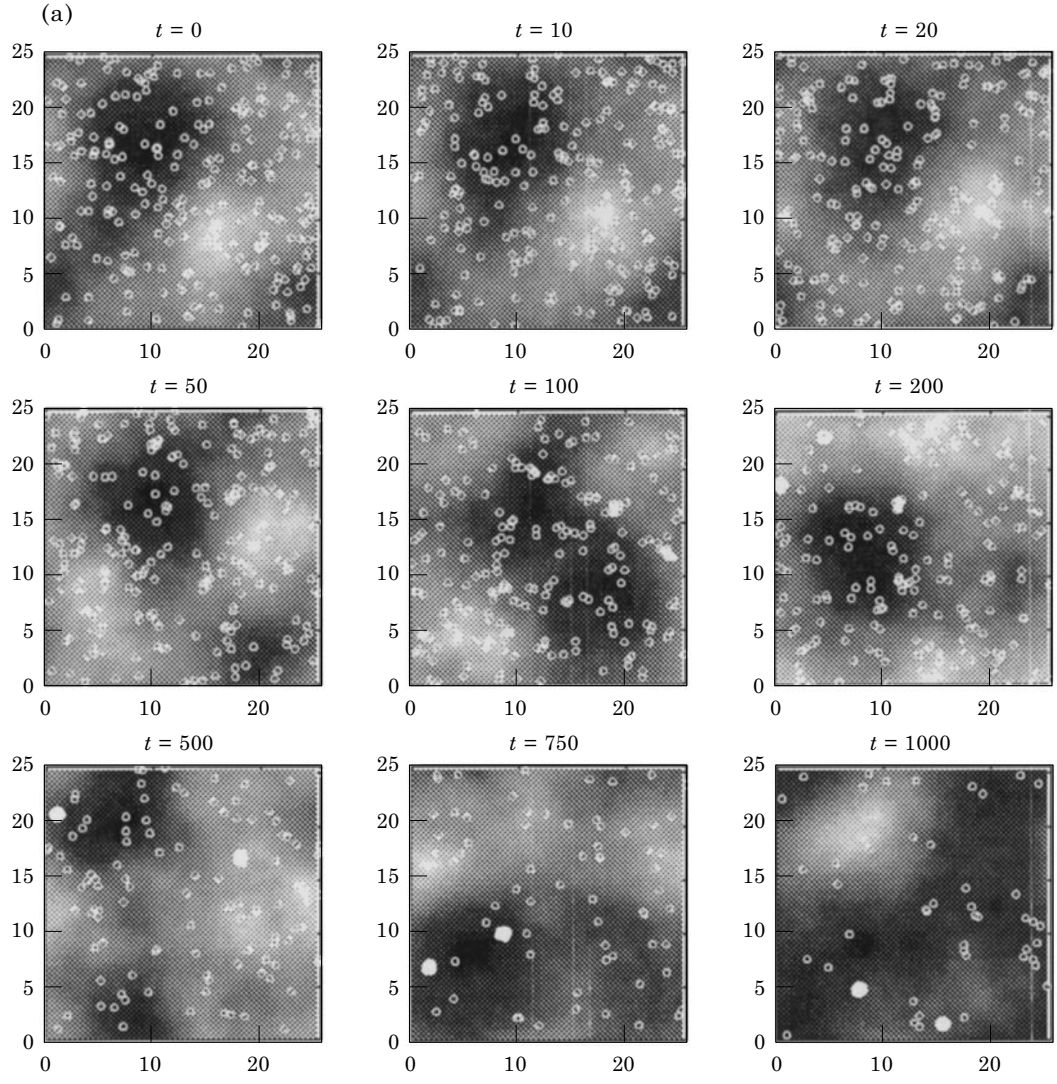


FIG. 18(a)—(*Caption on opposite page*).

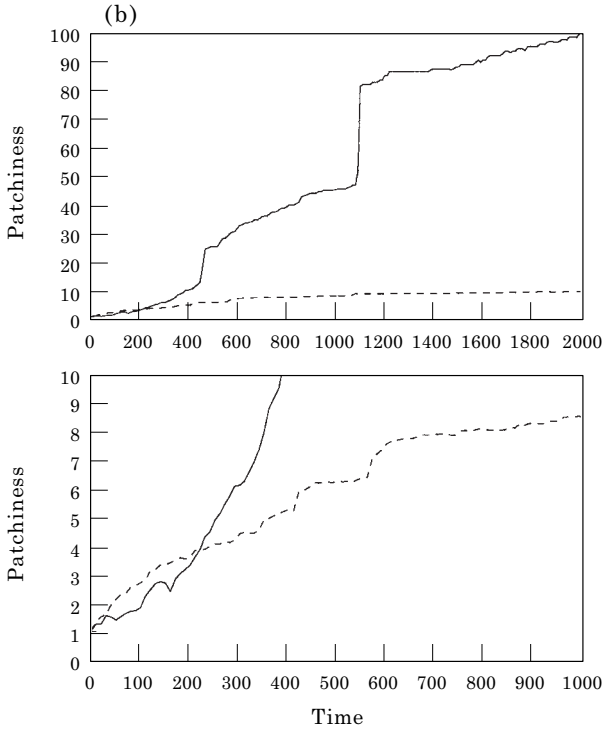
change of this expected value is given by the expected flux of animals across the boundary of the volume (and internal population dynamics changes). This leads back to the point-wise formulation (14).

But to calculate the flux, we need to compute the expected value of an animal's social responses. Let the social behavioral decision of an individual at position $\mathbf{x}$ (which may have consequences such as changes of speed or heading) be expressed as $f(\mathbf{x}, t, v_1, v_2, \ldots, v_M)$, a function of the numbers in a set of M volumes $\mathscr{V}_i$ near the organism's position. Virtually any social algorithm (at least without memory) can be written in this form (for example f can be stochastic, $M \to \infty$ can yield continuous variation in space, etc.). The decision function may be expanded in a Taylor series,

$$f(\mathbf{x}, t, v_1, v_2, \ldots, v_M) = f(\mathbf{x}, t, \bar{v}_1, \bar{v}_2, \ldots, \bar{v}_M) + v_i' \frac{\partial f}{\partial v_i} + \frac{1}{2} v_i' v_j' \frac{\partial^2 f}{\partial v_i \partial v_j} + \cdots$$

(summation convention used) where $v_i' = v_i - \bar{v}_i$ is the deviation from the average in a given realization, and the derivatives are evaluated at $\bar{v}_i$. Calculation of the density flux requires an expression for the expected value of the decision function, $\phi(\mathbf{x}, t) = E\{f\}$. In expanded form,

$$\phi(\mathbf{x}, t) = f(\mathbf{x}, t, \bar{v}_1, \bar{v}_2, \ldots, \bar{v}_M) + \frac{1}{2} E\{v_i' v_j'\} \frac{\partial^2 f}{\partial v_i \partial v_j} + \cdots. \quad (15)$$



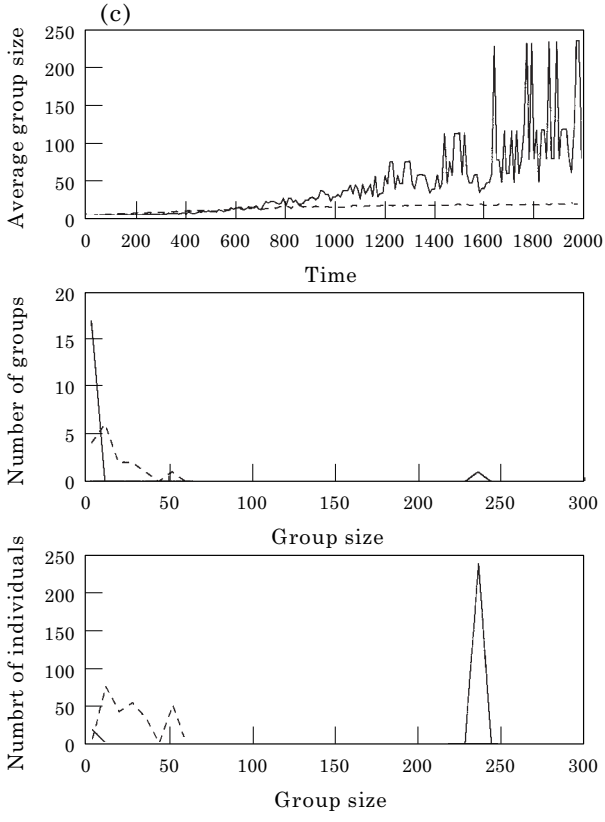


FIG. 18(b) and (c)—(*Caption opposite*).

In (15), nonlinear behavioral terms appear in the expected decision as functions of joint probability distributions between volumes and/or higher order statistics within volumes.

4.1.1. *Canonical one-parameter distributions*

We can test two canonical statistical distributions as models for how neighbors occur in simulation models, spanning the spectrum of the "randomness" present in a system. One extreme is to assume *no variance*, i.e. individuals exist in a crystallized state with little or no relative motion; thus all higher moments in (15) vanish, allowing the expected value of the behavioral decision to be estimated from the density alone. The opposite alternative is the maximally random case, i.e. that neighbors occur as Poisson points governed by the mean density distribution (Papoulis, 1984). For such a set, the probability of observing k individuals in V_i is Poisson distributed,

$$Pr(v_i = k) = e^{-\bar{v}_i} \frac{v_i^k}{k!}. \tag{16}$$

All higher moments, and thus the expected value of the social behavior, can be calculated from (16).

The stochastic processes giving rise to Poisson points are defined in part by the requirement that the number of individuals observed within non-overlapping intervals must be independent random variables. Unlike density-independent behaviors such as chemotaxis, density-dependent social behaviors give rise to distributions of individuals in which non-overlapping intervals are not independent. In fact, many social behaviors violate this requirement explicitly. For example when individuals maintain a minimum distance from neighbors, the observation of an individual at a given position implies a lower probability of occupation of immediately adjacent positions. In this case, we expect that using an assumption of Poisson points to predict

FIG. 18 (a) Social behavior in turbulent flow. (b) Patchiness with turbulent flow (—) and without (– – –). Lower figure is a magnified view of the upper. (c) Group sizes with flow (—) and without (– – –). Top: average group size; middle: number of groups of different sizes; bottom: number of individuals belonging to groups of different sizes.

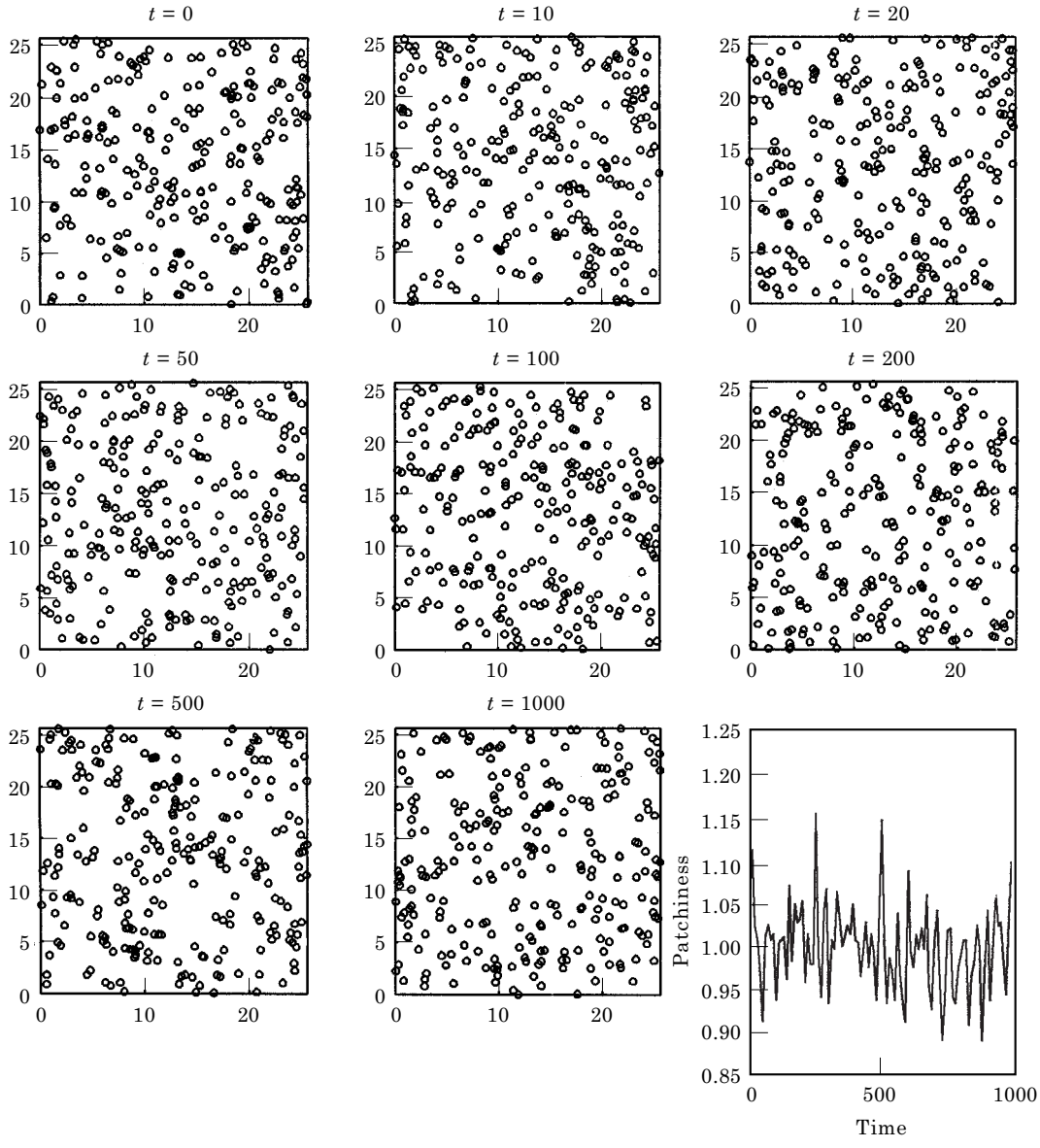


FIG. 19. Movement in the periodic domain with turbulent flow in the absence of grouping forces ($\mathscr{V} = 0$, $\alpha = 1$, $\beta = 0.2$ or $\kappa = 0.02$).

population-level movements can fail rather dramatically.

However, in other models of social behaviors, the extent to which the Poisson point assumption could apply as an approximation is less clear. For example in many social aggregation models, individuals respond to the total number of other individuals within surrounding neighborhoods. If the neighborhoods are large relative to typical spacing between individuals, and if the stochastic fluctuations in a neighborhood from one behavioral response to the next are of a sufficient magnitude, the incidence of individuals in individual-based simulation may closely resemble those generated by Poisson processes. This was found to be the case in the 1D social aggregation model of Grünbaum (1994), depending on the choice of parameters determining rates of stochastic fluctuation and behavioral adjustments to neighbors. The relevant question therefore seems not to be whether the Poisson point assumptions can be rigorously justified (for that is quite unlikely) but whether there are ways of defining local densities and utilizing

the Poisson point assumption to obtain acceptable predictions of population-level fluxes due to individual-level social behaviors.

We note that the level of stochasticity is to a large extent an "emergent phenomenon" in these types of dynamical systems: a large component of the stochastic forcing for any given individual is determined by the degree of stochasticity displayed by its neighbors, and vice versa. Behavioral conditions leading to large levels of stochasticity (and perhaps therefore validity of Poisson point assumptions) have not been sufficiently explored to allow *a priori* analytical predictions, but this would be a fruitful line of investigation.

Having made these comments, we now give an example of how the Poisson point assumption can be applied to obtain useful characteristics of the population-level fluxes.

4.1.2. *Example*

We can estimate the expected preferred velocity from the model (5) or (8) with the

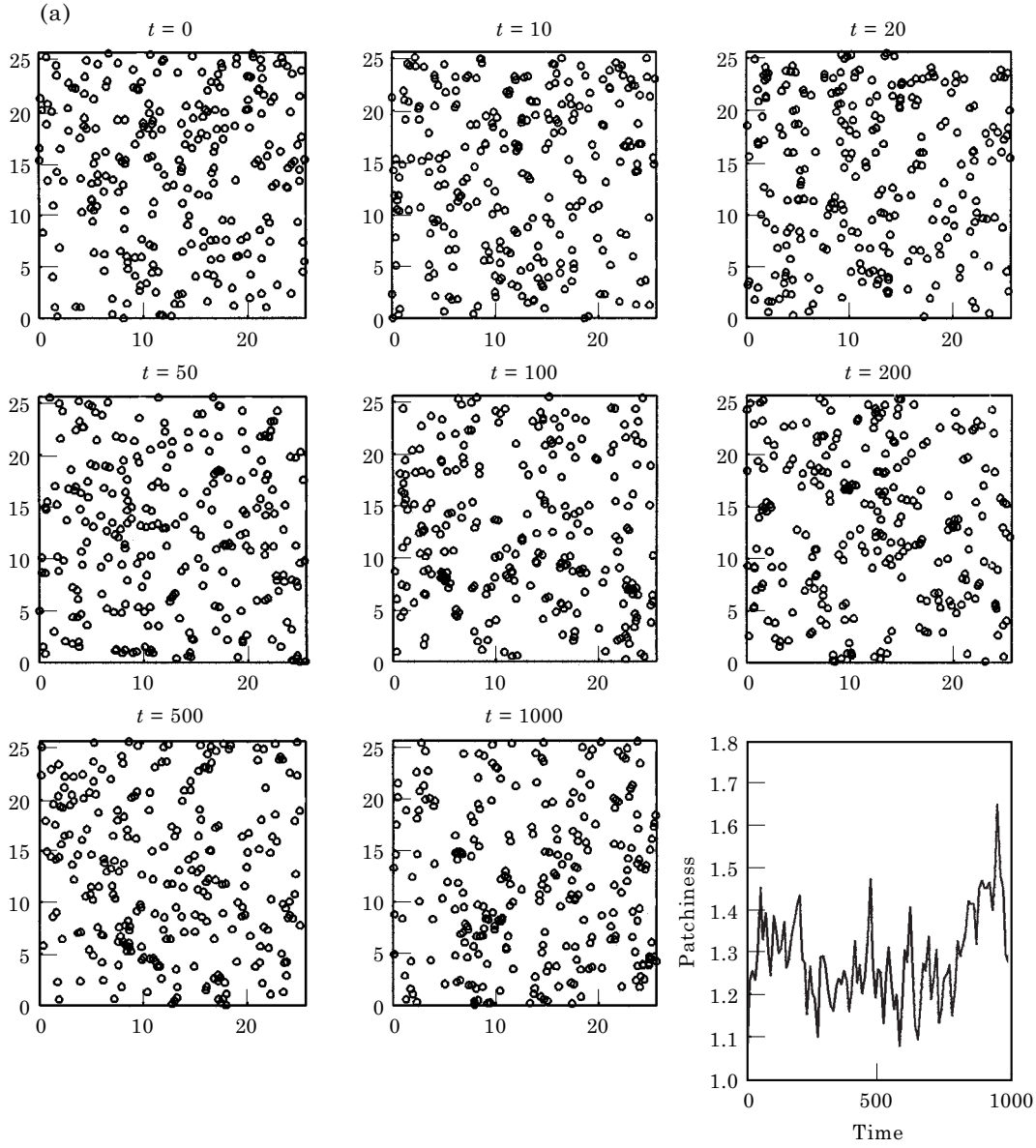


FIG. 20(a)—(*Caption on following page*).

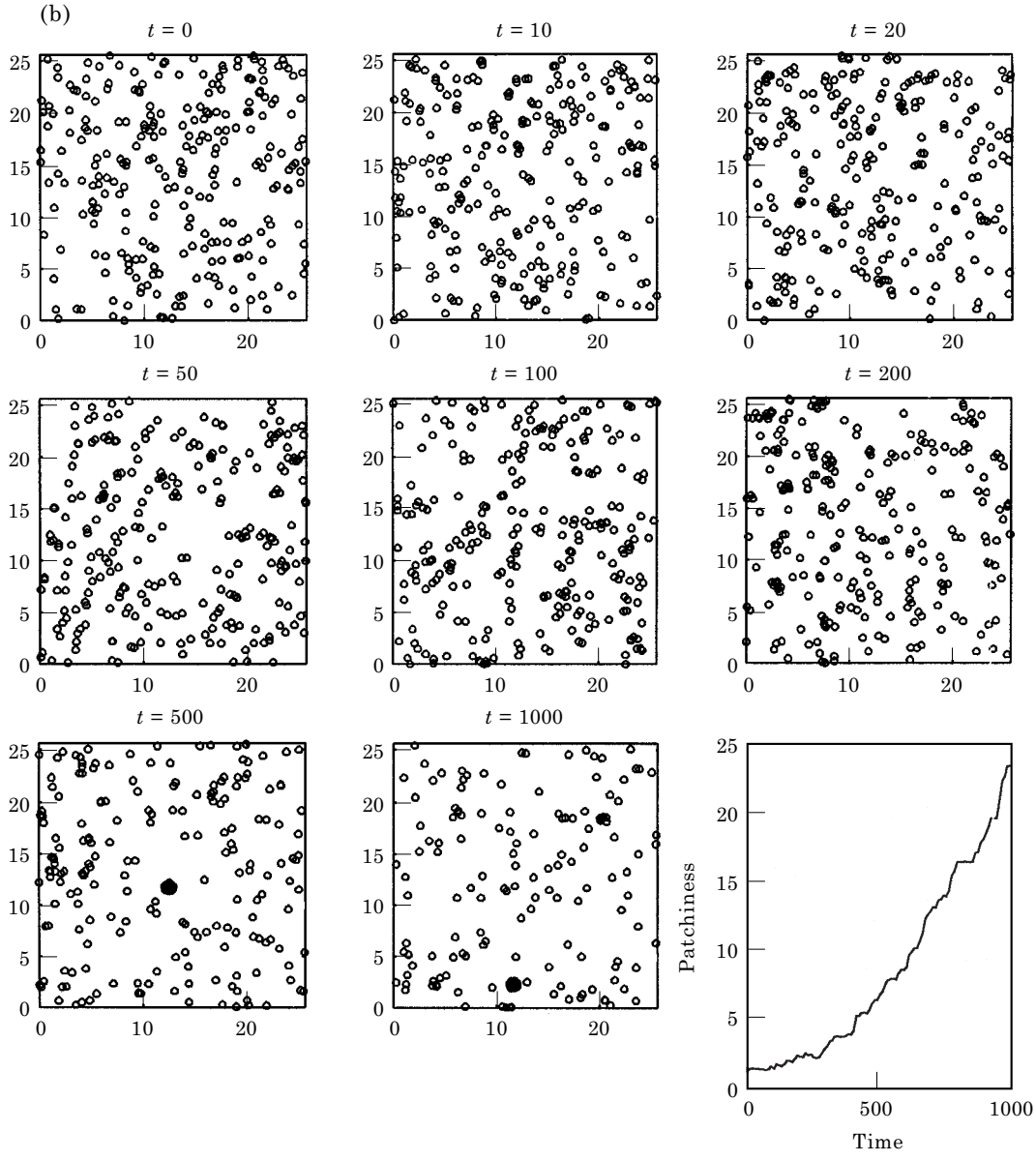


FIG. 20(b)

FIG. 20. (a) Weak grouping forces (70%) ($\mathscr{V} = 0.14$, $\alpha = 0.7$, $\beta = 0.117$ or $\kappa = 0.014$). (b) Slightly stronger grouping forces (75%) ($\mathscr{V} = 0.15$, $\alpha = 0.75$, $\beta = 0.130$ or $\kappa = 0.0150$).

assumption of uniform or Poisson distributed particles. In each case, the higher moments in (15) vanish, so that we can replace the sums by integrals and find

$$\mathbf{V}(\mathbf{x}) = \int_{|\mathbf{z}|<1} \mathrm{d}\mathbf{z}\rho(\mathbf{x}+\mathbf{z})\mathbf{W}(\mathbf{z})$$

When $\mathbf{W}$ has the form $\mathbf{W} = \mathbf{z}F(|\mathbf{z}|)$, we can rewrite it as $-\nabla\mathscr{W}(|\mathbf{z}|)$ and integrate by parts to find

$$\mathbf{V}(\mathbf{x}) = \int_{|\mathbf{z}|<1} \mathrm{d}\mathbf{z}\mathscr{W}(|\mathbf{z}|)\nabla\rho(\mathbf{x}+\mathbf{z})$$

$$= \nabla\int_{|\mathbf{z}|<1} \mathrm{d}\mathbf{z}\mathscr{W}(|\mathbf{z}|)\rho(\mathbf{x}+\mathbf{z}). \qquad (17)$$

When the density ρ varies slowly, we can approximate this by

$$\mathbf{V} = \left[\int_{|\mathbf{z}|<1} d\mathbf{z}\mathscr{W}(|\mathbf{z}|)\right]\nabla\rho$$

$$= \mathscr{V}\frac{\pi}{12}\nabla\rho$$

$$\mathscr{W} = \mathscr{V}\left(\frac{|\mathbf{z}|^2}{2} - \frac{|\mathbf{z}|^4}{4} - \frac{1}{4}\right) \quad \text{for (6)},$$

$$\mathscr{V} = \mathscr{V}\frac{11\pi}{160}\nabla\rho,$$

$$\mathscr{W} = 3.3\mathscr{V}\left(\frac{5|z|^4}{16} - \frac{|\mathbf{z}|^2}{8} - \frac{|\mathbf{z}|^6}{6} - \frac{1}{48}\right) \quad \text{for (8)},$$

showing that the preferred velocity is up-gradient.

4.2. DYNAMICS AND DENSITY DISTRIBUTIONS IN PHASE SPACE

In the theory of gases, there are three basic approaches to modeling the dynamics. The

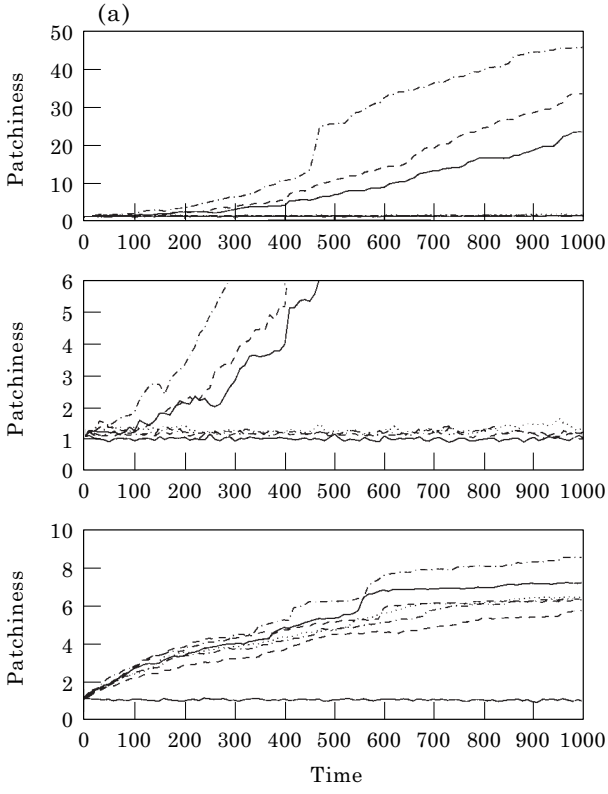
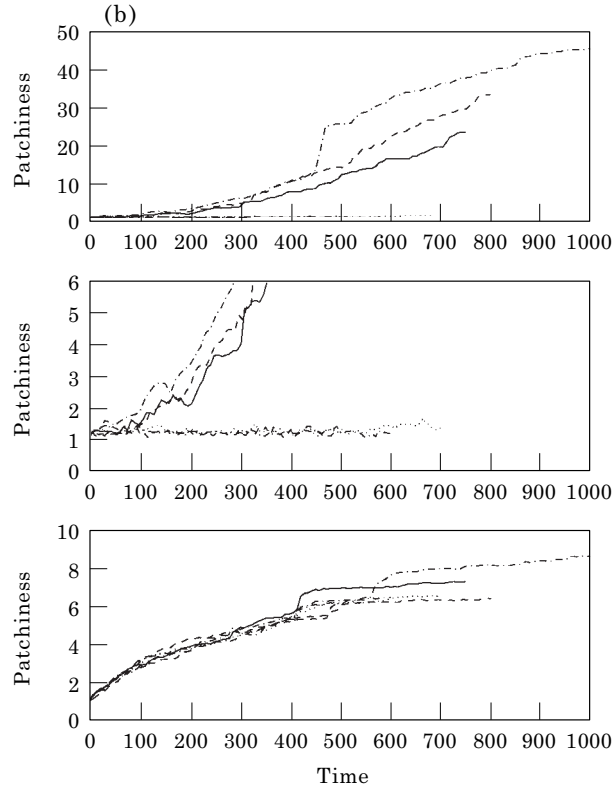


FIG. 21(a)—(*Caption opposite*).



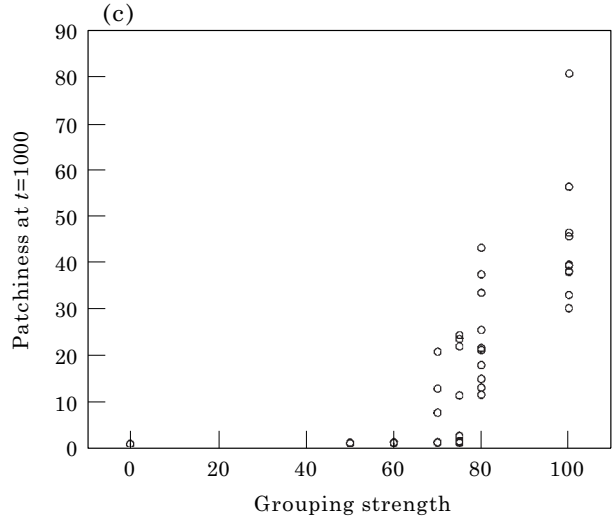


FIG. 21(b) and (c)

FIG. 21. (a) Patchiness vs. times for various grouping strengths (counting upwards: (—) = 0%, (– – –) = 50%, (–·–·–) = 60%, (. . . .) = 70%, second (—) = 75%, second (– – –) = 80%, second (–·–·–) = 100%). The second panel is a magnified version of the first; the third panel shows the case without flow. (b) Same but with time rescaled, so that the grouping strength is effectively fixed and the flow is reduced in strength: (—) = 0%, (– – –) = 50%, (–·–·–) = 60%, (. . .) = 70%, second (—) = 75%, second (– – –) = 80%, second (–·–·–) = 100%. (c) Patchiness at $t = 1000$ for 10 realizations as a function of the grouping strength.

simplest, but computationally overwhelming, approach is to simulate the motions of individual molecules by using the rules of Newtonian dynamics and representations of collision processes. This is directly equivalent to the individual-based simulations described above. At the opposite end of the spectrum are the Navier–Stokes equations that view the gas as a continuum and introduce viscosity and diffusion to incorporate effects of momentum and energy transfer via collisions. In between these two approaches lies the Boltzmann equation, which predicts the evolution in space and time of the probability distribution for the molecular velocities. This equation proves crucial in that it forges a direct link between molecular dynamics and continuum mechanics and makes possible calculation of the form and coefficients of the viscous/diffusive terms in the Navier–Stokes equations, as well as an approach to problems where corrections to these are needed. An analogous equation for organisms should play an equivalent and important role in aiding our understanding of the relationship between the continuum and individual-based models and the limits of validity of the former.

Thus, we identify several possible stages in developing a continuum model for the distribution of organisms. If we begin with the dynamics of the individual as outlined above, we can then develop an equation for the spatial/velocity distribution $n(\mathbf{x}, \mathbf{v}, t)$ where $n(\mathbf{x}, \mathbf{v}, t)\, d\mathbf{x}\, d\mathbf{v}/N$ is the probability that an individual is in the domain of size $d\mathbf{x}$ around $\mathbf{x}$ and has velocity in the range $d\mathbf{v}$ around $\mathbf{v}$. N is the total number of organisms. This is the Boltzmann equation. But if we wish to have a closed expression for the expected density of organisms, $\rho(\mathbf{x}, t) \equiv \int d\mathbf{v} n(\mathbf{x}, \mathbf{v}, t)$, we must reduce the equation further; this can be done using either a Fokker–Planck differential equation version of the Boltzmann equation or it can be done directly from the integral equation. In gas dynamics, the latter reduction procedure was formalized by Chapman (1916) and Enskog (1917).

We shall now discuss procedures for deriving continuum models from our individual-based grouping models, beginning with an examination of the equation for the evolution of the spatial/velocity density $n(\mathbf{x}, \mathbf{v}, t)$. Let $\mathscr{P}_{\delta \mathbf{V}}(\mathbf{z}; \mathbf{x}', \mathbf{v}', t)$ be the probability that a particle at position $\mathbf{x}'$ with velocity $\mathbf{v}'$ at time t has a random acceleration $\delta \mathbf{V} = \mathbf{z}$. Define a similar function $\mathscr{P}_{\delta \mathbf{X}}(\mathbf{z}; \mathbf{x}', \mathbf{v}', t)$ for the random hops. Then we find

$$n(\mathbf{x}, \mathbf{v}, t + \delta t) = \int d\mathbf{x}'\, d\mathbf{v}' \mathscr{P}_{\delta \mathbf{X}}(\mathbf{x} - \mathbf{x}' - \mathbf{v}'\delta t; \mathbf{x}', \mathbf{v}', t) \mathscr{P}_{\delta \mathbf{V}}(\mathbf{v} - \mathbf{v}' - \mathbf{a}\delta t; \mathbf{x}', \mathbf{v}', t) n(\mathbf{x}', \mathbf{v}', t) \tag{18}$$

under the assumption that the accelerations $\delta \mathbf{V}$ and hops $\delta \mathbf{X}$ are uncorrelated. The density is determined by the $\mathbf{v}$ integral of this equation:

$$\rho(\mathbf{x}, t + \delta t) = \int d\mathbf{x}'\, d\mathbf{v}' \mathscr{P}_{\delta \mathbf{X}}(\mathbf{x} - \mathbf{x}' - \mathbf{v}'\delta t; \mathbf{x}', \mathbf{v}', t) n(\mathbf{x}', \mathbf{v}', t). \tag{19}$$

4.2.1. *Fokker–Planck simplification*

The Fokker–Planck approach (cf. Chandrasekhar, 1943) can be used to simplify integrals such as (18) when one or both probability functions are sharply peaked. In particular, consider an integral

$$\begin{aligned} I(\mathbf{x}, \mathbf{v}', t) &= \int d\mathbf{x}' \mathscr{P}_{\delta \mathbf{X}}(\mathbf{x} - \mathbf{x}' - \mathbf{v}'\delta t; \mathbf{x}', \mathbf{v}', t) \phi(\mathbf{x}', \mathbf{v}', t) \\ &= \epsilon \int d\mathbf{z} \mathscr{P}_{\delta \mathbf{X}}(\epsilon \mathbf{z}; \mathbf{x} - \mathbf{v}'\delta t - \epsilon \mathbf{z}, \mathbf{v}', t) \phi(\mathbf{x} - \mathbf{v}'\delta t - \epsilon \mathbf{z}, \mathbf{v}', t) \end{aligned} \tag{20}$$

We assume $\mathscr{P}_{\delta \mathbf{X}}$ varies rapidly with the first argument (i.e. has order one variation as $\mathbf{z}$ varies), while the variation in the second argument and in ϕ are presumed to be slow—the

ratio of scales of variation being ϵ. Taylor-expanding with $\epsilon \sim \delta t^{1/2}$ leads to

$$I \simeq \epsilon \int d\mathbf{z} \mathscr{P}_{\delta \mathbf{X}}(\epsilon \mathbf{z}; \mathbf{x}, \mathbf{v}', t) \phi(\mathbf{x}, \mathbf{v}', t)$$

$$- \epsilon^2 \frac{\partial}{\partial x_i} \int d\mathbf{z} z_i \mathscr{P}_{\delta \mathbf{X}}(\epsilon \mathbf{z}; \mathbf{x}, \mathbf{v}', t) \phi(\mathbf{x}, \mathbf{v}', t)$$

$$+ \frac{\epsilon^3}{2} \frac{\partial^2}{\partial x_i \partial x_j} \int d\mathbf{z} z_i z_j \mathscr{P}_{\delta \mathbf{X}}(\epsilon \mathbf{z}; \mathbf{x}, \mathbf{v}', t) \phi(\mathbf{x}, \mathbf{v}', t)$$

$$- \epsilon \delta t \frac{\partial}{\partial x_i} v_i' \int d\mathbf{z} \mathscr{P}_{\delta \mathbf{X}}(\epsilon \mathbf{z}; \mathbf{x}, \mathbf{v}', t) \phi(\mathbf{x}, \mathbf{v}', t)$$

We can now perform the $\mathbf{z}$ integrals and arrive at

$$I \simeq \phi(\mathbf{x}, \mathbf{v}', t) + \delta t \frac{1}{2} \frac{\partial^2}{\partial x_i \partial x_j} [\sigma_{ij}(\mathbf{x}, \mathbf{v}', t) \phi(\mathbf{x}, \mathbf{v}', t)]$$

$$- \delta t \frac{\partial}{\partial x_i} [v_i' \phi(\mathbf{x}, \mathbf{v}', t)] \tag{21}$$

with

$$\sigma_{ij}(\mathbf{x}, \mathbf{v}', t) \equiv \frac{1}{\delta t} \int d\mathbf{Z} Z_i Z_j \mathscr{P}_{\delta \mathbf{X}}(\mathbf{Z}; \mathbf{x}, \mathbf{v}', t)$$

The assumption on the statistics of $\delta \mathbf{X}$ is that they have zero mean and the variance remains of order δt in the appropriate limit.

4.2.2. *Hopping movements*

The common reduction (cf. van Kampen, 1992, or Okubo, 1980, 1986) deals with situations in which the particle does not retain memory of the initial velocity, either because the damping is rapid, the time steps of interest are long compared with velocity decorrelation times, or hops dominate. In that case, we can apply (21) to (19) to find

$$\rho(\mathbf{x}, t + \delta t) = \rho(\mathbf{x}, t)$$

$$+ \delta t \frac{1}{2} \frac{\partial^2}{\partial x_i \partial x_j} \left[\int d\mathbf{v}' \sigma_{ij}(\mathbf{x}, \mathbf{v}', t) n(\mathbf{x}, \mathbf{v}', t) \right]$$

$$- \delta t \frac{\partial}{\partial x_i} \left[\int d\mathbf{v}' v_i' n(\mathbf{x}, \mathbf{v}', t) \right]$$

If $\mathscr{P}_{\delta \mathbf{X}}$ is independent of the initial velocities, then σ_{ij} is also, and the second term can be simplified. Likewise, we can assume that the expected velocities are the preferred value $\mathbf{V}(\mathbf{x}, t)$ so that $n(\mathbf{x}, \mathbf{v}, t) = \rho(\mathbf{x}, t) \delta(\mathbf{v} - \mathbf{V})$. In that case, we obtain the simple expression

$$\rho(\mathbf{x}, t + \delta t) = \rho(\mathbf{x}, t)$$

$$+ \delta t \frac{1}{2} \frac{\partial^2}{\partial x_i \partial x_j} [\sigma_{ij}(\mathbf{x}, t) \rho(\mathbf{x}, t)]$$

$$- \delta t \frac{\partial}{\partial x_i} [V_i \rho(\mathbf{x}, t)]$$

or

$$\frac{\partial}{\partial t} \rho(\mathbf{x}, t) = \frac{1}{2} \frac{\partial^2}{\partial x_i \partial x_j} [\sigma_{ij}(\mathbf{x}, t) \rho(\mathbf{x}, t)]$$

$$- \frac{\partial}{\partial x_i} [V_i \rho(\mathbf{x}, t)] \tag{22}$$

Thus the density flux is given by

$$J_i = u_i \rho(\mathbf{x}, t) - \kappa_{ij} \frac{\partial}{\partial x_j} \rho(\mathbf{x}, t) \tag{23}$$

with an advective velocity

$$u_i = V_i - \frac{1}{2} \frac{\partial}{\partial x_j} \sigma_{ij} \tag{23}$$

which can very well be convergent (and may depend on ρ). The diffusivity is

$$\kappa_{ij} = \frac{1}{2} \sigma_{ij} \tag{23}$$

It is worth noting that even for a Lagrangian model in which individual movement is governed by a nonlinear equation, the resulting equation for the probability density function (22) is linear in ρ. The nonlinearity is parametrically disguised in terms of the coefficients $\mathbf{V}$ and σ. We shall treat density dependence with the Fokker–Planck formalism by introducing a dependence of the phase transition probability on the animal density at the points of departure. (Okubo, 1986, shows that making the transition probability depend on the point of arrival leads to equations with the moments outside the derivatives.)

4.2.3. *Boltzmann equation*

When the stochastic term $\delta \mathbf{X}$ in the position equation (1a) is missing, so that the only random

variability is in the forces, the transition probability can be simplified

$$\mathscr{P}_{\delta \mathbf{X}}(\mathbf{x}-\mathbf{x}'-\mathbf{v}'\delta t;\mathbf{x}',\mathbf{v}',t)=\delta(\mathbf{x}-\mathbf{x}'-\mathbf{v}'\delta t)$$

and therefore $\sigma_{ij}=0$. Now applying the Fokker–Planck expansion procedure (21) to the original eqn (18) gives

$$n(\mathbf{x},\mathbf{v},t+\delta t)=\int d\mathbf{v}'\mathscr{P}_{\delta \mathbf{V}}(\mathbf{v}-\mathbf{v}'-\mathbf{a}\delta t;$$

$$\mathbf{x},\mathbf{v}',t)n(\mathbf{x},\mathbf{v}',t)-\delta t\frac{\partial}{\partial x_i}\int d\mathbf{v}'v_i'\mathscr{P}_{\delta \mathbf{V}}(\mathbf{v}-\mathbf{v}'$$

$$-\mathbf{a}\delta t:\mathbf{x},\mathbf{v}',t)n(\mathbf{x},\mathbf{v}',t)$$

For a peaked $\mathscr{P}_{\delta \mathbf{V}}$, the last term can be replaced to $O(\delta t)$ by

$$-\delta t\frac{\partial}{\partial x_i}v_i n(\mathbf{x},\mathbf{v},t)$$

and we can expand the left-hand side to obtain the Boltzmann equation

$$\frac{\partial}{\partial t}n(\mathbf{x},\mathbf{v},t)+\nabla\cdot[\mathbf{v}n(\mathbf{x},\mathbf{v},t)$$

$$=\int d\mathbf{v}'\psi(\mathbf{v};\mathbf{x},\mathbf{v}',t)n(\mathbf{x},\mathbf{v}',t) \quad (24a)$$

with the transition rate ψ given by

$$\psi(\mathbf{v};\mathbf{x},\mathbf{v}',t)=\frac{1}{\delta t}[\mathscr{P}_{\delta \mathbf{V}}(\mathbf{v}-\mathbf{v}'-\mathbf{a}\delta t;\mathbf{x},\mathbf{v}',t)$$

$$-\delta(\mathbf{v}-\mathbf{v}')]. \quad (24b)$$

This is the integral form of the Boltzmann equation and is applicable if our time steps are large enough that the kernel is well defined.

From the definition of ψ, it is clear that

$$\int d\mathbf{v}\psi(\mathbf{v};\mathbf{x},\mathbf{v}',t)=0 \quad (25)$$

so that the $\mathbf{v}$ integral of the Boltzmann equation (24) will lead to the density equation (14) with the flux being

$$\mathbf{J}=\int d\mathbf{v}\,\mathbf{v}n(\mathbf{x},\mathbf{v},t). \quad (26)$$

Note that the transition rate does not appear directly in the equation for the time rate of change of the density; thus we have two time scales, one for the readjustment of the velocity distribution (involving the rate determined by ψ) and one for spatial grouping, involving the characteristic velocities and spatial scales.

4.2.4. *General Fokker–Plank reduction*

If we consider our equation including both random hops and accelerations (under the assumption that $\delta\mathbf{V}$ and $\delta\mathbf{X}$ are uncorrelated), we can apply the reduction procedure (21) twice to show that

$$\frac{\partial}{\partial t}n(\mathbf{x},\mathbf{v},t)=-\frac{\partial}{\partial x_i}[v_i n(\mathbf{x},\mathbf{v},t)]$$

$$-\frac{\partial}{\partial v_i}[a_i n(\mathbf{x},\mathbf{v},t)]$$

$$+\frac{1}{2}\frac{\partial^2}{\partial x_i\partial \mathrm{x}_j}[\sigma_{ij}n(\mathbf{x},\mathbf{v},t)]$$

$$+\frac{1}{2}\frac{\partial^2}{\partial v_i\partial v_j}[\gamma_{ij}n(\mathbf{x},\mathbf{v},t)]$$

where γ is defined in terms of the second moments of $\delta\mathbf{V}$.

$$\gamma_{ij}(\mathbf{x},\mathbf{v},t)\equiv\frac{1}{\delta t}\int d\mathbf{Z}Z_iZ_j\mathscr{P}_{\delta \mathbf{V}}(\mathbf{Z};\mathbf{x},\mathbf{v},t)$$

When random hops do not occur, we obtain the differential form of the Boltzmann equation

$$\frac{\partial}{\partial t}n(\mathbf{x},\mathbf{v},t)=-\frac{\partial}{\partial x_i}[v_i n(\mathbf{x},\mathbf{v},t)$$

$$-\frac{\partial}{\partial v_i}[a_i n(\mathbf{x},\mathbf{v},t)]$$

$$+\frac{1}{2}\frac{\partial^2}{\partial v_i\partial v_j}[\gamma_{ij}n(\mathbf{x},\mathbf{v},t)]. \quad (27)$$

Each of the terms on the right-hand side has a straightforward interpretation: density in a particular region and velocity band changes by (1) spatial fluxes associated with advection, by (2) changes in velocity from the deterministic forces, and by (3) diffusion of velocity because of the stochastic forces (although the latter is not strictly a diffusion, since it appears within the two derivatives).

4.3. DENSITY FLUX ESTIMATES

We can use either the integral or differential forms of the Boltzmann equation to parameterize the density flux in the case where the animals are sensitive only to the vector distances to neighbors but not to their velocity (no arrayal forces). In that case, we may expect that the forces depend only on ρ and it would make sense to try and represent $\mathbf{J}$ in terms of the density.

4.3.1. *Moment method*

In our example, where

$$\mathbf{a} = -\alpha(\mathbf{x}, \rho, t)[\mathbf{v} - \mathbf{V}(\mathbf{x}, \rho, t)],$$

$$\gamma_{ij} = \delta_{ij}\beta^2(\mathbf{x}, \rho, t),$$

we can take moments of the differential equation (27):

$$\frac{\partial}{\partial t}\rho(\mathbf{x}, t) = -\frac{\partial}{\partial x_i}\int d\mathbf{v}[v_i n(\mathbf{x}, \mathbf{v}, t)]$$

$$= -\frac{\partial}{\partial x_i} J_i \quad (28a)$$

$$\frac{\partial}{\partial t} J_i = -\frac{\partial}{\partial x_j}\int d\mathbf{v}[v_i v_j n] - \alpha J_i + \alpha V_i \rho. \quad (28b)$$

The second moment is given by

$$\frac{\partial}{\partial t}\int d\mathbf{v}[v_i v_j n] = -\frac{\partial}{\partial x_j}\int d\mathbf{v}[v_i v_j v_k n]$$

$$-2\alpha\int d\mathbf{v}[v_i v_j n] + \alpha(V_i J_j + V_j J_i) + \gamma_{ij}\rho \quad (28c)$$

(using the symmetry of γ_{ij}). When the biological time-scales are short, so that α and γ are large, the density flux relaxes rapidly to the steady-state solution of (28b)

$$J_i \to V_i\rho - \frac{1}{\alpha}\frac{\partial}{\partial x_j}\int d\mathbf{v}[v_i v_j n] \quad (28d)$$

and, when the preferred velocity $\mathbf{V}$ is small, we can write the second moment in terms of the density

$$\int d\mathbf{v}[v_i v_j n] = \frac{1}{2\alpha}\gamma_{ij}\rho \quad (28e)$$

and thereby close the flux expression

$$J_i = V_i\rho - \frac{1}{\alpha}\frac{\partial}{\partial x_j}\frac{1}{2\alpha}\gamma_{ij}\rho \quad (28f)$$

corresponding to an advection velocity

$$u_i = V_i - \frac{1}{\alpha}\frac{\partial}{\partial x_j}\frac{1}{2\alpha}\gamma_{ij} \quad (28g)$$

and a diffusivity

$$\kappa_{ij} = \frac{1}{2\alpha^2}\gamma_{ij}. \quad (28h)$$

To estimate $\mathbf{V}$ and/or β, we replace the sums in (5), (10) by integrals over the density:

$$\mathbf{V}(\mathbf{x}) = \int d\mathbf{z}\mathbf{W}(\mathbf{z})\rho(\mathbf{x} + \mathbf{z})$$

$$= \nabla\int d\mathbf{z}\mathscr{W}(|\mathbf{z}|)\rho(\mathbf{x} + \mathbf{z}) \quad (29a)$$

and

$$\beta = \beta_0 \exp\left(-\int d\mathbf{z} W_b(\mathbf{z})\rho(\mathbf{x} + \mathbf{z})\right) \quad (29b)$$

or, when α is constant,

$$\kappa = \kappa_0 \exp\left(-2\int d\mathbf{z} W_b(\mathbf{z})\rho(\mathbf{x} + \mathbf{z})\right)$$

$$= \kappa_0 \exp\left(-\int d\mathbf{z} W_\kappa(\mathbf{z})\rho(\mathbf{x} + \mathbf{z})\right). \quad (29c)$$

The formulae (29), together with the expression for the flux

$$\mathbf{J} = \mathbf{V}\rho - \frac{1}{\alpha}\nabla\frac{\beta^2}{2\alpha}\rho$$

$$= \mathbf{V}\rho - \frac{1}{\alpha}\nabla(\alpha\kappa\rho)$$

$$= \left[\mathbf{V} - \frac{1}{\alpha}\nabla(\kappa\alpha)\right]\rho - \kappa\nabla\rho$$

complete the link between the microscopic behavioral rules and the macroscopic density equation.

Some sample results, shown in Fig. 22, indicate that the continuous model reproduces features of the discrete model quite well. We had

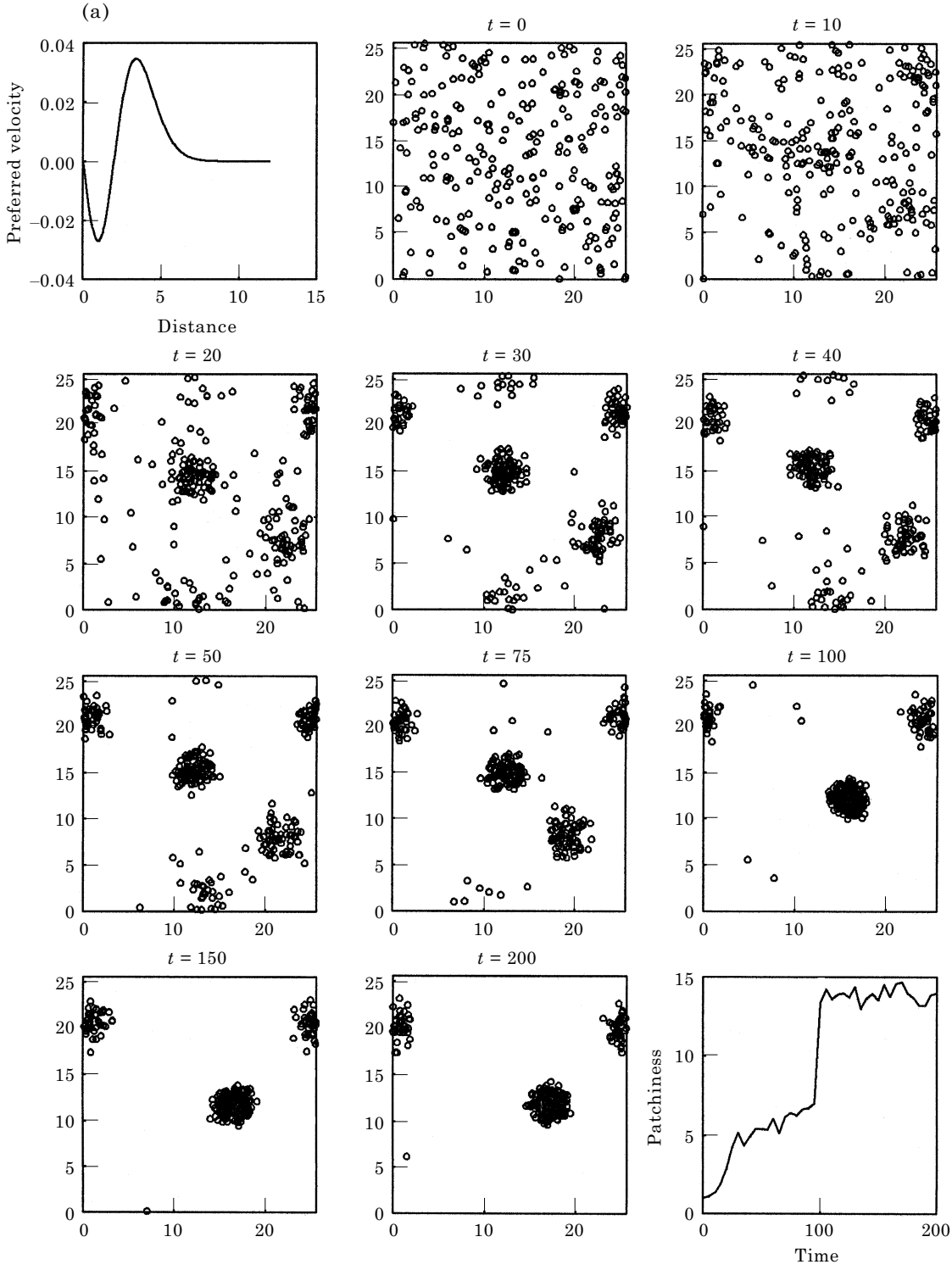


FIG. 22(a)—(*Caption on p*. 439)

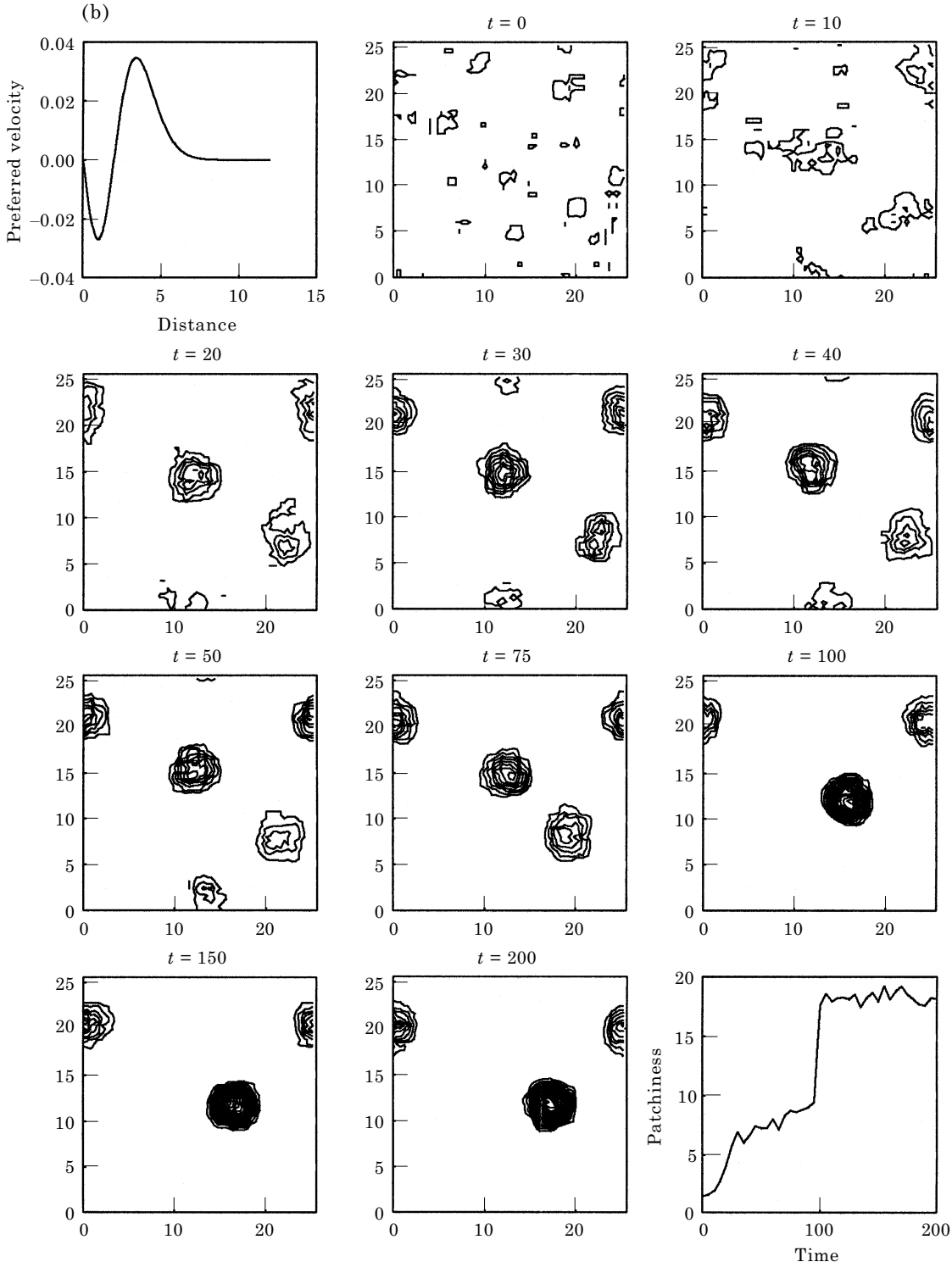


FIG. 22(b)—(*Caption on p.* 439)

to increase the perception radius to 4 rather than 1 to have adequate spatial resolution with a 64×64 pseudospectral model (with a grid spacing of 0.4)—the equations are rather stiff and numerical problems can arise as the density approaches zero. In addition, for ease of calculation, we have replaced the weighting function $\mathscr{W}_2$ with the sum of two Gaussians,

since their Fourier transform is easily defined [cf. Fig. 22(c)]

The major difference between the continuum model and the individual-based simulation is the rate of merger during the last half of the run. For some parameters, merger appears to occur earlier in the continuum model, while for others it occurs later. The details of a group merger

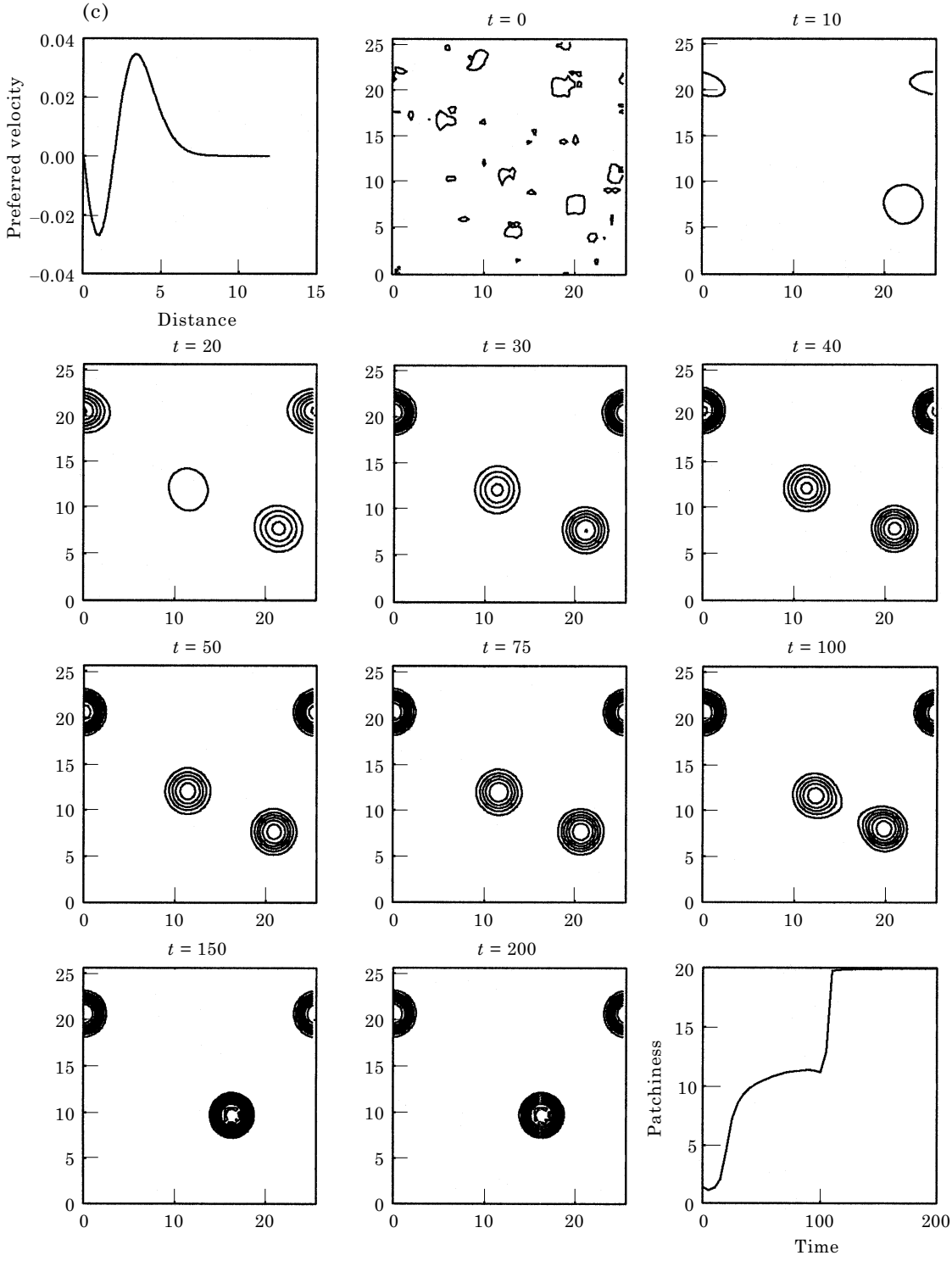


FIG. 22(c)—(*Caption on p*. 439)

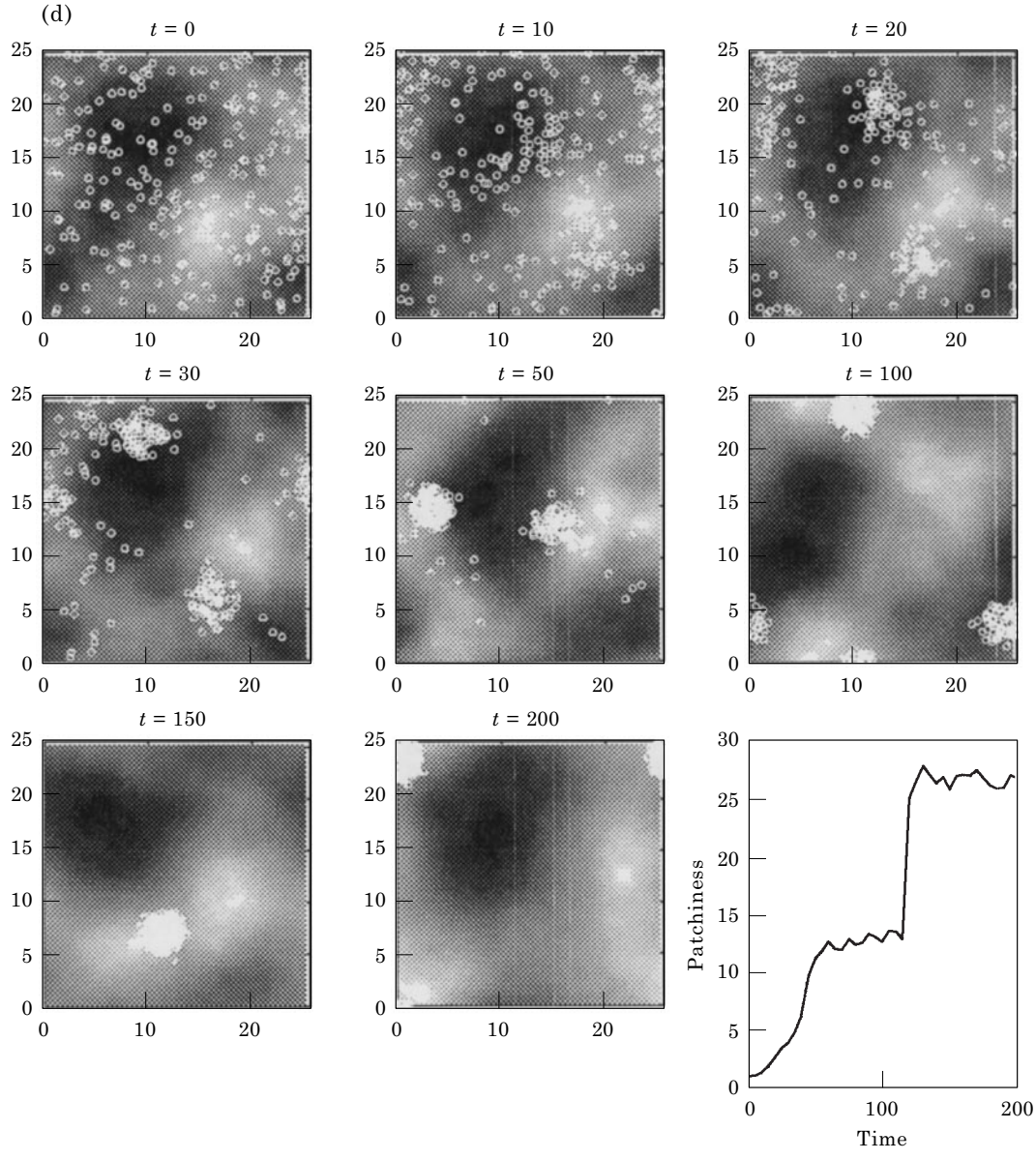


FIG. 22(d)—(*Caption on p.* 439)

event are shown in Fig. 23. The differences in merger events appear to be associated with the more random movement of the groups in the individual-based model, which enhances their encounter rate; on the other hand, the groups in the continuum model appear to have longer tails (with very low, but non-zero, densities, perhaps associated with the longer range of the Gaussians) so that the groups are always interacting, albeit weakly.

4.3.2. *Chapman–Enskog method*

In the alternative approach to deriving density fluxes, we deal directly with the integral equation (24) [although the same procedure can actually be used on the differential equation (27)]. Consider the case where the conditions are spatially uniform. Then

$$\frac{\partial}{\partial t} n(\mathbf{v}, t) = \int d\mathbf{v}' \psi(\mathbf{v}; \mathbf{v}', t) n(\mathbf{v}', t)$$

and we assume that the velocity decays to an equilibrium distribution. The rate of decay depends upon the eigenvalues λ of the integral operator

$$\int d\mathbf{v}'\psi(\mathbf{v};\mathbf{v}',t)e_\lambda(\mathbf{v}',t) = -\lambda e_\lambda(\mathbf{v},t)$$

All of the eigenvalues and eigenvectors may depend parametrically upon space and animal density. Note that there is one zero eigenvalue [since the constant 1 is a left eigenfunction solution by (25)]. For the Boltzmann equation to be reducible, all the other eigenvalues must be negative. [Note—in the gas problem, the zero eigenvalue is multiply degenerate (Harris, 1971); we shall assume for the biological dynamics that it is not.] The basic assumption is that the other eigenvalues are large: the velocity equilibrates rapidly compared with the group formation process. In effect, then, the velocity distribution remains nearly in local equilibrium. But the equilibrium distributions vary from place to place. Advection then causes weak disequilibria

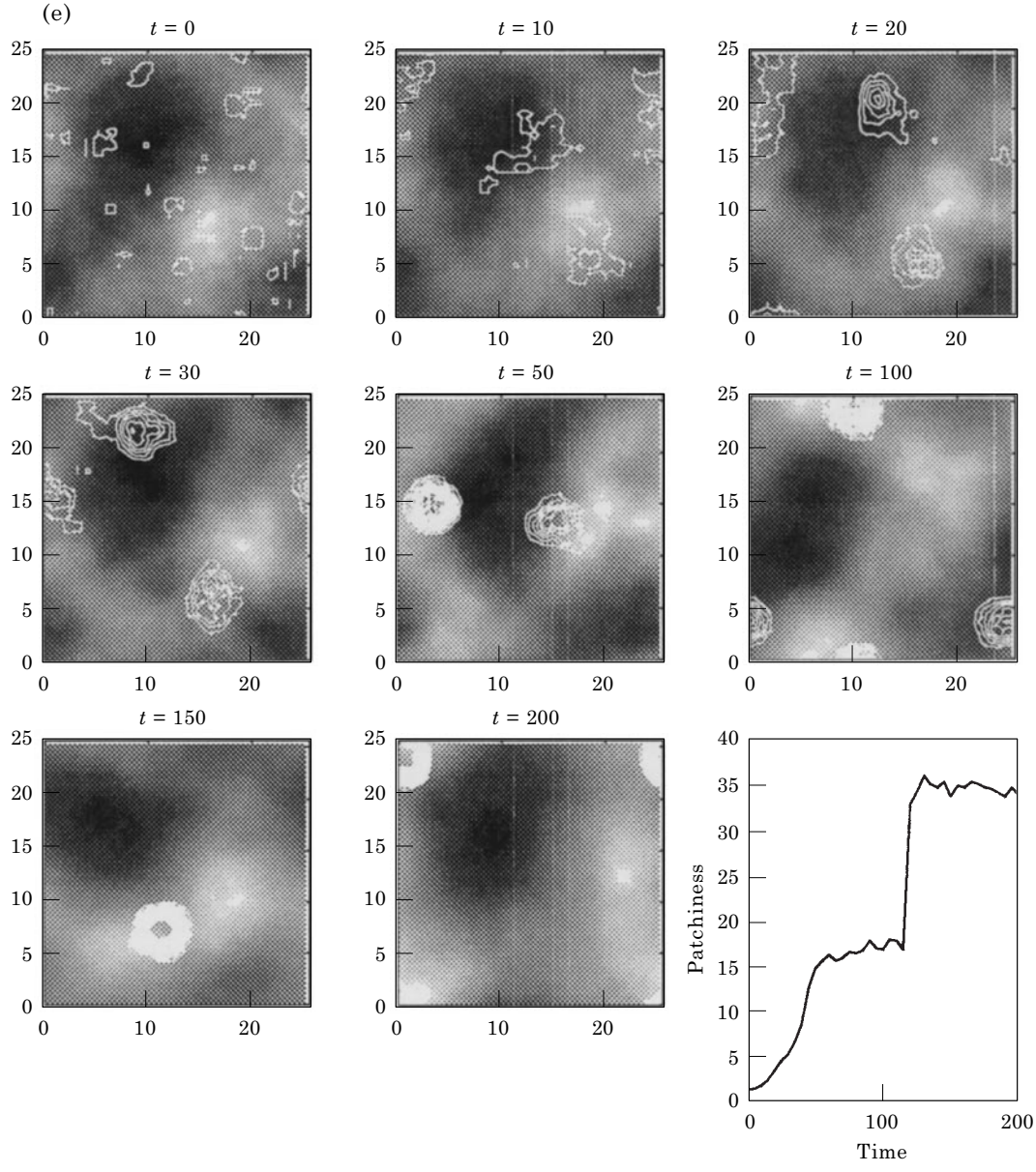


FIG. 22(e)—(*Caption opposite page*).

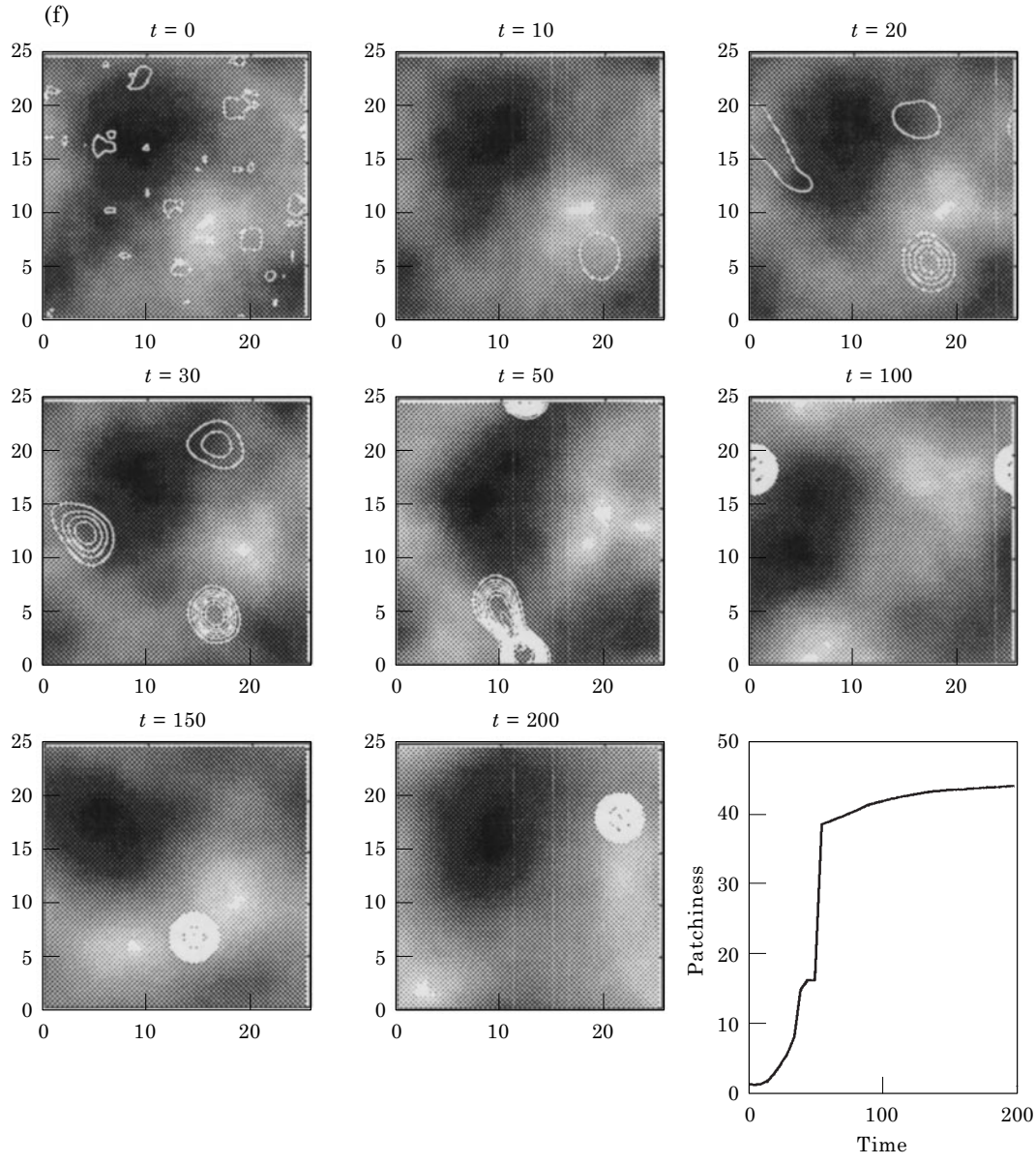


FIG. 22(f)

FIG. 22. (a) Individual-based grouping calculation using a model to (6), but with a radius of influence 4 and longer range sensing $\mathbf{W}_3 = -\nabla[0.28\exp(-\mathbf{z}^2/8) - 0.2277\exp(-\mathbf{z}^2/4)]$. The first frame shows the weighting function $\mathbf{W}_3$. The last frame shows the patchiness (using squares with side 2). Parameters are $\alpha = 1$, $\beta = 1$. (b) Density computed from the previous figure. At each of the 64×64 grid points, we count the number of individuals within a square with side length 2 centered on the gridpoint; then $\rho = N/4$. The last frame shows the patchiness estimated from

$$\tilde{p} = \frac{\iint d\mathbf{x}\rho(\mathbf{x})\iint_{\mathbf{z}\in(-1\ldots1,-1\ldots1)} d\mathbf{z}\rho(\mathbf{x}+\mathbf{z})}{4\,\frac{1}{\text{Area}}\left[\iint d\mathbf{x}\rho(\mathbf{x})\right]^2}$$

While the values are somewhat different from those in Fig. 22(a), the pattern is very similar. (c) Continuum model with the weighting function $\mathbf{W}_3$ shown in the first frame. The contour intervals are the same. The initial conditions are from the $t = 0$ frame in 22(b). The final frame shows the patchiness index; it shows different development in the early stages, but similar growth rates as the patches develop. The continuum model does not show the merger events occurring as frequently. (d) Individual-based grouping calculation in the presence of turbulent flow. (e) Density computed from the previous figure. (f) Continuum model for grouping in turbulent flow.

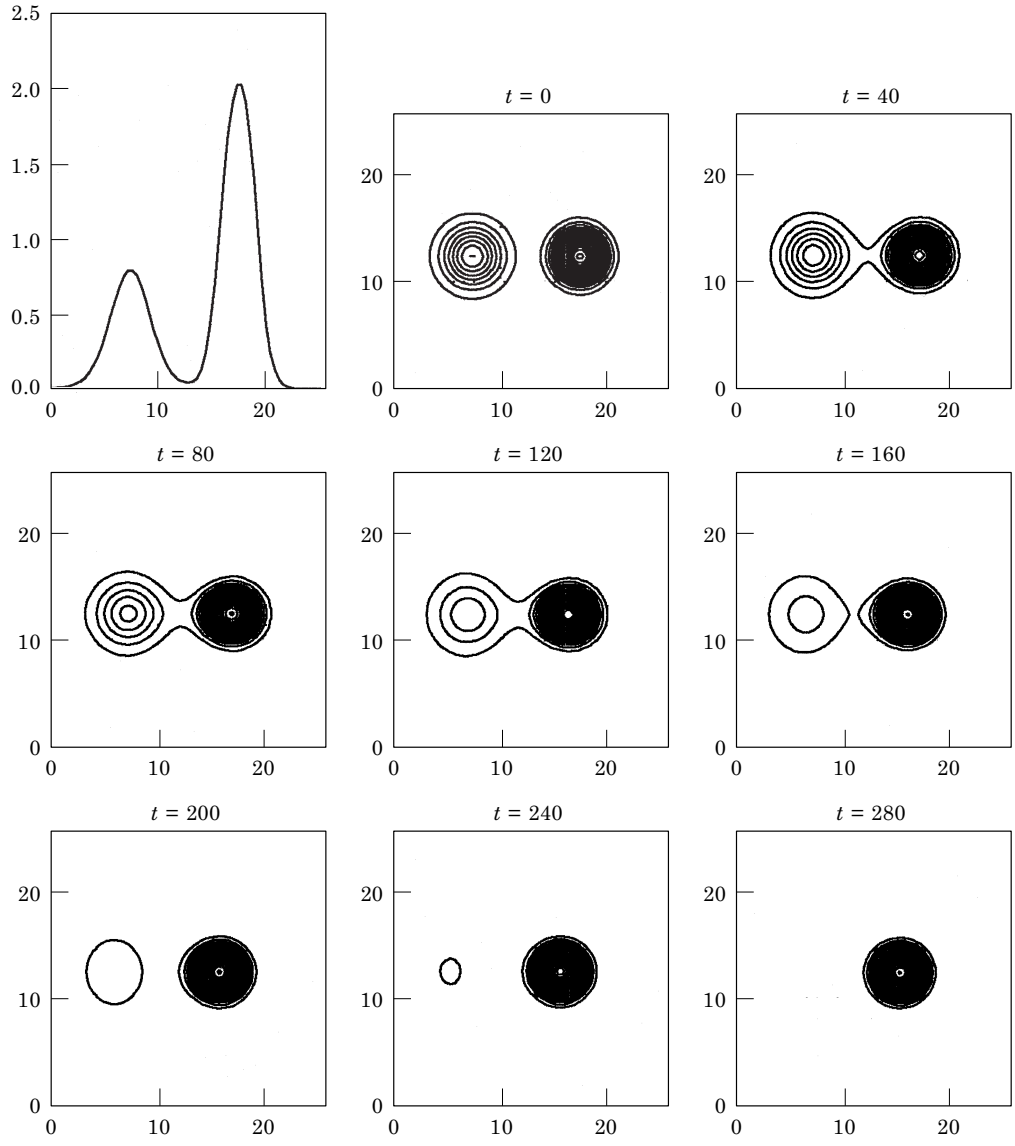


FIG. 23. Merger in the continuum model. Each group was produced by starting with an initial isolated patch and letting it evolve for 200 time units. The final states were then added together to produce the initial condition. The first frame shows a section through the groups at the initial time.

and leads to asymmetric distributions, which, by (26), result in density fluxes.

When the biological time-scales are rapid compared with grouping times, the first approximation to (24) is

$$n(\mathbf{x}, \mathbf{v}, t) = \rho(\mathbf{x}, t)\bar{n}(\mathbf{x}, \mathbf{v}, t)$$

where $\bar{n}$ is the right eigenfunction, e_0, for $\lambda = 0$

$$\int d\mathbf{v}'\psi(\mathbf{v}; \mathbf{x}, \mathbf{v}', t)\bar{n}(\mathbf{x}, \mathbf{v}', t) = 0$$

and is normalized by $\int d\mathbf{v}\bar{n}(\mathbf{x}, \mathbf{v}, t) = 1$.

There are two cases we can now consider, depending upon whether $\bar{n}(\mathbf{x}, \mathbf{v}, t) = \bar{n}(\mathbf{x}, -\mathbf{v}, t)$ or not. The symmetric case will arise if the transition rate from velocity $\mathbf{v}' \to \mathbf{v}$ is the same as the transition rate from $-\mathbf{v}' \to -\mathbf{v}$. In the symmetric case, the flux from $\int d\mathbf{v}v_i\bar{n}(\mathbf{x}, \mathbf{v}, t)$ is zero and we must consider higher order corrections. These lead to "kinesis"—the animals can move up or down a gradient (as we shall see) without being able to sense the gradient or determine a direction. In contrast, in the "taxis" case, the organisms determine their direction based on conditions and $\bar{n}$ is not

symmetric. The interesting limit is the case when the asymmetry is weak, so that its effects enter at the same order as those of kinesis.

In this case, we split ψ into the part associated with a symmetric equilibrium ψ^s and the rest ψ^u. We choose ψ^s so that $\mathbf{e}_0^\dagger = 1$ is still a left null eigenfunction (and therefore it is also a null vector of ψ^u). If we define $n = \rho\bar{n} + n'$, we find the correction to the local equilibrium velocity distribution satisfies

$$\begin{aligned}
&\int d\mathbf{v}'\psi^s(\mathbf{v}, \mathbf{x}', \mathbf{v}', t)n'(\mathbf{x}, \mathbf{v}, t) \\
&\quad = -\int d\mathbf{v}'\psi^u n(\mathbf{x}', \mathbf{v}', t) + \frac{\partial}{\partial t} n(\mathbf{x}, \mathbf{v}, t) \\
&\qquad + \nabla \cdot [\mathbf{v}n(\mathbf{x}, \mathbf{v}, t)] \\
&\quad \simeq -\rho \int d\mathbf{v}'\psi^u \bar{n}(\mathbf{x}', \mathbf{v}', t) + \frac{\partial}{\partial t}[\rho(\mathbf{x}, t)\bar{n}(\mathbf{x}, \mathbf{v}, t)] \\
&\qquad + \nabla \cdot [\mathbf{v}\rho\bar{n}(\mathbf{x}, \mathbf{v}, t)]. \qquad (30)
\end{aligned}$$

The $\partial\rho/\partial t$ term is higher order; therefore, the r.h.s. is still orthogonal to $\mathbf{e}_0^\dagger$, (30) will be solvable for n', and we will be able to evaluate the density flux

$$J_i = \int d\mathbf{v} v_i n'(\mathbf{x}, \mathbf{v}, t). \qquad (31)$$

By the definition (31), we are only concerned with the antisymmetric part of n'; that will be generated by the antisymmetric parts of the r.h.s. of (30). We shall assume that ψ depends only upon n (if at all) through ρ. In that case $\bar{n}$ may depend on ρ, but the term $\rho\partial\bar{n}/\partial t$ is still symmetric and can be ignored. We can now see that the flux has a term proportional to ρ, acting as an advection, and a term proportional to $\nabla\rho$, giving diffusion.

$$J_i = u_i\rho - \kappa_{ij}\frac{\partial}{\partial x_j}\rho$$

Notice that the advecting velocity $\mathbf{u}$ is non-zero even when ψ has only symmetric transitions and local equilibrium. This derivation implies that aggregation processes are similar under many different assumptions about the microscopic behavior: as long as the time-scales for the velocity distribution to come to local equilibrium are short and the transition rates depend on (at most) the density of the organisms (along with external environmental cues), we end up with a diffusion and a convergent/divergent velocity. Of course, the rates do depend strongly upon microscopic assumptions.

4.4. INSTABILITIES AND THE DEVELOPMENT OF PATCHES

As Fig. 22 demonstrated, the spatially-uniform density state can be unstable to disturbances which develop into nearly isolated groups. We would like to consider the linear and nonlinear development for an example of patch growth, taking the case of social taxis.

$$\frac{\partial}{\partial t}\rho = -\nabla \cdot [\mathbf{u}\rho - \kappa\nabla\rho]$$

and modelling the social behavior by taking the velocity to be

$$\mathbf{u} = \nabla\phi$$

with

$$\phi = \iint d\mathbf{z}\mathscr{W}(\mathbf{z})\rho(\mathbf{x} + \mathbf{z})$$

Non-dimensionalizing these equations can be helpful in studying their properties. We scale lengths by L, the characteristic perception distance (the scale inherent in $\mathscr{W}$), time by L^2/κ, velocity by κ/L, ϕ by κ, and densities by ρ_0. The equations become

$$\begin{aligned}
&\frac{\partial}{\partial t}\rho = -\nabla \cdot [\mathbf{u}\rho - \nabla\rho] \\
&\mathbf{u} = \nabla\phi \\
&\phi = \epsilon \iint d\mathbf{z}\mathscr{W}(\mathbf{z})\rho(\mathbf{x} + \mathbf{z}) \qquad (32)
\end{aligned}$$

with the non-dimensional parameter being the Peclet number $\epsilon = \rho_0 W_0 L^2/\kappa = u_{bio}L/\kappa$. Here W_0 is the characteristic magnitude of $\mathscr{W}$ (defined as the area integral of $\mathscr{W}$ divided by L^2)—implying we have the normalization $\iint\mathscr{W} = 1$ for the non-dimensional weighting function. The parameter $u_{bio} \equiv \rho_0 W_0 L$ is the characteristic movement

speed for the organisms; thus the Peclet number compares movement rates to diffusive rates.

If we substitute $\rho = \bar{\rho} + \rho' \exp(i\mathbf{k}\cdot\mathbf{x} + st)$ and linearize ($|\rho'| \ll \bar{\rho}$), we find the growth rate is given by

$$s = |\mathbf{k}|^2[\bar{\rho}\epsilon\hat{\mathscr{W}}(k)] - |\mathbf{k}|^2$$

$$\hat{\mathscr{W}}(\mathbf{k}) = \iint \mathscr{W}(\mathbf{z}) \exp(i\mathbf{k}\cdot\mathbf{z}) \tag{33}$$

and the uniform state is unstable if s is positive. Since $\hat{\mathscr{W}}(0) = \iint \mathscr{W} = 1$, s will be positive as $|\mathbf{k}| = \to 0$ if $\bar{\rho}\epsilon > 1$.

To analyse the nonlinear behavior, we take solutions of the form

$$\rho = \sum_n \rho_n(t) \exp(in\mathbf{k}\cdot\mathbf{x}). \tag{34}$$

For this form, we can also find ϕ since the integral equation (32) is linear:

$$\phi = \epsilon \sum_n \rho_n(t)\hat{\mathscr{W}}(n\mathbf{k}) \exp(in\mathbf{k}\cdot\mathbf{x}). \tag{35}$$

Substituting (35) into the dynamical equations (32) and combining terms gives

$$\frac{\partial}{\partial t}\rho_n = n|\mathbf{k}|^2 \sum_m [\epsilon m\hat{\mathscr{W}}(m\mathbf{k})\rho_m\rho_{n-m} - n\rho_n]. \tag{36}$$

For a weakly nonlinear stability calculation, we truncate (36) to include $n = -2, -1, 0, 1, 2$ and find

$$\frac{\partial}{\partial t}\rho_0 = 0$$

$$\frac{\partial}{\partial t}\rho_1 = |\mathbf{k}|^2[\epsilon\hat{\mathscr{W}}(\mathbf{k})\rho_0\rho_1 + \epsilon(2\hat{\mathscr{W}}(2\mathbf{k}) - \hat{\mathscr{W}}(\mathbf{k}))\rho_2\rho_1^* - \rho_1]$$

$$\frac{\partial}{\partial t}\rho_2 = 2|\mathbf{k}|^2[2\epsilon\hat{\mathscr{W}}(2\mathbf{k})\rho_0\rho_2 + \epsilon\hat{\mathscr{W}}(\mathbf{k})\rho_1^2 - 2\rho_2]$$

(ρ_1^* being the complex conjugate of ρ_1) or

$$\frac{\partial}{\partial t}\rho_1 = s_1\rho_1 + |\mathbf{k}|^2\epsilon(2\hat{\mathscr{W}}(2\mathbf{k}) - \hat{\mathscr{W}}(\mathbf{k}))\rho_2\rho_1^*$$

$$\frac{\partial}{\partial t}\rho_2 = s_2\rho_2 + 2|\mathbf{k}|^2\epsilon\hat{\mathscr{W}}(\mathbf{k})\rho_1^2 \tag{37}$$

with the s_n being the growth rates for perturbations of the uniform state with $\bar{\rho} = \rho_0$ from (33). When $s_2 < 0$, we can expect a center manifold approximation to hold, so that $\rho_2 \simeq R(\rho_1)$ and the leading term will be quadratic $\rho_2 \simeq r\rho_1^2$. Substituting this in (37) and solving the second equation for r gives

$$\rho_2 \simeq \frac{2|\mathbf{k}|^2\epsilon\hat{\mathscr{W}}(\mathbf{k})}{2s_1 - s_2}\rho_1^2$$

and the first gives

$$\frac{\partial}{\partial t}\rho_1 \simeq [s_1 + N|\rho_1|^2\rho_1 \tag{38}$$

with

$$N = \frac{2|\mathbf{k}|^4\epsilon^2\hat{\mathscr{W}}(\mathbf{k})(2\hat{\mathscr{W}}(2\mathbf{k}) - \hat{\mathscr{W}}(\mathbf{k}))}{2s_1 - s_2}. \tag{39}$$

Thus, the nonlinearity will be *destabilizing* if $2\hat{\mathscr{W}}(2\mathbf{k}) > \hat{\mathscr{W}}(\mathbf{k})$ and $s_1 > 0$, $s_2 < 0$. Below, we shall look at a specific example; however, we can readily show that nonlinearity destabilizes long waves and that these are the waves excited when $\epsilon\rho_0 = 1 + \delta > 1$—the system is just above threshold. For long waves,

$$\hat{\mathscr{W}}(n\mathbf{k}) \simeq \iint \mathscr{W}(\mathbf{z})[1 + in\mathbf{k}\cdot\mathbf{z} - \frac{1}{2}(n\mathbf{k}\cdot\mathbf{z})^2]$$

$$\simeq 1 - \alpha n^2|\mathbf{k}|^2 \tag{40}$$

with $\alpha = 1/2\iint \mathscr{W}(\mathbf{z})(\hat{\mathbf{k}}\cdot\mathbf{z})^2$ being a constant in the case of a radially symmetric $\mathscr{W}$. Then the growth rates look like

$$s_n = n^2|\mathbf{k}|^2[(1+\delta)(1 - \alpha n^2|\mathbf{k}|^2) - 1]$$

$$\simeq n^2|\mathbf{k}|^2[\delta - \alpha n^2|\mathbf{k}|^2].$$

Therefore the unstable waves are long: $|\mathbf{k}|^2 < \delta/\alpha$. Returning to (40), we see that $2\hat{\mathscr{W}}(2\mathbf{k}) - \hat{\mathscr{W}}(\mathbf{k}) \simeq 1$ when $|\mathbf{k}|^2$ is small—the nonlinearity is destabilizing.

We have calculated the growth rates and nonlinear coefficients for some of the models used in Sections 2 and 3. Figure 24 shows examples of these. Near the critical value of $\bar{\rho}\epsilon$, the growing waves are long and nonlinearity is destabilizing, but the scale becomes order one as the parameter becomes larger. There is a range of k for which the nonlinearity is destabilizing (and in other regions, multiple wavenumbers are

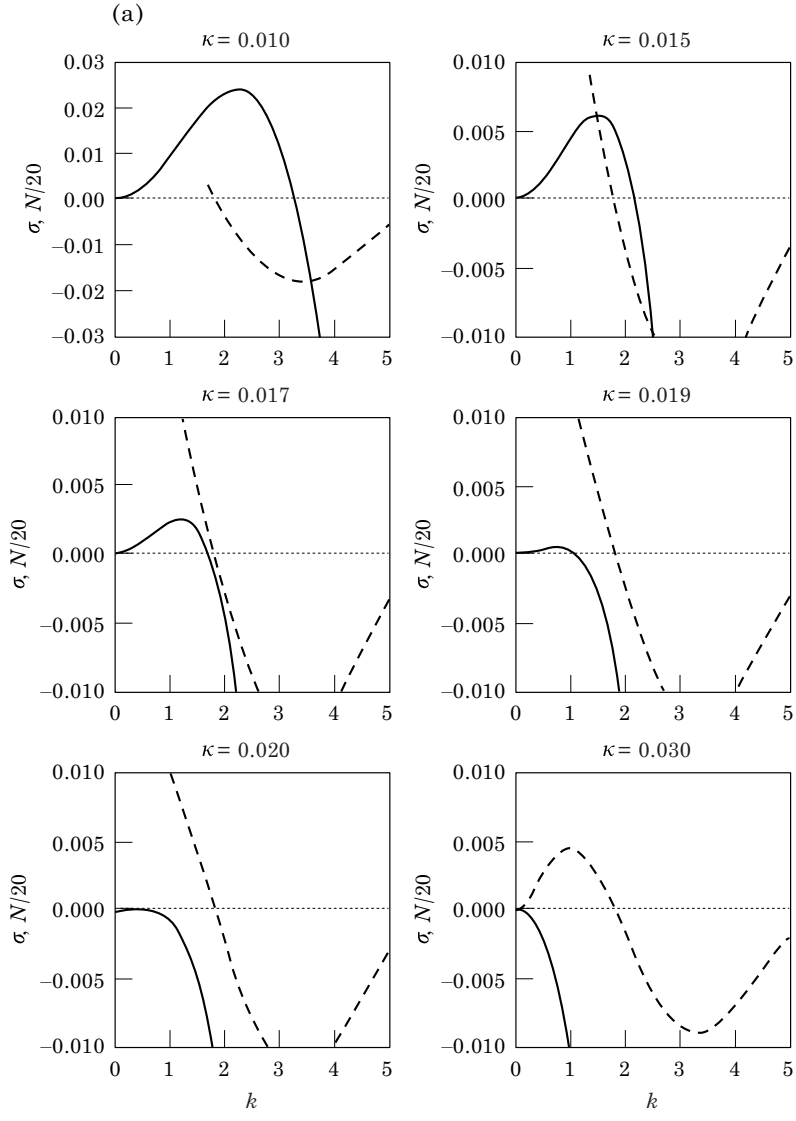


FIG. 24(a)—(*Caption on p.* 445)

unstable), so that instabilities, once started, will grow to large amplitude.

In view of analytical and numerical difficulties posed by partial integro-differential equations, it is of interest to consider the convolution form in relation to the regular differential forms associated with it via Taylor expansion (Alt, 1985). As an example, we compare the linear stability of this two-dimensional version of the Kawasaki (1978) model to that estimated from differential approximations to it. Approximating the non-local distribution of ρ in the convolution (32) by a Taylor series of n-th order gives an expression for velocity which is a linear function of derivatives of ρ, evaluated at x, and moments of the weighting function $\mathscr{W}$,

$$\phi \simeq \rho + \nabla_i \rho \iint z_i \mathscr{W} + \frac{1}{2} \nabla_i \nabla_j \rho \iint z_i z_j \mathscr{W} + \cdots \tag{41}$$

This approximation gives a growth rate (if the necessary moments exist) of

$$s = -|\mathbf{k}|^2 + |\mathbf{k}|^2 \bar{\rho} \epsilon \left[\hat{\mathscr{W}}|_{k=0} + ik_i \left. \frac{\partial \hat{\mathscr{W}}}{\partial k_i} \right|_{k=0} - \frac{k_i k_j}{2} \left. \frac{\partial^2 \hat{\mathscr{W}}}{\partial k_i \partial k_j} \right|_{k=0} + \cdots \right]$$

In general, for an n-th order of the Taylor expansion of (41), the approximate growth rate will be the n-th order expansion of $\hat{\mathscr{W}}(\mathbf{k})$ around $\mathbf{k} = 0$. The factor appearing in the growth rate derived from the integral equation, $\hat{\mathscr{W}}(k)$, is bounded as the wavenumber k approaches infinity, while the differential approximation gives a polynomial of finite order for the growth rate which goes to infinity (with the correct or incorrect sign depending on n) with increasing k. Thus (41) can give rapidly growing short wave modes which are not present in the integral equation form. We conclude that convolutions such as (31) cannot be replaced with differential forms, even if the true density distributions are smooth and well approximated locally by Taylor expansions of low order (see also Alt, 1985).

4.5. ISOLATED GROUPS

We can solve for steady-state groups that decay to a uniform exterior density, which we can normalize to $\rho = 1$ (cf. Fig. 25). For the exterior to be stable, we require $\epsilon < 1$. (If we want to define a larger Peclet number, we can choose a different background density value ρ_0). The isolated group will be radially symmetric and the flux must vanish at each r; thus,

$$\rho \frac{\partial}{\partial r} \phi = \frac{\partial}{\partial r} \rho$$

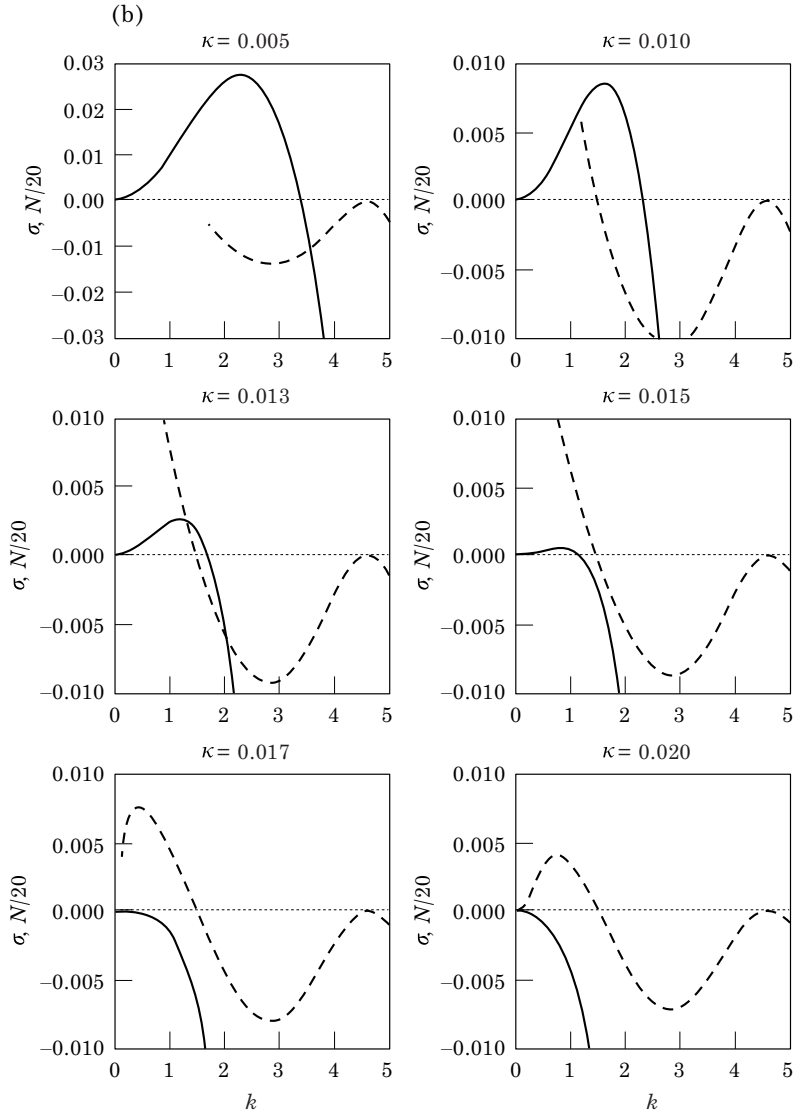


FIG. 24(b)—(*Caption on opposite page*).

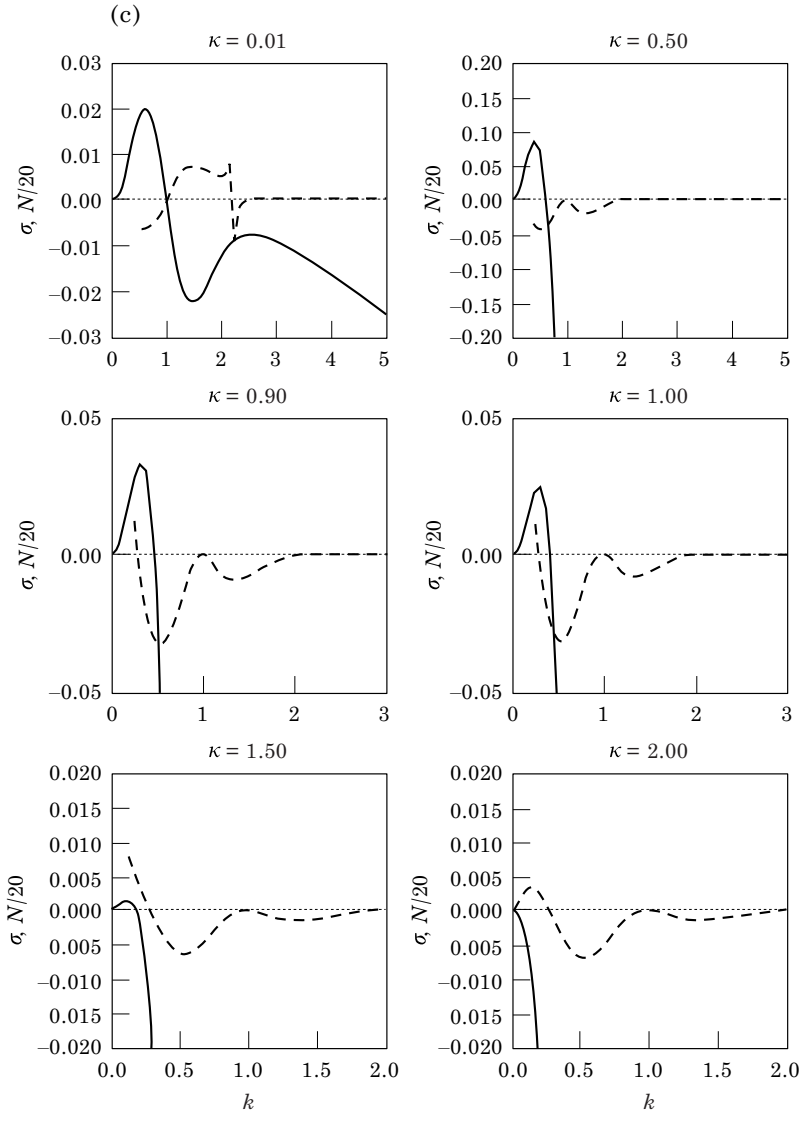


FIG. 24(c)

FIG. 24. Growth rates s_1 (—) and nonlinear coefficients N (– – –) for various κ values for the weighting (6), $\mathscr{W}(\mathbf{z}) = \mathscr{V}(0.25 - 05|\mathbf{z}|^2 + 0.25|\mathbf{z}|^4)$. These are dimensional, rather than non-dimensional. (b) Growth rates and nonlinear coefficients for various κ values for the weighting (7), $\mathscr{W}(\mathbf{z}) = \mathscr{V}(0.06875 + 0.4125|\mathbf{z}|^2 - 1.03125|\mathbf{z}|^4 + 0.55|\mathbf{z}|^6)$. (c) Growth rates and nonlinear coefficients for various κ values for the weighting used in (14c), $\mathscr{W}(\mathbf{z}) = 0.28\exp(-0.125|\mathbf{z}|^2) - 0.2277\exp(-0.25|\mathbf{z}|^2)$.

which implies

$$\phi = \epsilon + \ln \rho \tag{42}$$

where the integration constant is chosen to match the far-field conditions, $\rho \to 1$ and eqns (32) and the normalization condition on $\mathscr{W}$ have been used.

Thus, we can write the problem in terms of a single integral equation

$$\epsilon \iint \mathscr{W}(\mathbf{z})[\rho(\mathbf{x}+\mathbf{z}) - 1] = \ln \rho(\mathbf{x}). \tag{43}$$

4.5.1. *Green's function*

For particular choices of $\mathscr{W}$, we can rewrite (43) as a differential equation and look for solutions with a numerical shooting algorithm. One simple case is

$$\mathscr{W}(\mathbf{z}) = \frac{1}{2\pi} K_0(|\mathbf{z}|)$$

which satisfies

$$\nabla_z^2 \mathscr{W}(\mathbf{z}) - \mathscr{W}(\mathbf{z}) = -\delta(\mathbf{z})$$

with the subscript on ∇_z^2 serving as a reminder of which variables the Laplacian is acting upon. If we multiply by $-\phi(\mathbf{x}+\mathbf{z})$ and integrate over $\mathbf{z}$, we find

$$\begin{aligned}\phi(\mathbf{x}) &= -\iint \phi(\mathbf{x}+\mathbf{z})(\nabla_z^2 - 1)\mathscr{W}(\mathbf{z}) \\ &= -\iint \mathscr{W}(\mathbf{z})(\nabla_z^2 - 1)\phi(\mathbf{x}+\mathbf{z}) \\ &= -\iint \mathscr{W}(\mathbf{z})(\nabla_{x+z}^2 - 1)\phi(\mathbf{x}+\mathbf{z}). \quad (44)\end{aligned}$$

Comparison with (32) shows that

$$(\nabla_x^2 - 1)\phi(\mathbf{x}) = -\epsilon\rho(\mathbf{x}) = -\epsilon \exp(\phi - \epsilon)$$

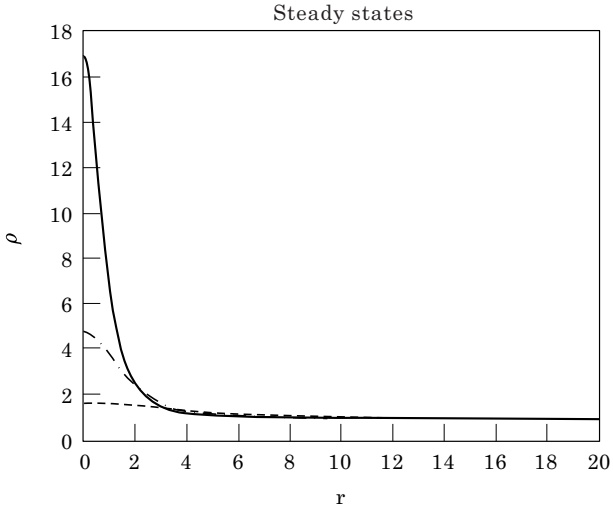


FIG. 25. Steady solutions for $\epsilon = 0.9$ (– – –), 0.7 (–.–.–) and 0.5 (—).

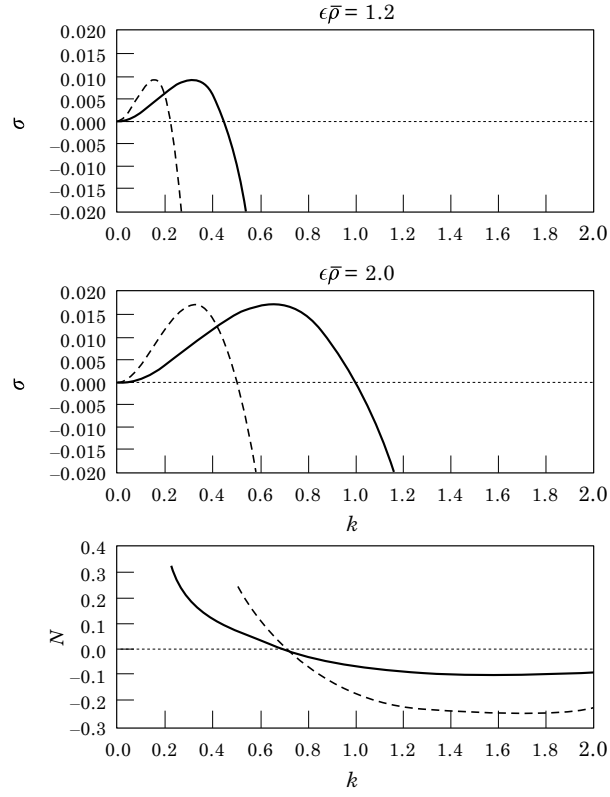


FIG. 26. Growth rates σ_1 for two different $\epsilon\bar{\rho}$ values: (– – –) show σ_2. The third plot shows the nonlinear coefficient N for $\epsilon\bar{\rho} = 1.2$ (—) and 2.0 (– – –).

with the last identity from (42). If we define $\phi' = \phi - \epsilon = \ln\rho$, we derive a simple nonlinear equation for the steady structure:

$$(\nabla^2 - 1)\phi'(\mathbf{x}) = \epsilon(1 - \mathrm{e}^{\phi'})$$

We have solved this in radial coordinates, using a second order Runge–Kutta integration, starting with the initial conditions $\phi = \phi_0$, $\phi_r = 0$ at $r = 0$. The constant ϕ_0 is adjusted to give well-behaved solutions as $r \to \infty$ (actually at $r = 40$). The solutions for two values of ϵ are shown in Fig. 25. Note that the groups get sharper and narrower as ϵ decreases.

4.5.2. *Growth rates/nonlinearity*

For this form of weighting, we can evaluate $\hat{\mathscr{W}}(k)$ explicitly [though the simplest way is to follow a procedure like (44)]

$$\hat{\mathscr{W}}(k) = \frac{1}{1 + |\mathbf{k}|^2}$$

allowing us to plot the growth rates [from (33)]

$$s = |\mathbf{k}|^2\left[\frac{\epsilon\bar{\rho}}{1+|\mathbf{k}|^2} - 1\right]$$

as shown in Fig. 26. We also show the nonlinear coefficient N from (39).

4.6. SUMMARY

Our derivations emphasize the generality of the advective–diffusive form for the flux and demonstrate that we can, at least in simple cases, find the forms of the velocities and diffusivities from the model for individual behavior. In general, we expect convergent velocity fields will arise, either from taxis or kinesis with chemical or social cues. The more random component of the motion can balance this convergence, resulting in organized groups. The continuum models show the development of such structures, as well as the phenomenon of group merger and (we expect) breakup.

At the same time, continuum models do require additional statistical assumptions which can only be validated as reasonable by detailed analysis and/or individual-based simulations. We also note that some processes, such as the slow random-walking of a group tend not to appear in the continuum models. Likewise, schooling behavior has not been examined, but we expect that a continuum description will require additional fields beyond density to capture the overall movement of the school (as discussed in the following section).

5. Population-level Movements through Group Size Distribution Dynamics

5.1. THE PURPOSE OF LOOKING AT GROUP SIZE DISTRIBUTIONS

One of the ultimate goals of modeling social grouping behaviors is to understand and predict the ecological role of such behaviors in shaping spatial and temporal distributions of social animals and their associated predator and prey species. These types of ecosystem-level questions typically deal with time and space scales much larger than those characterizing group movements. In many cases, the scales over which interesting population dynamics take place are so large that, relatively speaking, groups lose their identities (due to fragmentation and fission) over very brief periods. In situations where this difference of scales exists, modeling approaches that focus on explicit group-by-group dynamics seem to be limited by practical constraints to time and space scales too small to answer ecological questions. On the other hand, this difficulty suggests alternative mathematical approaches specifically based on multiple space/time-scales. One established line of mathematical development that falls into this category is the literature on non-spatial descriptions of group size distributions (Okubo, 1986; Gueron & Levin, 1993). Another relatively well studied type of analysis stemming from multiple scales is the translation of spatially-explicit, individual-based biased random walks to advection–diffusion equations for populations (Patlak, 1953; Keller & Segel, 1971; Alt, 1980; Othmer *et al*., 1988). In this section, we suggest an approach combining these two kinds of analysis to obtain population-level advection–diffusion equations that *implicitly* contain both the dynamics of group fusion and fission and the different biased random walk characteristics associated with groups of different sizes.

5.2. POPULATION-LEVEL EFFECT OF GROUPING

Non-spatial group size distribution models describe the dynamics after individuals have collected into schools or swarms which can move around in the area of interest, can interact with other groups, and can change sizes by losing or gaining members (Okubo, 1986; Gueron & Levin, 1993). Small schools are created by the fission of existing larger schools, large schools are created by the fusion of smaller schools, and so on. For a given set of social behavioral algorithms, school properties such as size and shape, average velocity, and cohesiveness typically vary with the number of school members. Thus, we expect that fission and fusion rates are functions of school size. Of course, for any particular school, the probability of splitting, or of encountering and fusing with another school, depends on the exact spatial arrangement of that school and its neighbors. However, for a relatively large area containing numerous

schools, an average rate of fusion and fission can be defined for schools of a given size, based on school properties and the overall density of other schools in the area.

Using these size-dependent fission and fusion rates, a non-spatial description of the local density of schools of size j, $g_j(t)$, can be written as a function of time in the form of a *dynamical system*,

$$\frac{\mathrm{d}\, g_j}{\mathrm{d}\, t} = \frac{1}{2}\sum_{k=1}^{j-1} a(k, j-k) g_k g_{j-k} - \sum_{k=1}^{\infty} a(j,k) g_k g_j + \sum_{k=j+1}^{\infty} b(k,j) g_k - \frac{1}{2}\sum_{k=1}^{j-1} b(j,k) g_j \quad (45)$$

where $a(j, k)$ is the rate at which schools of size j fuse with schools of size k, and $b(j, k)$ is the rate at which schools of size j split to form schools size k and $(j - k)$ (Okubo, 1986). In (45), the first and third terms on the right hand side represent creation of j-schools by fusion of smaller schools and fission of larger schools, respectively. The remaining terms represent disappearance of j-schools by fusion or fission.

5.3. FISSION AND FUSION RATES

In (45), fusion rates are assumed to be quadratic in school density, while fission rates are proportional to the density. These functional forms are based on the assumption that fission is essentially an intra-group phenomenon; i.e. it does not depend on the presence or absence of neighboring groups. In contrast, fusion is assumed to be principally determined by encounter rates between schools—how often groups come close enough to detect and respond to neighboring groups—and simple analyses suggest that any particular group's rate of encounter is simply proportional to the density of other groups in the area. This implies the quadratic functional form for fusion.

The quadratic form in turn suggests that as population densities become larger, fusion rates typically increase more quickly than fission rates. At low population densities, then, we expect a small mean school size at the equilibrium of (45); with increasing population density, we expect not only an increase in number of schools but also a shift towards larger modal school size.

5.4. AN EXAMPLE: SIMPLIFIED 2D SCHOOLING

Of course, dynamical systems descriptions of group size distribution like (45) do not of themselves solve the problem of tying individual social behavior to population fluxes. However, they can be useful as intermediate descriptions that can be coupled rigorously both to the lower level of individual behavior and to the higher level of moving populations. As a concrete (though highly incomplete) example of how this coupling might be attempted, we begin with a highly simplified model of schooling behavior, and carry out some of the steps required to derive the dynamical systems description.

For this example, we use the schooling model described briefly in Section 2 (Fig. 10); more details of behavior and some of the interactions between social and tactic behaviors are described in Grünbaum (1998a). We used simulations with large numbers of individuals (shown in Fig. 10), together with some basic analysis, to estimate the rates of fission and fusion appropriate to use in (45) for this schooling model. The group size distribution is assumed to be approximately stationary, with fusion roughly balancing fission at each group size.

5.5. FUSION RATES

In our example, we estimate fusion rates using a two-dimensional version of the encounter rate models of Koopman (1956), Gerritsen & Strickler (1977), and Rothschild & Osborne (1988). According to this model, $a(j, k)$, the rate at which j-sized schools will encounter k-sized schools, is given by

$$a(j,k) = \frac{4}{\pi}(R(j) + R(k))(U(j) + U(k))E\left(\frac{2\sqrt{U(j) * U(k)}}{U(j) + U(k)}\right). \quad (46)$$

In (46), $R(j)$ and $R(k)$ are the characteristic lengths of schools of size j and k, respectively, $U(j)$ and $U(k)$ are the characteristic velocities, and E is the complete elliptic integral of the second kind. Equation (46) makes it clear that we have to know the variation of R and U with

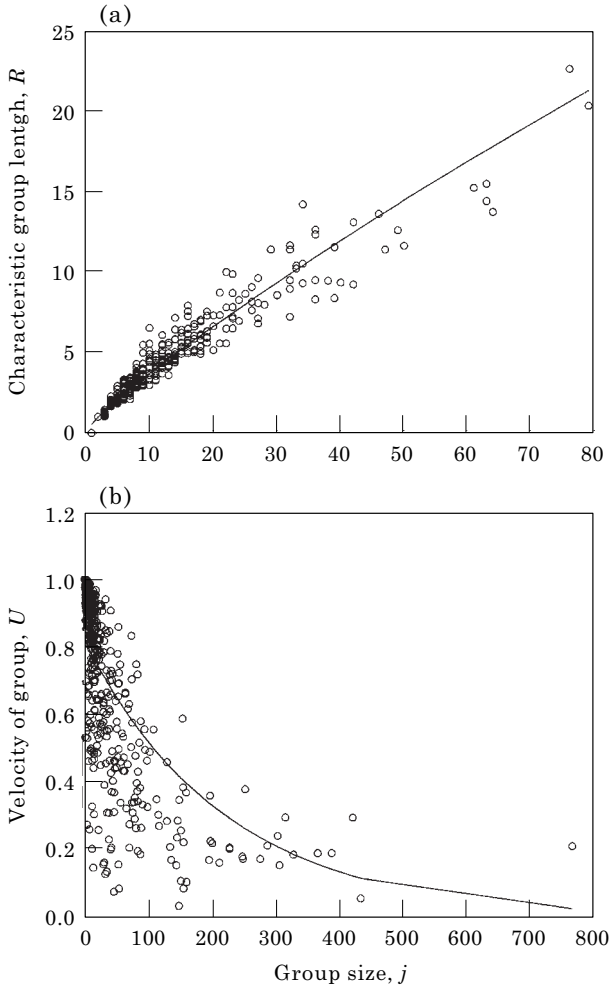


FIG. 27. (a) School length R vs. group size j. (b) School velocity U vs. group size j.

group size. Since we cannot calculate them analytically for our schooling behavior, these parameters must be measured from simulations. Figure 27 shows frequency diagrams for R and U observed in our simulations; we approximate these distributions by curve-fitting as

$$R(j) = L_0 j^{L_1}$$

$$U(j) = c_1\, 10^{-c_2 j}.$$

From our simulations, we estimate that $L_0 = 0.52$, $L_1 = 0.85$, $c_1 = 0.823$, and $c_2 = 0.002$.

5.6. FISSION RATES

Fission rates in the simulations are approximately linear as a function of school size, but have a non-zero intercept (Fig. 28),

$$b(j,k) = \begin{cases} b_0\left(\dfrac{j-v}{j-2v}\right) & j > vv \;\;\&\;\; v < k < j - v \\ 0 & \text{otherwise} \end{cases}. \tag{47}$$

This means that schools smaller than a threshold "unit" size (v) virtually never split, while larger schools have a relatively short "half-life." The unit size observed in the simulation corresponds to 2–3 times the number of neighbors perceived by an individual within a school ($v \cong 11$, $b_0 \simeq 0.00011$ in these simulations).

5.7. GROUP SIZE DISTRIBUTIONS

Using (46), (47) as the fusion and fission functions for (45), we can numerically estimate the group size distributions that would result from our schooling behavior in simulations with many more individuals than our computational capacity allows us to carry out. It turns out that the groups resulting from the simple schooling algorithm are very nearly exponentially distributed [Fig. 29(a)]. This is not altogether surprising—Okubo (1986) used a maximum entropy analysis to argue that the exponential distribution is in some sense a "natural" group

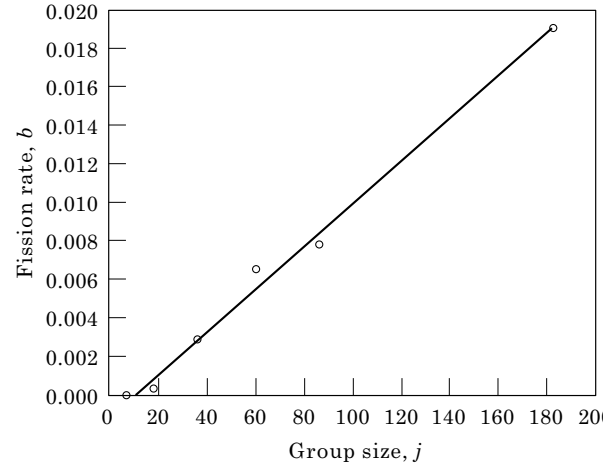


FIG. 28. Splitting rates vs. group size.

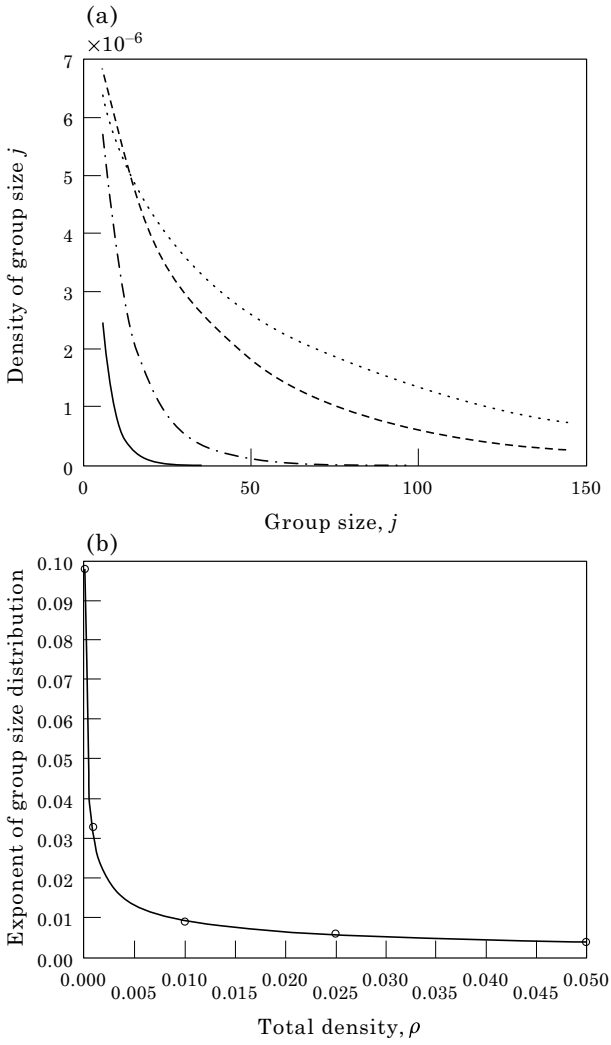


FIG. 29. (a) Group size distributions for various densities. (—) $p = 0.0001$, (–·–·–) $p = 0.001$, (– – –) $p = 0.01$, (. . .) $p = 0.025$. (b) Group size exponent vs. density.

size distribution, and went on to document numerous instances of this type of distribution observed in social animals. In our case, it is a particularly convenient result, because it allows us to characterize the group size distribution with a single density-dependent exponent [Fig. 29(b)]. Note that increasing density has the expected result of shifting the group size distribution towards larger groups.

5.8. A HEURISTIC APPROACH TO POPULATION-LEVEL FLUXES

As shown in Fig. 27, group speeds tend to decrease with increasing group sizes; and, since groups tend to get larger as total population density, ρ, increases, we expect that the movement of the population due to the random motion of groups will change also (probably decrease as density increases). As an attempt to anticipate the consequences of schooling behavior for population fluxes, we present a heuristic discussion of how a diffusion equation with density-dependent coefficients might describe such a situation. A rigorous analysis would most likely follow derivations from the random walk literature, which use perturbation techniques (Grünbaum, 1998b). To be maximally useful, such an analysis would also take into account not only environmental advection terms but also directional biases corresponding to taxis at the group level.

Our heuristic argument is as follows: the density g_j of any given group size j will diffuse due to the random motion of the groups (Okubo, 1980, 1986; Edelstein-Keshet, 1988; Murray, 1990) as

$$\frac{\partial}{\partial t} g_j = -\nabla(D_j \nabla g_j). \tag{48a}$$

In (48), the size-specific diffusivity, D_j, is estimated as

$$D_j = \frac{1}{2} U^2(j)\tau_j. \tag{48b}$$

τ_j, the directional persistence time for a school of size j, should be measured from school trajectories, but since we do not have this information, we assume it to be constant, $\tau_j = \tau$. In practice, large groups probably have greater directional persistence than small groups; however, in our simulations, the half-life of larger groups is so short that they almost invariably split before turning, so we expect the error from this simplification to be small. We then sidestep the whole issue by scaling our time variable by τ.

Multiplying both sides of (48) by j, summing over all group sizes, and rearranging gives

$$\frac{\partial}{\partial t} \rho = -\nabla(\bar{D}(\rho)\nabla\rho). \tag{49a}$$

where

$$\bar{D}(\rho) = \sum_{j=0}^{\infty} jD_j(\rho \frac{\partial}{\partial \rho} f_j + f_j), \quad f_j(\rho) = \frac{1}{\rho} g_j. \tag{49b}$$

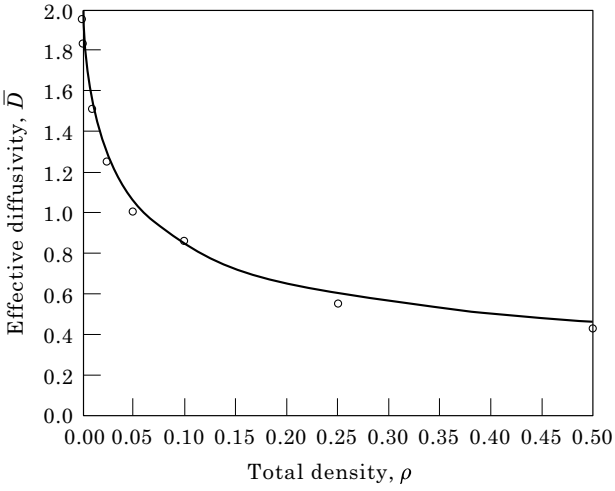


FIG. 30. Effective diffusivity vs. density.

In (49), f_j is the *relative frequency* of group size j at total density ρ, and $\bar{D}(\rho)$ is the *effective diffusivity* at population density ρ.

Equation (49) is our heuristic estimate of the diffusive flux of population density for individuals schooling according to our simple behavior, and is the principal result of this section. The effective diffusivity based on the parameters from our simulation results is a strongly decreasing function of school size (Fig. 30). This confirms the expectation that a dense *social* population diffuses relatively more slowly than the same density of an *asocial* species would (i.e. a population composed entirely of group size one). Figure 31 shows how social behavior affects the spread of an initially clumped population. Compared to an asocial population, the population concentration is dispersed more slowly, but diffusion is rapid at low densities at the fringes. Asocial diffusion from an initially clumped distribution results in the standard Gaussian; in contrast, effective diffusivity in a social population apparently results in a *leptokurtic* distribution (Zar, 1996).

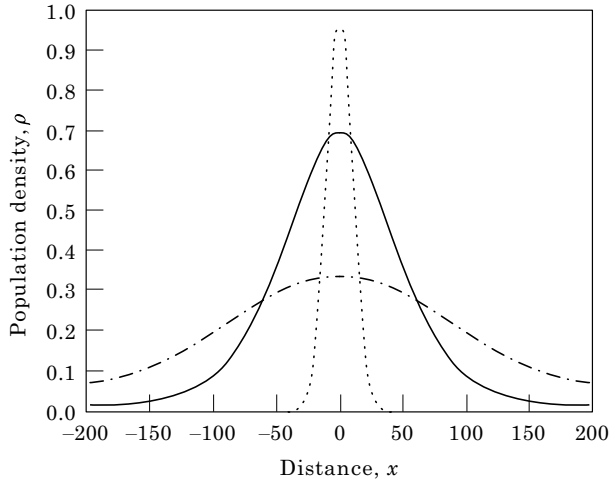


FIG. 31. Diffusion of an initial concentration (. . .) in the case without social behavior (–·–·–) and with social behavior (—). In the latter case, the individuals are organized into many small schools.

6. Concluding Remarks

The problems of understanding the collective dynamics of ensembles of individual agents is a fundamental one throughout science. Animal aggregations provide an ideal system for studying such processes because the responses of individuals to their local environments can be well characterized in terms of observations, or via hypotheses developed to explore ecological or evolutionary processes, and because a remarkable diversity of macroscopic patterns are observed in Nature.

The evolutionary determinants of grouping behavior differ from system to system; depending on the circumstances, the ultimate causes may involve improved foraging for resources, defense against predators, or mating success. Whatever the explanations, however, the mechanisms of aggregation are almost always mediated proximally in terms of individual responses to their local environments. One of the central challenges in the subject, therefore, is to relate macroscopic dynamics to local behavior, and to scale from the individual to the group. The approach we take is a well established one (see, e.g. Grünbaum & Okubo, 1994), beginning from a Lagrangian representation of an individual's acceleration as a sum of forces due to intrinsic and extrinsic factors, including those deriving from the fluid dynamics, chemical and resource gradients, and inter-individual attraction and repulsion. By varying the importance of the various factors, and the responsiveness to them, we explore their interplay, and distinguish the determinants of pattern on different scales.

We have reviewed the fundamental approach and assumptions in individual-based (Lagrangian) and continuum (Euler) models of animal aggregation, as well as methods for

relating one to the other. In the individual-based approach, the forces, both deterministic and stochastic, are specified (along with hopping movements if applicable), and the equations for each organism's velocity and position are integrated. Aggregation can result from either environmental or social responses. Both directional responses (taxis), where the accelerations depend on the gradient of the cue field, and non-directional (kinesis), where only the local concentration of the cue field matters, can lead to group formation. We have shown that environmental variability or flow may not seriously impede grouping and, indeed, may enhance it by increasing the encounter and merger rates of groups. Turbulent flows, which might be expected to break groups apart, do not have the same effects as enhanced diffusivity. Instead, because the small scale shears are weak, groups can form readily and merge more rapidly.

The continuum models allow analytical approaches to questions of spontaneous emergence of patches and the structure of isolated patches. Patch merger rates may likewise be accessible via perturbation approaches. Although derivation of continuum equations from individual-based models does require statements about the statistics that are not strictly accurate for social behavior, our experiments show that the continuum model nevertheless reproduces the results of the original individual-based model quite well. Whether or not this will hold for other forms of social interaction remains an open question.

Finally, the study of individual-based models can lead to formulation of models for group-size distributions. We can extend these to a spatial domain with environmental variability; however, they will have multiple state variables at each point. In effect, we are supplementing the density with information about the group statistics. More study is required to gain insight into the minimal amount of information which must be carried in large-scale models in order to represent properly the aggregation process.

We have not dealt here with the effects of grouping on predator–prey or competitive interactions. That investigation is ongoing. The common biological models for interacting species assume, in effect, that the distributions are Poisson. Thus, it seems clear that such interactions will be strongly influenced by the non-random distributions induced by social behavior or environmental responses (including response to prey or predator distributions). We believe that the combined use of individual-based, continuum, and group statistics models will allow us to assess the effects of spatial patchiness on species interaction and to understand how such behaviors might evolve.

We would like to acknowledge support from the Office of Naval Research, under ONR-URIP N00014-92-J-1527, "Modelling Biological-Physical Interactions: a Population Biological Approach." In addition, GF has support from NSF grant OCE-9525809, "Biogeochemical cycling of carbon in the southern ocean." DG would like to acknowledge the support from the Killam Foundation and N.S.F. Grant 9503478, "Special Year in Mathematical Biology at the University of Utah," and SL is pleased to acknowledge the support of the National Aeronautics and Space Administration through grants NAGW-4688 and NAG5-6422 and of the Alfred P. Sloan Foundation through grant 97-3-5. The tuna work was augmented by access to data from M. Lutcavage at the New England Aquarium. We are especially grateful for the extremely careful and thoughtful comments by our reviewer; they have been of great help in clarifying the paper.

REFERENCES

Abbott, M. R. & Zion, P. M. (1987). Spatial and temporal variability of phytoplankton pigment off northern California during CODE-1. *J. Geophys. Res.* **92,** 1745–1756.

Alt, W. (1980). Biased random walk models for chemotaxis and related diffusion approximations. *J. Math. Biol.* **9,** 147–177.

Alt, W. (1985). Degenerated diffusion equations with drift functionals modelling aggregation. *Nonlinear Anal.* **9,** 811–836.

Aoki, I. (1982). A simulation study on the schooling mechanism in fish. *Bull. Japan. Soc. Sci. Fish.* **48,** 1081–1088.

Berg, H. C. & Brown, D. A. (1972). Chemotaxis in *Escherichia coli* analyzed by three-dimensional tracking. *Nature* **239,** 500–504.

Chandrasekhar, S. (1943). Stochastic problems in physics and astronomy. *Rev. Modern Phys.* **15,** 1–89.

Chapman, S. (1916). On the law of distribution of molecular velocities, and on the theory of viscosity and thermal conduction in a non-uniform simple monatomic gas. *Phil. Trans. Roy. Soc. A* **216,** 279; see also Harris (1971).

Cohen, D. S. & Murray, J. D. (1981). A generalized diffusion model for growth and dispersal in a population. *J. Math. Biol.* **12,** 237–249.

Craig, C. L. (1987). The ecological and evolutionary interdependence between web architecture and web silk spun by orb web weaving spiders. *Biol. J. Linnean Soc.* **30,** 135–162.

Dal Passo, R. & de Mottoni, P. (1984). Aggregative effects for a reaction–advection equation. *J. Math. Biol.* **20,** 103–112.

Davis, C. S., Flierl, G. R., Wiebe, P. H. & Franks, P. J. S. (1991). Micropatchiness, turbulence, and recruitment in plankton. *J. Mar. Res.* **49,** 109–151.

Diekmann, O., Heesterbeek, J. A. P. & Metz, J. A. J. (1990). On the definition and the computation of the basic reproduction ratio Ro in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28,** 365–382.

Edelstein-Keshet, L. (1988). *Mathematical Models in Biology*, New York: Random House.

Enskog, D. (1917). Kinetische Theorie der Vorgänge in mässig verdünnten Gaser. Doctorial Dissertation, Uppsala University, Uppsala, Sweden; see also Harris (1971).

Epstein, A. W. (1996). Physical processes and zooplankton distribution in the Great South Channel: observational and theoretical studies. Ph.D. Dissertation, MIT-WHOI Joint Program in Physical Oceanography, 198 pp.

Franks, P. J. S. (1992). Sink or swim: accumulation of biomass at fronts. *Mar. Ecol. Prog. Ser.* **82,** 1–12.

Gerritsen, J. & Strickler, J. R. (1977). Encounter probabilities and community structure in zooplankton: a mathematical model. *J. Fish. Res. Board Canada* **34,** 73–82.

Grindrod, P. (1991). *Patterns and Waves* 239 pp. Oxford: Oxford University Press.

Grünbaum, D. (1992). Local processes and global patterns: biomathematical models of bryozoan feeding currents and density dependent aggregations in Antarctic krill. Ph.D. Thesis, Cornell University, Ithaca, NY.

Grünbaum, D. (1994). Translating stochastic density-dependent individual behavior to a continuum model of animal swarming. *J. Math Biol.* **33,** 139–161.

Grünbaum, D. (1997). Schooling as a strategy for taxis in a noisy environment. In: *Animal Groups in Three Dimensions* (Parrish, J. K. & Hamner, W. M., eds), pp. 257–281. Cambridge: Cambridge University Press.

Grünbaum, D. (1998a). Schooling as a strategy for taxis in a noisy environment. *Evol. Ecol.* **12,** 503–522.

Grünbaum, D. (1998b). Advection–diffusion equations for generalized tactic searching behaviors. *J. Math. Biol.* (in press); also Advection–diffusion equations for internal state-mediate random walks. *SIAM J. Appl. Math.* (accepted).

Grünbaum, D. & Okubo, A. (1994). Modelling social animal aggregations. In: *Frontiers of Theoretical Biology*, Lecture Notes in Biomathematics, Vol. 100 (Levin, S. A., ed.). Berlin: Springer-Verlag.

Gueron, S. & Levin, S. A. (1993). Self-organization of front patterns in large wildebeest herds. *J. theor. Biol.* **165,** 541–552.

Gueron, S., Levin, S. A. & Rubenstein, D. I. (1996). The dynamics of mammalian herds: from individual to aggregations. *J. theor. Biol.* **182,** 85–98.

Harris, S. (1971). *An Introduction to the Theory of the Boltzmann Equation* 221 pp. New York: Holt, Rinehart & Winston.

Hastings, A. & Powell, T. (1991). Chaos in a three-species food chain. *Ecology* **72,** 896–903.

Hofmann, E. E. (1993). Coupling of circulation and ecosystem models. In: *Patch Dynamics* (Levin, S. A., Powell, T. H. & Steele, J. H., eds), pp. 136–161. Berlin: Springer-Verlag.

Humston, R., Ault, J., Lutcavage, M. & Olson, D. (1998). Schooling and migration of large pelagic fishes relative to environmental cues (in prep).

Huth, A. & Wissel, C. (1990). The movement of fish schools: a simulation model. In: *Biological Motion*, Lecture Notes in Biomathematics, Vol. 89 (Alt, W. & Hoffmann, G., eds), pp. 577–590. Berlin: Springer-Verlag.

Huth, A. & Wissel, C. (1992). The simulation of the movement of fish schools. *J. theor. Biol.* **156,** 365–385.

Inagaki, T., Sakamoto, W. & Kuroki, T. (1976). Studies on the schooling behavior of fish = D1. II: Mathematical modelling of schooling form depending on the intensity of mutual force between individuals. *Bull. Japan. Soc. Sci. Fish.* **42**(3), 265–270.

Kawasaki, K. (1978). Diffusion and the formation of spatial distribution. *Math. Sci.* **16**(183), 47–52.

Keller, E. F. & Segel, L. A. (1971). Model for chemotaxis. *J. theor. Biol.* **30,** 225–234.

Kierstead, H. & Slobodkin, L. B. (1953). The size of water masses containing plankton bloom. *J. Mar. Res.* **12,** 141–147.

Koopman, B. O. (1956). The theory of search. I. Kinematic bases. *Oper. Res.* **4,** 324–346.

Leibovich, S. (1993). Spatial aggregation arising from convective processes. In: *Patch Dynamics* (Levin, S. A., Powell, T. H. & Steele, J. H., eds), pp. 110–124. Berlin: Springer-Verlag.

Levin, S. A. & Segel, L. A. (1976). Hypothesis for origin of planktonic patchiness. *Nature* **259,** 659.

Levin, S. A. & Segel, L. A. (1985). Pattern generation in space and aspect. *SIAM Rev.* **27,** 45 – 67.

Levin, S. A., Morin, A. & Powell, T. H. (1989). Patterns and processes in the distribution and dynamics of Antarctic krill. In: *Scientific Committee for the Conservation of Antarctic Marine Living Resources Selected Scientific Papers Part* 1, *SC-CAMLR-SSP*/5, *CCAMLR* pp. 281–299.: Hobart, Tasmania, Australia.

Lloyd, M. (1967). Mean crowding. *J. Anim. Ecol.* **36,** 1–30.

Lobel, P. S. & Robinson, A. R. (1988). Larval fishes and zooplankton in a cyclonic eddy in Hawaiian waters. *J. Plankton Res.* **10,** 1209–1223.

Mackas, D. D. (1984). Spatial and autocorrelation of plankton community composition in a continental shelf ecosystem. *Limnol. Oceanog.* **29,** 451–471.

Maltrud, M. E. & Vallis, G. K. (1991). Energy spectra and coherent structures in forced two-dimensional and beta-plane turbulence. *J. Fluid Mech.* **228,** 321–342.

Mann, K. H. & Lazier, J. R. N. (1991). *Dynamics of Marine Ecosystems* 446 pp. Cambridge: Blackwell.

Matsuda, K. & Sannomiya, N. (1980). Computer simulation of fish behaviour in relation to fishing gear. *Bull. Japan. Soc. Sci. Fish.* **46,** 689–697.

Matsuda, K. & Sannomiya, N. (1985). Computer simulation of fish behaviour in relation to a trap model. *Bull. Japan. Soc. Sci. Fish.* **51,** 33–39.

Mimura, M. & Yamaguti, M. (1982). Pattern formation in interacting and diffusing systems in population biology. *Adv. Biophys.* **15,** 19–65.

Mollison, D. (1977). Spatial contact models for ecological and epidemic spread. *J. R. Statist. Soc. B* **39,** 283–326.

MOLLISON, D. & LEVIN, S. A. (1995). Spatial dynamics of parasitism. In: *Ecology of Infectious Diseases in Natural Populations* (Grenfell, B. T. & Dobson, A. P., eds), pp. 384–398. Cambridge: Cambridge University Press.

MURRAY, J. D. (1990). *Mathematical Biology* 767 pp. Heidelberg: Springer-Verlag.

NEILL, W. H. (1979). Mechanisms of fish distribution in heterothermal environments. *Am. Zool.* **19,** 305–317.

OKUBO, A. (1974). Diffusion-induced stability in model ecosystems. *The Johns Hopkins Univ. Tech. Report* 86. Chesapeake Bay Institute.

OKUBO, A. (1980). *Diffusion and Ecological Problems: Mathematical Models. Biomathematics*, Vol. 10, 254 pp. Heidelberg: Springer-Verlag.

OKUBO, A. (1986). Dynamical aspects of animal grouping: swarms, schools, flocks, and herds. *Adv. Biophys.* **22,** 1–94.

OLSEN, D. B. & BACKUS, R. H. (1985). The concentrating of organisms at fronts: a cold-water fish and a warm-core ring. *J. Mar. Res.* **43,** 113–137.

OLSEN, D. B. & PODESTA, G. P. (1988). Oceanic fronts as pathways in the sea, In: *Signposts in the Sea* (Hernkind, W. F. & Thistle, A. B., eds), Proceedings of a multidisciplinary workshop on Marine Animal Orientation, May 29–31, 1986, Florida State University, Tallahassee, FL, pp. 1–14.

OLSON, D., HUMSTON, R., PODESTA, G., SAMUELS, G. & LATCAVAGE, M. (1996). Bluefin tuna distribution and migration relative to ocean fronts in New England waters. 1993–95. *Proc. 47th Ann. Tuna Conf.* p. 18, Lake Arrowhead, CA.

OTHMER, H. G., DUNBAR, S. R. & ALT, W. (1988). Models of dispersal in biological systems. *J. Math. Biol.* **26,** 263–298.

PAPOULIS, A. (1984). *Probability, Random Variables, and Stochastic Processes* 576 pp. New York: McGraw-Hill.

PARRISH, J. K. & HAMNER, W. M. (eds) (1997). *Animal Groups in Three Dimensions*. Cambridge: Cambridge University Press.

PASCUAL-DUNLAP, M. (1995). Some nonlinear problems in plankton ecology, 152 pp. Ph.D. Thesis, MIT-WHOI Joint Program in Oceanography.

PATLAK, C. S. (1953). Random walk with persistence and bias. *Bull. Math. Biophys.* **15,** 311–338.

PLATT, T. & DENMAN, K. L. (1975). A general equation for the mesoscale distribution of phytoplankton. *Mem. Soc. R. Sci. Liege* **7,** 31–42.

REYNOLDS, C. W. (1987). Flocks, herds, and schools: a distributed behavior model. *Comp. Graphics* **21,** 25–34.

RING GROUP—BACKUS, FLIERL, KESTER, OLSON, RICHARDSON, WIEBE, WORMUTH (1981). Gulf Stream rings: their physics, chemistry and biology. *Science* **212** (4499), 1091–1100.

ROBINSON, A. R. (ed.) (1983). *Eddies in Marine Science* 609 pp. Berlin: Springer-Verlag.

ROTHSCHILD, B. J. & OSBORN, T. R. (1988). Small-scale turbulence and plankton contact rates. *J. Plankton Res.* **10,** 465–474.

SAKAI, S. (1973). A model for group structure and its behavior. *Biophysics* **13,** 82–90 (Japanese).

SEGEL, L. A. & JACKSON, J. L. (1972). Dissipative structure: and explanation and an ecological example. *J. theor. Biol.* **37,** 545–559.

SEGEL, L. A. & LEVIN, S. A. (1976). Applications of nonlinear stability theory to the study of the effects of dispersion on predator–prey interactions. In: *Selected Topics in Statistical Mechanics and Biophysics*, Conference Proceedings Number 27 (Piccirelli, R., ed.), pp. 123–152. New York: American Institute of Physics.

STOMMEL, H. (1949). The trajectory of small bodies sinking slowly through convection cells. *J. Mar. Res.* **8,** 24–29.

SUNDERMEYER, M. A. (1995). Mixing in the North Atlantic tracer release experiment: observations and numerical simulations of Lagrangian particles and passive tracer. M.S. Thesis, Dept of Earth, Atmos. and Planetary Sci., MIT, Cambridge, MA.

TAYLOR, G. I. (1921). Diffusion by continuous movement. *Proc. London Math. Soc. Ser.* 2 **20,** 196–212.

TURCHIN, P. (1989). Population consequences of aggregative movement. *J. Anim. Ecol.* **58,** 75–100.

VAN KAMPEN, N. G. (1992). *Stochastic Processes in Physics and Chemistry* (revised and enlarged edition) 464 pp. Amsterdam: North-Holland.

WARBARTON, K. & LAZARUS, J. (1991). Tendency-distance models of social cohesion in animal groups. *J. theor. Biol.* **150,** 473–488.

YAMAZAKI, H. & OSBORN, T. R. (1988). Review of oceanic turbulence: implications for biodynamics. In: *Towards a Theory on Biological-Physical Interactions in the World Ocean* (Rothschild, B. J., ed.), pp. 215–234. Dordrecht: Kluwer.

ZAR, J. H. (1996). *Biostatistical Analysis* 3rd Edn. Upper Saddle River, NJ: Prentice-Hall.