

The use of radial distribution and pair-correlation functions to analyze and describe biological aggregations

Kelly Younge³, Brendan Johnston³, Cory Christenson³, Amit Bohara³, Joshua Jacobson³, Nancy M. Butler^{1*}, and Paul Saulnier²

¹Kutztown Univ. of Pennsylvania, P.O. Box 730, Kutztown, PA 19530, USA.

²Gustavus Adolphus College, 800 West College Avenue, Saint Peter, MN 56082, USA.

³All other authors listed were undergraduate research associates at Gustavus Adolphus College.

Abstract

Patterns are pervasive in nature with many examples being found in both living and inanimate systems. While researchers recognize the importance of the behavior of individuals to the structure and shape of an aggregation, a major hurdle in describing aggregated organisms has been the difficulty of tracking the movement of individuals over time. Here we present an innovative application of an analytical technique derived from statistical mechanics (a subfield of physics) to describe the spatial distribution of grouped organisms. Radial distribution and pair-correlation functions are traditionally used by physicists to describe inert particle dynamics. This novel biological application allows one to infer the behavioral characteristics of individuals within a group based solely on the spatial distribution of the aggregate population. Additionally, the method allows one to determine the correlation length, the average maximum distance over which one individual may exert an influence on another member of the aggregation. The analytical technique presented here is also important in that it minimizes two problems that typically plague studies of grouped organisms: it eliminates the need to track the movements of individuals, and it partially takes into account the presence of occluded individuals. This technique also permits quantitative comparison between aggregations formed under various environmental and/or experimental conditions. Thus, this technique may be of value to resource managers, ecologists, and others working with grouped organisms (e.g., plankton swarms, schooling fish, flocking birds, or migratory mammals) who seek to gain information about factors influencing the structure and behavior of such groups.

Nature is characterized by aggregations of organisms. Sometimes these aggregations are temporary, such as birds congregating in a noisy flock on a birdfeeder and then scattering when the food has been depleted. Others are more long lasting, such as the herds of caribou that migrate across the Arctic. Some aggregations occur on scales of millimeters, such as the clusters of flagellated algae that form within the water column of a lake or ocean, while others occur on much larger spatial scales, such as schools of fish covering hundreds of meters. All of these aggregations, however, have one thing in common—the formation and structure of the aggregation is the result of the collective behaviors of the individuals.

Traditionally, there are two primary methods that scientists use to gain insight into the dynamics of grouped organisms:

*Corresponding author

Acknowledgments

The Gustavus Adolphus College authors are pleased to acknowledge partial support from the National Science Foundation under Award No. 0244681 and from the Petroleum Research Fund, administered by the American Chemical Society.

through predictive (i.e., theoretical) mathematical models that attempt to simulate the patterns observed in nature and through data analysis techniques that may be applied to the actual patterns observed in nature.

Predictive models have been developed to simulate such characteristics of grouped organisms as swarm cohesiveness and size (Domeier and Collin 1997; Gazi and Passino 2004), how individuals are positioned within a group (Breder 1954; Pfister 1990; Stevens 1990; Dill et al. 1997; Grimm and Berger 2003; Lutscher 2003; Mogilner et al. 2003), how individuals move within a group (Huth and Wissel 1990; Yamazaki and Kamykowski 2003), the role of diffusion and transport processes in swarm formation (Okubo and Levin 2001; Alt and Hoffmann 1990), the effects of local and non-local swarm interactions, and the influences of external drivers on group structure and dynamics (Fish et al. 1991; Hamner and Parrish 1997; Ritz 1997; Caparroy 2003; Lin et al. 2003; Skillingstad 2003). In many of these cases, the parameters present in the mathematical model are adjusted until a simulation is produced that mimics the patterns observed in nature. The optimized values of these adjustable parameters are then used to

infer the behavioral characteristics of the individuals comprising the aggregation.

There have also been many data analysis techniques developed over the years that are used to quantitatively describe the observed patterns in nature. These include variance-to-mean ratio (David and Moore 1954), the negative binomial parameter (Waters 1959), and Lloyd's indices of mean crowding and patchiness (Lloyd 1967), (see Pielou 1977 for a more extensive review). More sophisticated analysis techniques have also been developed. These include measurements based on nearest-neighbor distances (Kennedy and Crawley 1967) and the tracking of individuals within an aggregation (Turchin 1996; Wiens *et al.* 1995), to name a few (see Parrish and Hamner 1997 for further discussion). Many of these techniques, while helpful, are limited in that a single parameter is often used to describe the whole aggregation, leading to the possibility that internal sub-structure within the group may be overlooked.

A major challenge in developing experimental analysis techniques that describe the behavior and characteristics of swarms in nature lies in how to assess the role of individual swarm members. One hurdle in describing the role of the behavior of individual group members in determining the structure and shape of an aggregation has been the difficulty of recording the movement of individuals over time (Ikawa and Okabe 1997; Jaffe 1997; Osborn 1997). Investigators have successfully used three-dimensional (3D) acoustic imaging to record the movements of individual zooplankters and fish *in situ* (McGehee and Jaffe 1996; Greene and Wiebe 1997; DeRobertis *et al.* 2003) while others have successfully used 3D video imaging in the lab and *in situ* (Parrish and Turchin 1997; Yen and Bundock 1997; Banas *et al.* 2003; Benfield *et al.* 2003). Another hurdle lies in analyzing the movement of the individuals in these recordings. Although researchers have traditionally relied upon manual frame-by-frame analysis of the film to note the x - y and y - z coordinates of each, investigators are now turning to image processing software to obtain these data (Benfield *et al.* 2003).

Here we present an analytical data analysis technique that can be used to describe aggregations and aid in the development of descriptive behavioral models. The power of this technique is that it does not require tracking each member of an aggregation; rather it provides a quantitative description of the structure of the aggregation from which the dynamic behaviors of individual members may be inferred. Furthermore, as it provides information about all locations within an aggregation (and not just a single number used to describe the whole aggregation), it enables one to perceive, describe, and compare patterns or swarm sub-structures that are not immediately obvious visually. Another advantage of this technique is its ability to partially compensate for occluded individuals (see below for a further discussion). Thus, besides providing information that other data analysis techniques do not, the technique presented here minimizes two problems that typically plague experimental studies of grouped organisms: it

eliminates the need to track the movements of individuals and it at least partially takes into account the presence of occluded individuals.

The technique presented here uses radial distribution and pair-correlation functions to quantitatively describe the structure and dynamics of an aggregation. Although these functions and the associated analytical techniques have traditionally been used by physicists to study inert particles, we will demonstrate here their utility in the study of grouped organisms.

The radial distribution function provides information on how individuals are distributed across increasing distances from a central individual in an aggregation. It yields a local measure of how close the observed distribution is to a uniform one (i.e., a distribution where all positions are equally likely). The pair correlation function, on the other hand, provides a measure of the spatial variation of the influence that one group member exerts on others in terms of both the magnitude and direction (i.e., a positive or negative influence) of this spatially varying influence. Together these functions also provide information about the average distance over which one swarm member may influence another, referred to as the correlation length.

As an example of the information provided by these functions, consider a large room (e.g., a gymnasium) filled with people, all wearing blindfolds and walking in random directions throughout the room. Viewed from above, the people appear to be a dynamic two-dimensional swarm. This "swarm" could be photographed and analyzed using radial distribution and pair-correlation functions. These functions may be used to infer the behavioral characteristics of a typical swarm member as well as its influence on other members of the swarm. For example, in this case, because all members are blindfolded, they are able to sense and respond to the presence of another individual in the swarm only when they come in direct contact with each other, a fact that would be reflected in the resulting analyses. These analyses would also provide information on indirect interactions between swarm members. When two group members collide, the collision affects not only their motion but also that of their neighbors, which will, in turn, influence the motion of more distant swarm members. The range, strength, and direction (positive or negative) of this indirect influence depends not only on the number of people in the "swarm" but also on their common behavioral characteristics (e.g., in this case, the fact that all of the people are blindfolded). The radial distribution and pair-correlation functions together provide a measure of how this influence or correlated behavior varies with position within the swarm. Locations within the swarm where the correlation values are negative indicate regions that are less likely to be occupied by a swarm member, while locations with positive correlation values are more likely to be occupied.

Now imagine how swarm structure and dynamics would change if we repeated this exercise using people who were not blindfolded. As in the example above, the group members are

walking about the room in random directions. However, in this case, because they now have the benefit of sight, each group member will try to maintain a “personal space” by avoiding collisions before they take place. Whereas in the previous example, an individual’s effective space (the distance over which direct interactions take place) was equal to the individual’s actual physical size, in this second scenario, the effective size of an individual is greater than their actual physical size. The effect of this increased size will propagate outward from an individual through interactions with neighbors such that the range and strength of the influence of a single individual will be greatly enhanced. Thus, because of differences in the behavioral characteristics of the group members, a low density swarm whose members can perceive and react to the presence of other members from a distance may behave like a much higher density swarm whose members are blindfolded and can only perceive and interact with other members through direct contact. The exact shape of the radial distribution and pair-correlation functions for sighted swarms would depend on such factors as their visual acuity and their sense of acceptable “personal space.” These are factors that may be influenced by environmental conditions. For example, if a fire suddenly broke out in one corner of the room, the people would rapidly move away from it, and their notion of acceptable “personal space” would be greatly altered as they crowd around the exit. This would have a profound influence on the corresponding radial distribution and pair-correlation functions and the related swarm correlation length.

Radial distribution and pair-correlation functions have many advantages over other techniques used to analyze grouped organisms. Researchers can infer behavioral characteristics of individuals based on group structure alone. As mentioned above, the position histories—i.e., $x(t)$, $y(t)$, $z(t)$ —of individual swarm members are not required to conduct the analysis; rather, global changes in group structure in response to external stimuli are used to infer the behavioral characteristics of the individuals.

One traditional technique used to describe the distribution of individuals within a group is nearest-neighbor analysis. While this technique is useful in providing general information about the average distance between individuals within the group, it provides little detailed information about the structure of the aggregation, such as whether all members are evenly distributed, the distance over which one individual may influence the behavior of another, the spatial variation in both magnitude and direction (positive or negative) of this influence, the conditions under which this influence is greatest (or least), and the presence of regions within the aggregation that are qualitatively different from other regions. Analyses using radial distribution and pair-correlation functions provide a means of quantitatively elucidating such group characteristics.

Another significant advantage of the proposed analysis technique is its ability to help minimize the occluded animal problem. In an aggregation, some members will block a cer-

tain fraction of the population from the researcher’s (or camera’s) view. Although occluded animals are not seen in a particular video frame, their influence is nevertheless reflected in the behavior of neighboring animals whose locations can be noted. This influence affects the swarm structure, which, in turn, is reflected in the corresponding radial distribution and pair-correlation functions. Thus, the effect of the presence of occluded animals is at least partially taken into account by the very nature of the analysis technique.

Materials and procedures

Radial distribution and pair-correlation functions may be computed for a variety of systems exhibiting group behavior by using videographic analysis. The basic procedure consists of recording a time-lapse video of the system to be studied (living or inanimate), extracting group member locations from each frame, computing the radial distribution function for each frame, and finally, averaging these results over many frames. As noted above, the frame-to-frame tracking of each animal comprising the swarm is not required as each frame is treated as an independent spatial configuration.

When implementing the described method, aggregations should be filmed in such a way as to provide a high degree of contrast between the group members and the background and at a speed sufficient to ensure that each frame presents a sufficiently distinct spatial configuration of the aggregation. Thus, the time between frames should be adequate to allow a typical swarm member to move a non-negligible distance (e.g., a distance equivalent to approximately 5% of the aggregation’s diameter). Filming should extend over a length of time sufficient to produce a number of unique frames that may be averaged so as to result in a smooth radial distribution function (typically a few hundred to a few thousand frames). The number of frames recorded and analyzed may be increased as desired so as to provide an arbitrarily small random variance in the resulting radial distribution function (i.e., arbitrarily small statistical noise). The number of frames required to achieve a desired noise level varies depending on swarm density; with low density swarms requiring that more frames be averaged, because of the statistics of small numbers present in such swarms. In all the experimental examples presented in the Assessment section, a sufficient number of frames were averaged so that the noise fluctuations present in the resulting radial distribution function graphs were less than about 3%.

The collected video is broken into individual frames using any one of the many commercially available products designed for such a purpose, with each frame saved as a separate file for subsequent analysis. (We used VirtualDub version 1.5.10, see www.VirtualDub.org). The coordinates describing the location of each group member can be determined for each frame manually by clicking on each with a mouse pointer, automatically using custom software, or by using a commercial software package. We wrote custom software to automatically extract animal locations from each frame using the following methodology.

A video frame is read into the computer from its individually saved file and each pixel in the image is assigned a value based on its video brightness and the weighted brightness of the pixels surrounding it. The surrounding pixels' brightness values are weighted by the inverse square of their distance from the pixel under consideration (at $[x,y]$) and that pixel is assigned a value using the following algorithm:

$$PV(x,y) = VB(x,y) + \sum_{i=-1.5R}^{1.5R} \sum_{j=-1.5R}^{1.5R} \frac{VB(x+i,y+j)}{[i^2 + j^2]} \quad (1)$$

where $PV(x,y)$ is the pixel value assigned to the location (x,y) , $VB(x,y)$ is the video brightness of the pixel at location (x,y) , R is the average radius of the animals under study (in pixels), and the sums on i,j take on integer values over the range indicated. In this way, the centers of individuals will have the highest pixel values (as they have the greatest number of bright pixels surrounding them) and will thus be correctly identified and their location extracted.

Once the positions of all group members in the frame are determined, these data are then used to compute the radial distribution function $g(\mathbf{r})$ using the following conceptual definition:

$$g(\mathbf{r}) = \frac{\text{Observed probability of finding an animal in a given swarm sub-region}}{\text{Probability of finding an animal in that region if all were uniformly distributed}} \quad (2)$$

where \mathbf{r} represents a position vector from an animal near the center of the swarm (designated as the coordinate origin) to a small sub-region within the swarm. Thus, the radial distribution function is a local measure of how close the observed animal distribution is to a uniform distribution (i.e., a distribution where all animal locations are equally likely). A $g(\mathbf{r})$ value of one (unity) implies that the probability of finding any particular animal in the region under consideration is equivalent to that which would be obtained if the animals were uniformly distributed, whereas $g(\mathbf{r})$ values greater/less than unity imply an enhanced/reduced probability relative to a uniform distribution. The conceptual Eq. 2 may be represented mathematically as (see Pathria 1986),

$$g(\mathbf{r}) = \frac{\left(\frac{dN}{N}\right)}{\left(\frac{dV}{V}\right)} \quad (3)$$

where dN is the actual (i.e., observed) number of animals located in the swarm sub-volume dV at position \mathbf{r} and where N is the total number of animals located in the entire swarm volume V . This process is repeated for all sub-regions that make up the swarm so that a radial distribution function value at all swarm locations is obtained. Each frame is similarly analyzed and the resulting $g(\mathbf{r})$ averaged to produce a smooth curve.

The pair-correlation function $v(\mathbf{r})$ is given as $v(\mathbf{r}) = g(\mathbf{r}) - 1$. This function tells the researcher the strength and the direction of the influence that the animal at the origin exerts on more distant animals. If $v(\mathbf{r})$ is greater/less than zero in a par-

ticular region then the influence of the animal at the origin is positive/negative and that region is favored/avoided. Conversely, if $v(\mathbf{r})$ is zero in a particular region, the motion of the animals in that region is not correlated in any way with the animal at the origin.

Finally, it should be noted that the choice of which animal is designated as the origin is arbitrary as long as they are sufficiently far from the swarm's edge so as to avoid any finite size/edge effects. Thus, the resulting radial distribution and pair-correlation functions are independent of the choice of origin and depend only on the animal characteristics and environmental conditions.

Assessment

As a proof of concept and utility demonstration, we applied the method described in this paper to three two-dimensional systems, one inanimate and two living. We will describe each experiment and the associated results in turn.

This method was first used in an inanimate application involving small metallic spheres of radius R (standard copper BBs) (Younge, et al. 2004). We placed the BBs at the bottom of a Petri dish and used a vibrating platform to induce random motion of the BBs in a two-dimensional plane. Five different densities of BBs were tested, representing a range of the percent area covered by the BBs (i.e., area fractions) from 10% to 50%. A digital video camera was positioned above the cell and used to record the motion of the BBs using time-lapse photography. The video camera recorded six frames/s for approximately 20 min, providing thousands of unique frames for analysis. Typical single video frames are shown in Fig. 1a-e; the corresponding time-lapse videos may be viewed at <http://physics.gac.edu/~psaul/swarm/>.

As discussed above, a commercial software package was used to separate each video clip into individual video frames. For each frame, the BB centers were extracted using our custom algorithm and the radial distribution function calculated. The process of computing the radial distribution function for any system consists of subdividing the system into a large number of small regions and applying Eq. 3 to each of these regions. For this system, the aggregation was mentally divided into many thin concentric annular rings (centered on the BB chosen as the origin). These thin rings constitute the sub-volumes or sub-regions to which Eq. 3 is applied. The implementation of Eq. 3 for this system consists of determining the number of BBs found in the particular annular ring being considered (dN) and then dividing this value by the total number of BBs in the aggregation (N). This quotient represents the numerator of Eq. 3 and is then divided, according to that equation, by the ratio of the annular ring volume (dV) (actually the ring area in this case as we are considering a two-dimensional example) to the total volume (V) (total area of the cell in this case). The result of this computation is the radial distribution function $g(r)$ for the ring/sub-region located a distance r from the BB at the center of the aggregation. This calculation is repeated for each annu-

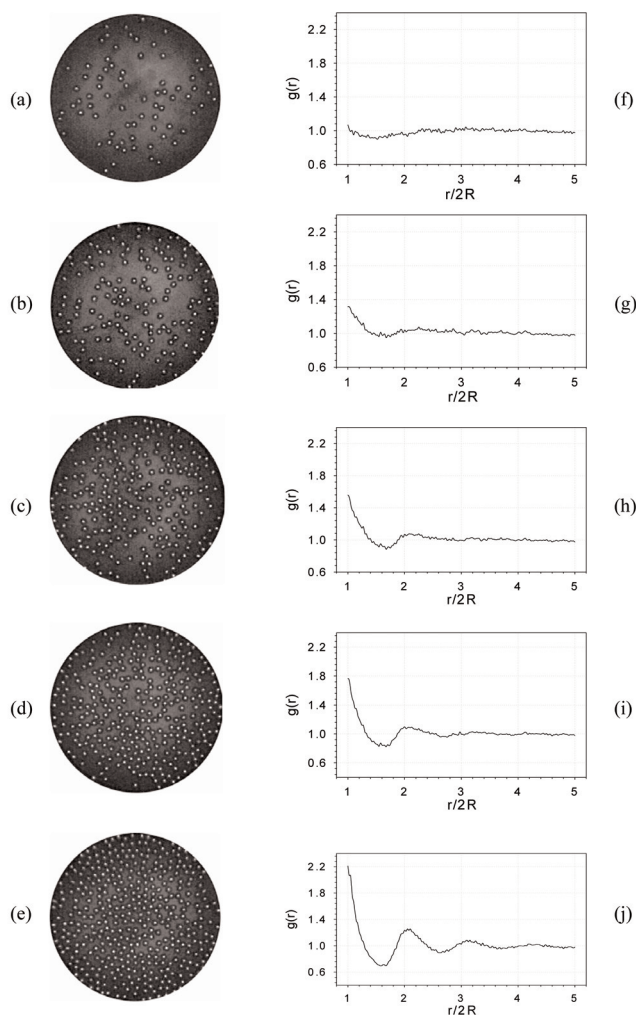


Fig. 1. Representative video frames showing the spatial distribution of BBs at (a) 10%, (b) 20%, (c) 30%, (d) 40%, and (e) 50% area fractions. The bright spots represent the center of the BBs against the dark background of the sample cell. The coordinates of the BB centers were extracted, and the radial distribution functions for each area fraction were calculated as described and are shown to the right: (f) 10%, (g) 20%, (h) 30%, (i) 40%, and (j) 50%. These radial distribution functions $g(r)$ are plotted versus the distance r from a central BB. This distance is scaled $2R$, by the diameter of a BB sphere. See text for a detailed discussion of the significance of these results.

lar ring in every video frame. These computations were performed on a computer, and the resulting radial distribution functions for each frame were averaged. The experimentally obtained radial distribution functions (RDFs) $g(r)$ for the two-dimensional hard sphere BB “swarms” with area fractions ranging from 10% through 50% are shown in Fig. 1 (f. 10%, g. 20%, h. 30%, i. 40%, j. 50%). These radial distribution functions are plotted versus the distance r from a central BB. This distance is scaled by $2R$, the diameter of a BB sphere.

To understand the significance of the results depicted in Fig. 1, one must recall the conceptual definition of $g(r)$,

namely, that $g(r)$ is a local measure of how closely a distribution resembles a uniform one. A value of unity for $g(r)$ indicates a uniform distribution of objects (inanimate or living) while values that differ from unity indicate an enhanced or suppressed probability relative to the uniform distribution. It is noted (see Fig. 1f) that $g(r)$ is approximately 1 for all distances from the central BB in the 10% area fraction sample, while $g(r)$ differs considerably from unity for the 50% area fraction sample (see Fig. 1j). One can conceptually understand these results by considering a process in which a collection of inert objects is assembled. As the first object (BB) is placed into an empty sample cell it is free to be located at any position (i.e., all locations are equally likely, corresponding to a uniform probability distribution), while the next BB added is free to be located anywhere except within a distance of $2R$ of the first BB. This process continues until all N particles (BBs) have been added to the sample cell. At each successive step in this process, the free area available to the particle being added is less than that which was available to the previous particle. Hence, its location is restricted and is therefore not given by a uniform distribution but is dependent on the positions of all the preceding particles. Thus, as the area fraction of particles in the system increases, this excluded area effect becomes more pronounced, and the influence of the particle at the origin extends (through intermediaries) to increasingly greater distances, causing the RDF to deviate from unity well beyond the central particle’s nearest neighbors. This explanation is consistent with the experimentally observed increased deviation from unity of the RDF as the area fraction increases. This indicates that as the location of any BB becomes more restrictive (due to the presence of other BBs) the spatial distribution of BBs becomes less uniform in character.

Furthermore, the oscillatory behavior of the radial distribution function for the 50% area fraction sample may also be explained by a simple argument. This oscillatory behavior implies that there is an alternation between enhanced and reduced probability of finding a BB at various distances r from the central BB. The regions of enhanced probability, $g(r)$ greater than unity, represent the “natural” or favorable locations within a collection of particles while locations with $g(r)$ less than unity represent unfavorable BB locations. Fig. 2 depicts the approximate “natural” particle positions located at distances from the particle at the origin of $2R$, $4R$, $6R$, and $8R$, etc. These “natural” particle positions arise from the fact that BBs at these locations experience, on average, a reduced frequency and strength of interactions with their neighbors. Particles (BBs) in the first nearest neighbor ring (at the approximate distance of $2R$ from the origin, $r/2R \approx 1$) cannot get any closer to the origin due to the hard spherical BB already present there. Similarly, the particles in the second nearest neighbor ring (at approximate distance of $4R$, $r/2R \approx 2$) would impinge upon the particles in the first neighbor ring should they attempt to move closer to the origin. This fact leads to the observed dip in the RDF at $r/2R \approx 1.5$. Thus, the presence of naturally occur-

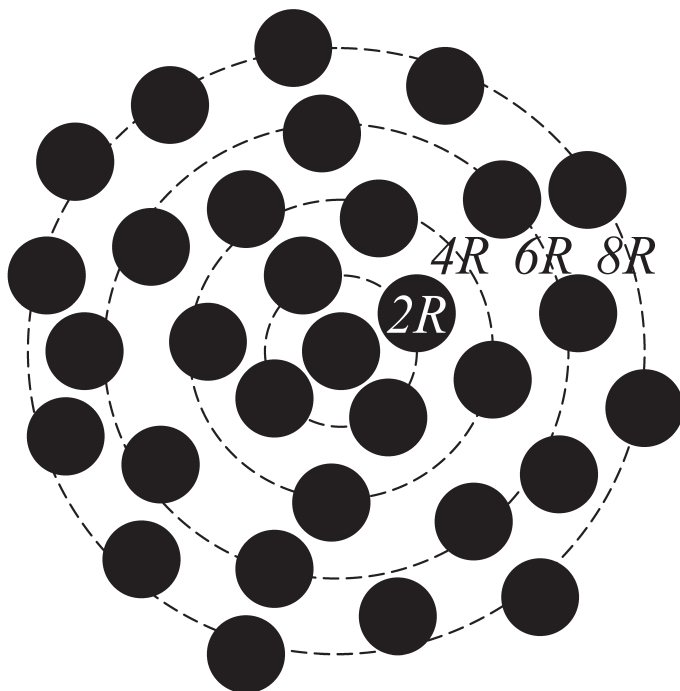


Fig. 2. The concentric circles of even number radii indicate the approximate “natural” positions of hard spheres (BBs) in an aggregation. These locations correspond to peaks in the radial distribution function. Particles are more likely to be found at approximate distances of $2R$, $4R$, $6R$, etc. from the particle (BB) at the origin due to the excluded area/volume effect leading to naturally occurring nearest neighbor rings. Conversely, they are less likely to be found in the intermediate regions due to an increased frequency of collision at these locations.

ring nearest neighbor rings is directly observed as oscillations in the radial distribution function.

While the excluded area (or volume) effect is a fundamental type of interaction between inert particles, many other types of interactions exist and may greatly influence the observed radial distribution function. By studying the radial distribution function of systems, insight into the underlying interactions may be gained and compared with systems prepared under different conditions. The radial distribution functions obtained in the BB study are consistent with hard sphere interactions and, in fact, knowing only the BB radius one could infer (from the radial distribution function alone) that the interactions between these spheres only occur during a collision (i.e., no long range direct interaction). Thus, the global structure, represented by the radial distribution function, may be used to infer the behavior of the individual objects making up an aggregation. It is precisely this fact that makes these analytical methods so attractive for biological applications (see below for the results of the biological studies).

Finally, another question that may be addressed by using the radial distribution function is, over what distance does the influence of one swarm member affect the behavior of another (the so-called correlation length)? The answer to this

question can most easily be obtained by using the pair-correlation function, $v(r)$, defined as,

$$v(\mathbf{r}) = g(\mathbf{r}) - 1 \quad (4)$$

One can see that, with this definition, when $g(r) = 1$, indicating a uniform distribution, $v(r) = 0$, stating that no correlations exist among the swarm members (i.e., the behavior of one member does not influence any other). Spatial regions where $g(r)$ is greater/less than unity imply regions of positive/negative correlation where the influence of the member at the origin is positive/negative. Thus, the correlation length, the average distance over which the swarm member at the origin influences others, is just the distance required for the radial distribution function to attain a value of unity. For example, this happens at about $8R$ ($r/2R \approx 4$) for our 50% area fraction hard sphere sample, implying that a BB in this sample influences its neighbors up to a distance of four sphere diameters away.

As mentioned above, the methods presented in this paper have significant potential application to biological systems. We have demonstrated this fact by applying the described method to aggregations of brine shrimp and ostracods. In both cases the animals were confined to a two-dimensional space for simplicity. The method was applied in the standard fashion, namely, videotaping the animals' motion, extracting their locations in the video frame via an automated procedure, computing the radial distribution function (in the same way as was done for the BBs), and averaging over many frames. Example videos may again be found at <http://physics.gac.edu/~psaul/swarm/> and are discussed below.

The results were dramatic as the radial distribution function for the BBs was consistent with that expected for particles that interact via collisions only (refer again to Fig. 1), while the radial distribution function for the brine shrimp exhibited a radically different character, consistent with objects whose interaction is long range and repulsive (see Fig. 3). The behavior of the ostracods leads to a radial distribution function that is consistent with strong short-range attractive interactions (see Fig. 4).

It is clear that the radial distribution function curves obtained from the brine shrimp aggregation are very different from those of the BBs (compare the curves obtained from the 10% aggregation of each). Specifically, for the 10% area fraction shrimp swarm, we see that there is a region of significantly decreased likelihood of finding a shrimp ($g(r) < 1$) out to approximately three shrimp diameters from the reference shrimp at the origin. This implies that the shrimp try to avoid being within close proximity to each other, a negative correlation, $v(r) = g(r) - 1 < 0$. After this distance, however, the RDF is fairly constant at unity, which implies that the distribution of shrimp is uniform, and no positions are more likely than others, meaning that the influence of the central shrimp on those distant shrimp is no longer significant, $v(r) = g(r) - 1 \approx 0$. The eventual leveling of the RDF to unity is a common feature

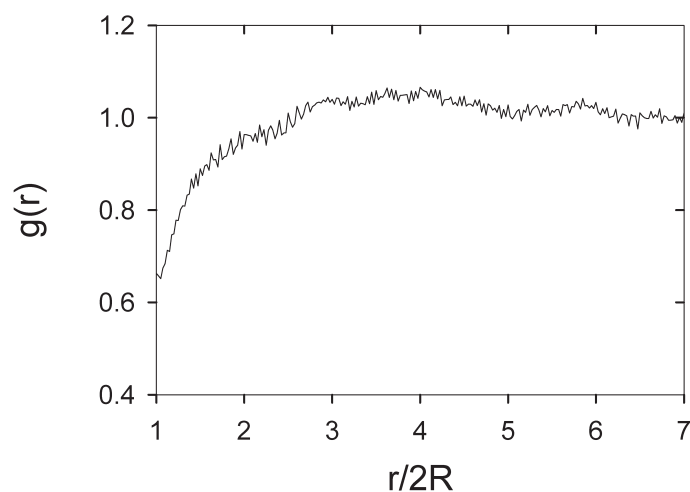


Fig. 3. The experimentally obtained radial distribution function $g(r)$ for a two-dimensional 10% area fraction swarm of brine shrimp. The radial distribution function is plotted versus distance r from a central shrimp. This distance is scaled by $2R$, the average diameter of a shrimp. This result indicates that brine shrimp exhibit a strong repulsive behavior ($g(r) < 1$) even at low densities (as similar results were found for 4% area fraction shrimp swarms as well) and that this repulsive behavior extends to approximately three shrimp diameters away ($r/2R=3$) at 10%.

for any system that only possesses short-range order. Thus, we are able to infer that the behavioral characteristics of brine shrimp, as individuals, based solely on the observed behavior of the aggregation (i.e., brine shrimp do not like to be in close proximity to each other and a single shrimp will influence others out to a distance of three shrimp diameters).

It is clear from Fig. 4 that the radial distribution function obtained from the ostracod aggregation is significantly different from that obtained in the study of the BBs or brine shrimp, implying that their behavior (as individuals) is also very different. We see that this distribution exhibits a sharp spike at a distance of one ostracod diameter from the central animal (at $r/2R \approx 1$ with a peak height of 2.0 for the 9.7% aggregation). The radial distribution function in this region implies that there is a strong affinity for ostracods to spend a significant amount of their time next to each other. This spike is followed by a broader dip, which arises due to the fact that other ostracods cannot easily be in this region lest they overlap the nearest neighbor ostracod ring already present there. As before, the radial distribution curve levels off to one as we examine regions farther from the central animal, as its influence diminishes, $v(r) = g(r) - 1 \approx 0$. It should also be noted that ostracods strongly interact with each other over a much shorter distance than the shrimp both relative to their size and in absolute terms.

Thus, our proof of concept work clearly demonstrates that the described method may be applied to inanimate or living systems and that the resulting radial distribution functions (derived from the aggregation as a whole) can be used to infer the behavioral characteristics of the individuals comprising

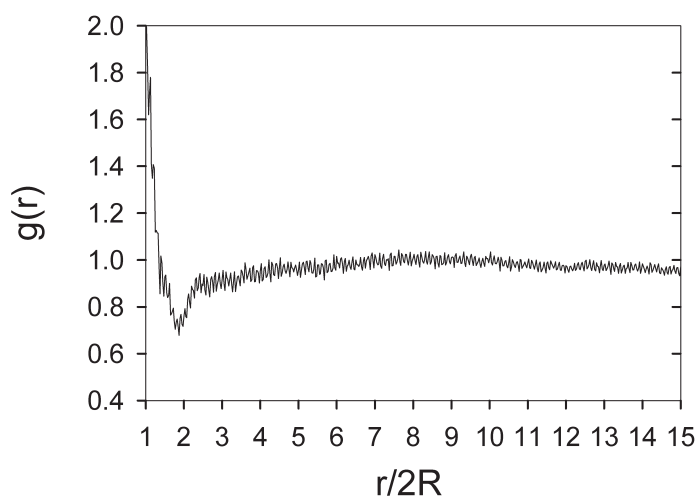


Fig. 4. The experimentally obtained radial distribution function $g(r)$ for a two-dimensional 9.7% area fraction swarm of ostracods. The radial distribution function is plotted versus distance r from a central ostracod. This distance is scaled by $2R$, the average diameter of an ostracod. This result indicates that ostracods exhibit a strong short-range attractive behavior as $g(r)$ shows a large initial spike ($g(r) > 1$). This spike (indicating the strong likelihood of other ostracods in close proximity) is followed by the typical excluded volume depletion zone, indicated by the adjacent dip in the RDF.

the swarm, leading to significant application for biological systems. Furthermore, these results are unique in that they are not duplicated by other data analysis techniques currently being employed by researchers.

Discussion

Scientists who study grouped organisms are often interested in understanding not only the placement of the organisms in space, but also their placement relative to other members of the group and the dynamic nature of their placement. They often are also interested in how aggregations respond to variation in environmental conditions.

For objects that are relatively static (such as trees in a forest or rocks in a stream bed), describing the location of individual objects within the group is quite straightforward. However, when dealing with large groupings of moving objects, this task becomes more difficult, and is made more so by the dynamic nature of the interactions between organisms within the group as well as between the organisms and their surrounding environment.

The analytical technique presented here reduces or eliminates many of the problems that plague investigators seeking to understand the dynamic nature of grouped objects. As demonstrated herein, the radial distribution function may be used to provide a quantitative description of the amount of order present in a system for both inanimate and living aggregations. Furthermore, it allows the investigator to study the magnitude of the influence that one object exerts on others within an aggregation as a function of distance $v(r)$. The average maximum distance over which one object influences

another (i.e., the correlation length) may also be obtained. Through the study of radial distribution functions the behavioral characteristics of the individual members of an aggregation may be inferred and the underlying physical or biological reasons investigated. For example, the influence of environmental cues on aggregation behavior may be studied using radial distribution functions. Finally, and most significantly, this technique does not require that the movements of individual members of the aggregation be tracked. Eliminating the need to track the movements of individuals within aggregations allows investigators to study not only small clusters of organisms, but also groups with large numbers of individuals, such as phytoplankton swarms or schools of fish.

One can apply this technique in an experimental setting to determine the relative importance of different environmental factors in determining swarm structure. For example, if the presence of a predator affects swarm formation and dynamics, one would expect to see changes in the radial distribution and pair-correlation functions reflecting that influence. Similarly, if light intensity, habitat structure, or population size affects swarm formation and dynamics, one would again expect changes in the radial distribution and pair-correlation functions as each of those factors are altered. And the relative importance of each of these factors in determining swarm formation and dynamics can be determined by comparing the degree to which the radial distribution and pair-correlation functions change for each factor. This technique can also be applied to increase our understanding of the impact of such factors as chemical pollutants or increasing water temperatures on swarm formation and dynamics. It is important to note that, while this technique allows one to perceive subtle patterns in nature, it does not in and of itself identify the underlying mechanisms by which these patterns develop. However, once a factor has been identified as influencing swarm formation and structure, further experimentation can be undertaken to elucidate the mechanism by which the observed response is manifest. For example, manipulating water clarity or light levels might provide insight into the role of vision with respect to how individuals position themselves relative to each other. The role of kairomones in swarm structure and function might be explored by comparing radial distribution and pair-correlation functions of the swarm when water has been added from an aquarium containing predators to those of the swarm when water is added from a tank without predators. Thus, analysis of aggregations through the use of radial distribution and pair-correlation functions enable researchers to perceive subtle swarm sub-structures for which they can then design experiments to elucidate the underlying biological mechanisms responsible for such structures. Hence, resource managers, ecologists, evolutionary biologists, and others interested in how biotic and abiotic factors influence the structure and behavior of aggregations may find this technique to be particularly useful.

Comments and recommendations

Most swarms that occur in nature exist in three-dimensional space. As stated earlier, the radial distribution function technique eliminates the need to track individual swarm members from frame-to-frame and thus removes one major challenge. To get a sense of the three-dimensional character of the swarms under study, they may be filmed from orthogonal directions (x - y and y - z). These cameras may be time-correlated to allow later simultaneous viewing of each orthogonal direction. While filming in three dimensions presents certain technical challenges, most of these can be successfully overcome in the design of the filming protocol. For example, researchers have successfully recorded in three dimensions the movement of particles past tethered copepods (e.g., Fields and Yen 1993; van Duren and Videler 2003) and the movement of free-swimming zooplankton (e.g., Doall et al. 1998, 2002; Seuront et al. 2004).

Two perceptual challenges that three-dimensional swarms present are occluded individuals and the lack of photographic depth perception. The occluded individual problem only becomes significant for swarms containing a high density of animals while the depth perception problem is always present. Both of these difficulties may be overcome by using an illumination scheme whereby only a single plane or "slice" of the swarm is illuminated at any given time. The illuminated plane may then be translated to construct the three-dimensional character of the swarm. The light used for this illumination will typically be outside the perceptual range of the population under study (typically the near-infrared). The required infrared sources (e.g., low-power laser diodes) and light sheet generating optical elements are inexpensive and readily available.

Finally, it should be noted that once the (x , y , z) positions of animals comprising a three-dimensional swarm are known, the computation of the corresponding radial distribution and pair-correlation functions present no more difficulty than a two-dimensional example. The main difference between a two-dimensional and a three-dimensional radial distribution function computation is that the swarm sub-regions to which Eq. 3 is applied are small swarm sub-areas or sub-volumes, respectively. The remainder of the computation is identical regardless of the dimensionality of the space in which the swarm lives.

References

- Alt, W., and G. Hoffmann [eds.]. 1990. Biological motion. Springer-Verlag.
- Banas, N. S., D.-P. Wang, and J. Yen. 2003. Experimental validation of an individual-based model for zooplankton swarming, p. 161-180. *In* L. Seuront and P. Strutton [eds.], Handbook of scaling methods in aquatic ecology. CRC Press.
- Benfield, M. C., C. J. Schwehm, R. G. Fredericks, G. Squyres, S. F. Keenan, and M. V. Trevorrow. 2003. Measurement of zooplankton distributions with a high-resolution digital camera system, p. 17-30. *In* L. Seuront and P. Strutton [eds.],

- Handbook of scaling methods in aquatic ecology. CRC Press.
- Breder, C. 1954. Equations descriptive of fish schools and other animal aggregations. *Ecology* 35:361-370.
- Caparroy, P. 2003. Discrete events-based Lagrangian approach as a tool for modeling predator-prey interactions in the plankton, p. 559-573. *In* L. Seuront and P. Strutton [eds.], *Handbook of scaling methods in aquatic ecology*. CRC Press.
- David, F. N., and P. G. Moore. 1954. Notes on contagious distributions in plant populations. *Ann. Bot. (Lond.)* 18:47-53.
- De Robertis, A., C. Schell, and J. S. Jaffe. 2003. Acoustic observations of the swimming behavior of the euphausiid *Euphausia pacifica* Hansen. *J. Mar. Sci.* 60:885-898.
- Dill, L. W., C. S. Holling, and L. H. Palmer. 1997. Predicting the three-dimensional structure of animal aggregations from functional considerations: The role of information, p. 207-224. *In* J. K. Parrish and W. M. Hamner [eds.], *Animal groups in three dimensions*. Cambridge Univ. Press.
- Doall, M. H., S. P. Colin, J. R. Strickler, and J. Yen. 1998. Locating a mate in 3D: the case of *Temora longicornis*. *Phil. Trans. R. Soc. Lond. B* 353:681-689.
- , J. R. Strickler, D. M. Fields, and J. Yen. 2002. Mapping the free-swimming attack volume of a planktonic copepod, *Euchaeta rimana*. *Mar. Biol.* 140:871-879.
- Domeier, M. L. and P. L. Colin. 1997. Tropical reef fish spawning aggregations: Defined and reviewed. *Bull. Mar. Sci.* 60:698-726.
- Fields, D. M., and J. Yen. 1993. Outer limits and inner structure: The 3-dimensional flow field of *Pleuromamma xiphas* (Calanoida: Metrinidae). *Bull. Mar. Sci.* 53:84-95.
- Fish, F. E., J. F. Fegely, and C. J. Xanthopoulos. 1991. Burst-and-coast swimming in schooling fish (*Notemigonus crysoleucas*) with implications for energy economy. *Comp. Biochem. Physiol.* 100A:633-637.
- Gazi, V., and K. M. Passino 2004. A class of attractions/repulsion functions for stable swarm aggregations. *Int. J. Control* 77:1567-1579.
- Greene, C. H., and P. H. Wiebe. 1997. Acoustic visualization of three-dimensional animal aggregations, p. 61-67. *In* J. K. Parrish and W. M. Hamner [eds.], *Animal groups in three dimensions*. Cambridge Univ. Press.
- Grimm, V., and U. Berger. 2003. Seeing the forest for the trees, and vice versa: pattern-oriented ecological modeling, p. 411-428. *In* L. Seuront and P. Strutton [eds.], *Handbook of scaling methods in aquatic ecology*. CRC Press.
- Hamner, W. M., and J. K. Parrish. 1997. Is the sum of the parts equal to the whole: the conflict between individuality and group membership, p. 165-173. *In* J. K. Parrish and W. M. Hamner [eds.], *Animal groups in three dimensions*. Cambridge Univ. Press.
- Huth, A., and C. Wissel. 1990. The movement of fish schools: a simulation, p. 577-590. *In* W. Alt and G. Hoffman [eds.], *Biological motion*. Springer-Verlag.
- Ikawa, T., and H. Okabe. 1997. Three-dimensional measurements of swarming mosquitoes: A probabilistic model, measuring system, and examples results, p. 90-104. *In* J. K. Parrish and W. M. Hamner [eds.], *Animal groups in three dimensions*. Cambridge Univ. Press.
- Jaffe, J. S. 1997. Methods for three-dimensional sensing of animals, p. 17-35. *In* J. K. Parrish and W. M. Hamner [eds.], *Animal groups in three dimensions*. Cambridge Univ. Press.
- Kennedy, J. S., and L. Crawley. 1967. Spaced-out gregariousness in sycamore aphid, *Drepanosiphum platanoides* (Schrank). *J. Anim. Ecol.* 36:147-170.
- Lin, C. -C., T. -W. Chen, and C. -C. Lai. 2003. The economics of honeybee swarming. *Reg. Sci. Urban Econ.* 33:581-594.
- Lloyd, M. 1967. Mean crowding. *J. Anim. Ecol.* 36:1-30.
- Lutscher, F. 2002. Modeling alignment and movement of animals and cells. *J. Math. Biol.* 45:234-260.
- McGehee, D., and J. S. Jaffe. 1996. Three-dimensional swimming behavior of individual zooplankters: observations using the acoustical imaging system FishTV. *J. Mar. Sci.* 53:363-369.
- Mogilner, A., L. Edelstein-Keshet, L. Bent, and A. Spiros. 2003. Mutual interactions, potentials, and individual distance in a social aggregation. *J. Math. Biol.* 47:353-389.
- Okubo, A., and S. A. Levin. 2001. *Diffusion and ecological problems: modern perspectives*, 2nd ed., Springer-Verlag.
- Osborn, J. 1997. Analytical and digital photogrammetry, p. 36-60. *In* J. K. Parrish and W. M. Hamner [eds.], *Animal groups in three dimensions*. Cambridge Univ. Press.
- Parrish, J. K., and W. M. Hamner [eds.]. 1997. *Animal groups in three dimensions*. Cambridge Univ. Press.
- Parrish, J. K., and P. Turchin. 1997. Individual decisions, traffic rules, and emergent patterns in schooling fish, p. 126-142. *In* J. K. Parrish and W. M. Hamner [eds.], *Animal groups in three dimensions*. Cambridge Univ. Press.
- Pathria, R. K. 1986. *Statistical mechanics*. Pergamon Press.
- Pfister, B. 1990. A one-dimensional model for the swarming behavior of *Myxobacteria*, p. 556-563. *In* W. Alt and G. Hoffman [eds.], *Biological motion*. Springer-Verlag.
- Pielou, E. C. 1977. *Mathematical ecology*, 2nd ed. Wiley.
- Ritz, D. A. 1997. Costs and benefits as a function of group size: Experiments on a swarming mysid *Paramesopodopsis rufa* Fenton, p. 194-206. *In* J. K. Parrish and W. M. Hamner [eds.], *Animal groups in three dimensions*. Cambridge Univ. Press.
- Seuront, L., F. G. Schmitt, M. C. Brewer, J. R. Strickler, and S. Souissi. 2004. From random walk to multifractal random walk in zooplankton swimming behavior. *Zool. Stud.* 43:498-510.
- Skyllingstad, E. D. 2003. The effects of Langmuir circulation on buoyant particles, p. 445-452. *In* L. Seuront and P. Strutton [eds.], *Handbook of scaling methods in aquatic ecology*. CRC Press.
- Stevens, A. Simulations of the gliding behavior and aggregation of *Myxobacteria*, p. 548-555. *In* W. Alt and G. Hoffman [eds.], *Biological motion*. Springer-Verlag.

- Turchin, P. 1996. Fractal analyses of animal movement: a critique. *Ecology* 77: 2086-2090.
- Van Duren, L. A., and J. J. Videler. 2003. Escape from viscosity: the kinematics and hydrodynamics of copepod foraging and escape swimming. *J. Exp. Biol.* 206:269-279.
- Waters, W. E. 1959. A quantitative measure of aggregation in insects. *J. Econ. Ent.* 52:1180-1184.
- Wiens, J. A., T. O. Crist, K. A. With, and B. T. Milne. 1995. Fractal patterns of insect movement in microlandscape mosaics. *Ecology* 76:663-666.
- Yamazaki, A. K., and D. Kamykowski. 2003. Modeling planktonic behavior as a complex adaptive system, p. 543-557. *In* L. Seuront and P. Strutton [eds.], *Handbook of scaling methods in aquatic ecology*. CRC Press.
- Yen, J., and E. A. Bundo. 1997. Aggregate behavior in zooplankton: Phototactic swarming in four developmental stages of *Coullana canadensis* (Copepoda, Harpacticoida), p. 143-162. *In* J. K. Parrish and W. M. Hamner [eds.], *Animal groups in three dimensions*. Cambridge Univ. Press.
- Younge, K., C. Christenson, A. Bohara, J. Crnkovic, and P. Saulnier. 2004. A model system for examining the radial distribution function. *Am. J. Phys.* 72:1247-1250.

Submitted 10 February 2006

Revised 1 September 2006

Accepted 2 September 2006