Statistical tests for analysing directed movement of self-organising animal groups

A. Merrifield *, M. R. Myerscough & N. Weber

School of Mathematics and Statistics, F07, University of Sydney, NSW 2006, Australia

Abstract

We discuss some theory concerning directional data and introduce a suite of statistical tools that researchers interested in the directional movement of animal groups can use to analyse results from their models. We illustrate these tools by analysing the results of a model of leader-less groups moving under the duress of certain indistinguishable individuals, that arises in the context of honeybee (Apis mellifera) swarming behaviour.

Key words: Spherical hypothesis tests, directional data, swarm guidance, self-organisation, honey bee, Apis mellifera.

1 Introduction

Spherical data takes the form of a set of directions in space or positions on the surface of a sphere and is used in areas of science including earth sciences, biological sciences and astrophysics. The contexts may vary, but the statistical methodology is common to most situations. Closely related is the area of circular statistics, which concerns data distributed on a circle.

In this article we introduce appropriate tests for spherical data and use them to analyse the behaviour of a migrating swarm of honeybees where the data has been generated by a computer simulation. Real data on honeybee movement is starting to become available [1] and we forsee that the methods discussed in this paper could be used to analyse real information in a similar way to the analysis of simulated data presented here.

* Corresponding author.
Email address: alistair@maths.usyd.edu.au (A. Merrifield).
A swarm of honeybees are guided to a pre-determined nest site, by a small number of bees who have prior knowledge of the location of the site \[2,3\]. The informed bees are indistinguishable from the ignorant members of the swarm and guide the uninformed swarm to their new home by flying continuously through the swarm with flight paths aligned in the direction of the new home, using visual cues \[4\].

\[5\] and \[6\] have produced models to explain the phenomena of directed movement of a leader-less group. \[7\] formulated a three-dimensional self-organising model of group formation based on simple rules of avoidance, attraction and alignment to examine the spatial dynamics of animal groups. A self-organising process is one where the collective actions of the group result from numerous interactions amongst the individuals of the group \[8\]. \[5\] extend the model of \[7\] to investigate the issue of a select few group members, privy to information regarding the location of a goal, guiding the uninformed group members. The concept of coordinated movement of groups, induced by a select few educated group members, also has applications in robotics \[9,10\] and similar fields.

Statistical methods are needed to assess whether the knowledgeable individuals (the scout bees) have been able to influence the ignorant (workers) and guide them from an arbitrary travel path to one aligned with the goal (the new nest-site). Are the uninformed individuals moving in a direction consistent with the goal and are these individuals moving in a coherent group towards this goal?

In order to answer these questions, \[5\] and \[6\] have applied measures of accuracy to the results of their simulations. Neither sets of authors consider the dispersion (cohesiveness) of the group of uninformed individuals in their analysis. \[6\] fix their goal some distance along the \(x\)-axis, away from the initial positions of the group. The authors assess the accuracy of the group by extending a plane perpendicular to the goal position and record the point where the group first crosses this plane. Accuracy is defined as the distance between this point and the goal. They also compare the length of the trajectory of the centre of the group with a direct path between the origin and goal position, to obtain a ratio indicating how linear the trajectory is. \[5\] use the normalised angular deviation of the group direction around the direction of the goal, as a measure of accuracy. The authors calculate the group direction as a vector extending toward the position of the centre of the group at the end of the simulation, from the position fifty timesteps beforehand. The three dimensional system is relegated to two dimensions, by calculating the circular angular deviation in the plane common to the group direction and the preferred direction.

We wish to avoid possible situations where the group’s direction is aligned with that of the goal by chance, due to a large dispersal (such as is found in swarm
conditions, [7]). Hence, we use measures of dispersion to indicate how concentrated the vectors are around the group direction; the smaller the measure of dispersion, the more likely the individual direction vectors are to be clustered around the group average direction (indicative of a coherent group formation). It is straightforward to analyse directional data in the three dimensional context in which it belongs, rather than lose information by reducing the problem to lower dimensions.

Studies of directional data date back to the beginnings of statistical theory [11,12]. The analysis of spherical statistics essentially started with [13], where a distribution for angular errors on a sphere and methods for inferences of mean directions and dispersions were developed. From the mid-1950’s, interest in directional data increased [14-18]. Systematic accounts of theories and methodologies for circular and spherical data have been published in [19-22]. Recent times have seen an interest in adapting conventional linear statistical techniques to find analogues in the circular and spherical realms [23-26]. Directional data are found in seismology (in the determination of earthquake fault planes, [27]) and structural geology [28]. Meteorology associates vectorial data with wind directions [29], oceanography with ocean current directions [30] and astrophysics with the arrival directions of cosmic ray showers [31]. Biological sciences utilises directional methodology in the context of animal movements [32].

Because the results of the self-organised models [5,6] are a set of directions of travel for each individual, spherical probability theory best reflects the physical situation being modelled. The use of conventional linear statistical techniques to analyse directional data leads to paradoxes and is inappropriate. Spheres have different algebraic structures to lines; the periodicity of the sphere needs to be taken into account. A simple illustration of this point occurs if we consider the circular case and take the average of two directions placed at $10^\circ$ and $350^\circ$. The univariate arithmetic average is the mid-point of the two, namely $180^\circ$, but this does not take into account that the circle is periodic and that the geometrically correct mean is $0^\circ$. To properly assess data, we need a theoretical approach that takes periodicity into account.

Spherical probability theory provides insight to those in univariate statistics. Spherical hypothesis tests can examine what (if any) qualitative features can be attributed to the population by the sample of directions. In the case of vectorial data, spherical theory can suggest if there is a single preferred direction amongst the sample of vectors (or amongst the sample of points distributed on the sphere). The hypothesis tests discussed in this article are intuitively familiar to someone schooled in univariate statistics and they have counterparts amongst linear hypothesis tests. The spherical tests are simple to implement and require minimal time to compute.
We commence by discussing some relevant spherical theory whose origins lie in [21]. We introduce a suite of statistical tools that researchers interested in directional animal movement can use to analyse results from their models. We illustrate this theory by analysing the results of a model of leader-less groups moving under the influence of indistinguishable (but informed) individuals, arising in the context of modelling honeybee swarming behaviour.

2 Background theory

We commence by introducing the multivariate central limit theorem and key notation. Let the random vector $\mathbf{x} \in \mathbb{R}^q$, $\mathbb{E}(\mathbf{x}) = \mathbf{\mu}$ and $\mathbb{E}(\mathbf{x} - \mathbf{\mu})(\mathbf{x} - \mathbf{\mu})^T = \Sigma$, a $q \times q$ matrix. Let $\mathbf{x}_i$ ($i = 1, \ldots, n$) be a random sample of the vector $\mathbf{x}$. Define the resultant $\mathbf{S}_n = \sum_{i=1}^n \mathbf{x}_i$. Then the multivariate central limit theorem states that as $n \to \infty$,

$$\sqrt{n} \left( \frac{\mathbf{S}_n}{n} - \mathbf{\mu} \right) \sim \mathcal{N}_q(0, \Sigma)$$

where $\mathcal{N}_q(\mathbf{\mu}, \Sigma)$ denotes a Gaussian distribution in $\mathbb{R}^q$ with mean vector $\mathbf{\mu}$ and covariance matrix $\Sigma$. See, for example, [34] for proof.

Define $\Omega_q = \{ \mathbf{x} : \mathbf{x} \in \mathbb{R}^q, \ |\mathbf{x}| = 1 \}$ and suppose $\mathbf{x} \in \Omega_q$. The spherical mean direction of the $n$ observations $\mathbf{x}_i$ ($\mathbf{x}_i \in \Omega_q$) is estimated by $\hat{\mathbf{\mu}} = \mathbf{S}_n/|\mathbf{S}_n|$. If the observations are clustered around a particular direction, the value that $|\mathbf{S}_n|$ assumes will be close to $n$. Conversely, if the observations are dispersed, $|\mathbf{S}_n|$ will be small. Consequently, $|\mathbf{S}_n|$ can be used as a measure of the concentration (dispersion or cohesiveness) of the sample about the mean direction $\mathbf{\mu}$ [22].

Let the set $\{ \mathbf{e}_1, \ldots, \mathbf{e}_q \}$ be an orthonormal basis for $\mathbb{R}^q$. Set $\mathbf{e}_q = \mathbf{\mu}$ and define $\mathbf{\xi}_{q-1}$ as a unit vector in the space spanned by $\{ \mathbf{e}_1, \ldots, \mathbf{e}_{q-1} \}$. Then $\mathbf{x}$ can be decomposed as

$$\mathbf{x} = t\mathbf{\mu} + (1 - t^2)^{\frac{1}{2}} \mathbf{\xi}_{q-1}$$

(1)

where $t$ is a scalar random variable ($-1 \leq t \leq 1$). Due to rotational symmetry of $\mathbf{x}$ in (1), $\mathbb{E}(\mathbf{x}) = t\mathbf{\mu} \mathbb{E}(t)$. Thus $\mathbb{E}(t)$ is a measure of the concentration of the distribution on $\Omega_q$ about $\mathbf{\mu}$ and can be estimated by $|\mathbf{S}_n|/n$. 

4
2.1 Spherical correlation

Let \( x \) and \( y \) be two random unit vectors in \( \Omega_\mathbf{q} \). Let \((x_i, y_i) (i = 1, \ldots, n)\) be \( n \) random samples of \( x \) and \( y \). We want to estimate some measure of the association between \( x \) and \( y \) [34]. Following the approach of [35], we define the spherical correlation coefficient as a measure of this association:

\[
\rho_q = \frac{|E(XY^T)|}{\sqrt{|E(XX^T)||E(YY^T)|}}
\]  
(2)

where \( X \) is a matrix whose columns consist of the vectors \( x_i \). We can estimate \( \rho_q \) with [22]:

\[
\hat{\rho}_q = \frac{|\sum_{i=1}^{n} x_i y_i^T|}{\sqrt{|\sum_{i=1}^{n} x_i x_i^T||\sum_{i=1}^{n} y_i y_i^T|}}
\]  
(3)

2.2 Hypothesis test of a single direction

From the multivariate central limit theorem, we can develop a test that the spherical mean of the group of vectors assumes the direction \( \mu_0 \). Consider the decomposition of the resultant vector onto the hypothesised direction:

\[
S_n = \text{Proj}_{\mu_0} S_n + S_n^\perp \mu_0
\]  
(4)

where the orthogonal component \( S_n^\perp \mu_0 = (I_q - \mu_0 \mu_0^T) S_n \). If the null hypothesis is true, we can argue this orthogonal component will be small (as \( S_n \) points in the direction of the mean, \( \mu_0 \)).

**Lemma 1**

\[
\left(1 - E(t^2)\right)^{-1} \frac{1}{n} \left(|S_n^\perp \mu_0|^2\right) \sim \chi^2_{\mathbf{q}-1}
\]  
(5)

for \( n \to \infty \), where \( \chi^2_{\mathbf{q}-1} \) denotes a Chi-squared distribution with \( \mathbf{q} - 1 \) degrees of freedom [21].

We can estimate \( E(t^2) \) by \( \sum_{i=1}^{n} \left(\mu^T x_i\right)^2 / n \).
Let $S_{n_j}$ be the resultant vector from the $j$-th sample ($j = 1, \ldots, m$). Let us consider the statistic:

$$ T = \sum_{j=1}^{m} w_j |S_{n_j}| - |\sum_{j=1}^{m} w_j S_{n_j}| $$. (6)

where the positive weights $w_j$ take into account the variabilities of the $S_{n_j}$. The motivation for the use of the statistic $T$ is that $T$ will tend to be small when the individual resultant vectors align. The required form of the weights $w_j$ is $w_j = \left( ((q-1) E(t_j)) / \left(1 - E(t_j^2) \right) \right)$. Asymptotic approximations can be made for the two summations in $T$. Once this is done, we define $z_{n_j}$ as:

$$ z_{n_j} = \left( \frac{w_j}{n_j E(t_j)} \right)^{\frac{1}{2}} S_{n_j}^{\perp} $$. (7)

which has a multivariate Gaussian distribution. We can write

$$ 2T \approx \sum_{j=1}^{m} |z_{n_j}|^2 - |\sum_{j=1}^{m} \lambda_j^{\frac{1}{2}} z_{n_j}|^2 $$. (8)

where $\lambda_j = \left( n_j w_j E(t_j) \right) / \left( \sum_{j=1}^{m} n_j w_j E(t_j) \right)$ such that $\lambda_j > 0$ and $\sum_{j=1}^{m} \lambda_j = 1$. These $\lambda_j$ components effectively standardise such that $2T$ becomes the sum of squared independent standard multivariate Gaussian random variables. Consequently,

**Lemma 2** If the null hypothesis of equality of mean directions across samples is true, then for large $n_j$ ($j = 1, \ldots, m$), $2T$ is distributed approximately like $\chi^2_{(m-1)(q-1)}$ [21].

### 2.4 Hypothesis test of equality of concentrations of two groups

It can be shown from the multivariate central limit theorem that as $n \to \infty$, $\sqrt{n} \left( |S_n|/n - E(t) \right) \overset{D}{\to} N(0, \text{Var}(t))$. From the additivity properties of the Normal distribution, we have the following result.

**Lemma 3** Suppose we have two independent samples (of sizes $n_1$ and $n_2$). Assuming that the concentrations from the two distributions are the same (i.e. $E(t_1) = E(t_2)$), then
\[
\frac{|S_{n_1}|}{n_1} - \frac{|S_{n_2}|}{n_2} \xrightarrow{D} N \left( 0, \frac{\text{Var}(t_1)}{n_1} + \frac{\text{Var}(t_2)}{n_2} \right)
\]

for large \(n_1\) and \(n_2\).

Alternatively,

\[
\left( \frac{\text{Var}(t_1)}{n_1} + \frac{\text{Var}(t_2)}{n_2} \right)^{\frac{1}{2}} \left( \frac{|S_{n_1}|}{n_1} - \frac{|S_{n_2}|}{n_2} \right) \approx N(0, 1). \tag{9}
\]

We can use the left hand side of (9) as a test statistic to assess the null hypothesis of equality of concentrations. We will need to estimate the variances of the concentrations. We can do this by:

\[
\widehat{\text{Var}}(t) = \hat{E}(t^2) - \hat{E}(t)^2
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} (\hat{\mu}^T x_i)^2 - \left( \frac{|S_n|}{n} \right)^2.
\]

3 Results and discussion

To illustrate the appropriateness of spherical theory for analysing vectorial samples, we take a random sample from a Fisher distribution (refer to [19] or [22]) which is centred on the direction \((0.8, .42, .42)\) and test the hypothesis that the true direction could be the \(x\)-axis \((1, 0, 0)\), using Lemma 1. The concentration parameter \((\kappa > 0)\) controls the alignment of the sample vectors relative to one another. As the value of the concentration parameter increases, the distribution degenerates to a point distribution centred on the mean (in this case, the direction \((.8, .42, .42)\)). The value of the concentration parameter is varied to compare the effect on the hypothesis test presented here and normalised angular deviation measures used by [5]. As the concentration parameter increases (and the sample becomes highly polarised), the hypothesis test correctly rejects the possibility of alignment with the \(x\)-axis (5% significance level). Normalised angular deviation fails to account for this change in dispersion and remains relatively invariant (Figure 1).

We use honeybee guidance as an example for illustrating the use of the previously discussed spherical hypothesis tests. Our basic model follows the same principles as the model of [7], which is based on simple behavioural rules. The simulations have 100 individuals present, where 5 individuals have prior knowledge of the location of the goal. This is consistent with experimental observations of honeybees, which indicate that 5% of the swarming bees have
prior knowledge of the location [36]. We firstly allow our uninformed individuals to form an organised group, travelling in an arbitrary direction. After some time in the simulation has elapsed, we introduce a small number of informed individuals to the group. The informed individuals move with paths of travel aligned with the x-axis and were dispersed in certain formations amongst the group. The informed individuals’ paths are perturbed by a random vector drawn from a Fisher distribution.

We illustrate the hypothesis tests from Lemmas 1, 2 and 3 in two ways. We use the three Lemmas to obtain evidence that the outcome of the model of leader-less directed motion is consistent with the hypothesis that the informed individuals can guide a mass of uninformed individuals by flying through the group, pointing the direction of the goal [4]. Secondly, we illustrate the usefulness of the techniques by comparing simulations with different configurations of knowledgeable individuals, within the cluster of ignorant individuals.

3.1 Analysis of the leader-less directed motion model

We commence testing the leader-less model by applying the test from Lemma 2 to assess whether there has been a significant change of direction to the ignorant group once the informed members have been introduced. We use the sample of directions of ignorant members prior to introduction of the knowledgeable members. We compare this with the sample of directions of the ignorant members at the conclusion of the simulation, when they have been under the influence of knowledgeable individuals for a considerable period of time. The knowledgeable individuals move through the centre of the group, one after another, with speeds larger than those of the ignorant members. The estimate of spherical correlation ($\hat{\rho}_3$) is 0.0997, meaning the two samples can effectively be regarded as independent for the hypothesis test. The approximating distribution of the test statistic $2T$ (from Lemma 2) follows a $\chi^2_2$ distribution ($q = 3$, $m = 2$). We calculate the test statistic to be 374.69 and the $p$-value to be approximately 0. We conclude that there has been a significant change in average group direction once the informed individuals have been introduced to the group. This does not allow us to assume that the direction of the ignorant group coincides with that of the goal.

However, we can test using Lemma 1 to see if the average direction of the group of uninformed individuals coincides with the direction of the goal (along the x-axis). We use the directions of the group of uninformed individuals at the conclusion of the simulation and calculate the test statistic to be 3.49 (corresponding to a $p$-value of 0.17). A large $p$-value indicates that the sample of directions is consistent with hypothesised direction. This is consistent with the null hypothesis that the average direction of the ignorant components
coincides with the direction of the goal. We compare the directions of the group of ignorant individuals at the conclusion of the simulation with the flight path of the scouts, using Lemma 2. This yields the test statistic to be 3.23 and a corresponding \( p \)-value of 0.20. Therefore, this indicates that there is no significant difference in average direction between the ignorant and knowledgeable groups at the end of the simulation.

Finally, we assess that we have a coherent group by testing the concentrations (dispersal) of the group initially and at the end of the simulation. To decide whether the group of uninformed individuals have formed an organised aligned group, we use Lemma 3. The two samples consist of the initial (random) and the final positions of the ignorant individuals in the simulation. The correlation between these two groups is estimated to be -0.0003. The null hypothesis is equality of concentrations. Using (9), we calculate the test statistic for the initial and final directions to be 10.74. The test statistic is significant (\( p \)-value is approximately 0). The concentration of ignorant individuals in the final positions is significantly larger than those of the same individuals in the initial positions. In fact, the concentration of the ignorant group is initially 0.06 and changes after the conclusion of the simulation to be 0.83. This reflects that the group has had a transition during the simulation from a group composed of individuals with random orientations (relative to one another), to one whose orientations are highly polarised.

To summarise, the statistical tests indicate that the uninformed group have had a significant change in direction once the knowledgeable individuals have been introduced to the system. The average direction of the ignorant individuals plausibly follows the direction of the goal and also coincides with the direction of the travel path of informed individuals. The group of ignorant individuals has changed significantly from a dispersed configuration to one of a more coherent nature. This is indicative of a cohesive group moving along a specific direction. Furthermore, this group has been influenced with the introduction of knowledgeable individuals and has shifted their travel path from one that is arbitrary, to one aligned with the flight direction of the informed individuals and the goal. The results of these tests support leader-less guidance using spatial cues.

3.2 Analysis of effective configurations of knowledgeable individuals

What configuration of knowledgeable individuals is most effective in guiding the mass of ignorant individuals to the goal? We run the simulation with five possible configurations of the informed entities. The first configuration constrains the informed to travel through the centre of the group (configuration 1). The second configuration allows the informed to be distributed around the
perimeter of the ignorant group (configuration 2). The third is a compromise between the first two, where the informed units are distributed within the group (same as that used in the previous simulation in Section 3.1, configuration 3). [4] suggest that the scout bees fly in the top portion of the swarm, sandwiching the swarm between themselves and the ground. We allow the informed individuals to fly in parallel formation on one side of the ignorant workers (configuration 4). For the final configuration, we include a reflective boundary in configuration 4 to simulate the effect of the ground (configuration 5). These configurations are shown in Figure 2.

We examine the uninformed groups’ final directions of travel using Lemma 2, to see whether there is a difference between the five configurations. The approximating distribution of the test statistic $2T$ follows a $\chi^2_q$ distribution ($q = 3$, $m = 5$). The result of the hypothesis tests indicates there is some significant difference between the resulting directions of the five simulations (p-value is small, test statistic is 5963.26). At least one of the five groups of ignorant individuals has a different average direction compared to the others.

We test each of the five configurations using Lemma 1, to assess if the resulting directions can be regarded as aligning with the $x$-axis (the results are summarised in Table 1). Results from two configurations stand out as having group directions that may coincide with the direction of the goal. Those configurations are configuration 1 (p-value is 0.38, test statistic is 1.95) and configuration 3 (see previously discussed work on assessing the validity of the guidance hypothesis, p-value is 0.17, test statistic is 3.49). From the results of the hypothesis tests, configuration 2 does not appear to be an effective one for facilitating guidance. We can test in a pairwise sense that the average directions of the results of configurations 1 and 3 are coincidental, the hypothesis of equal directions is accepted (p-value = 0.06, test statistic is 5.80).

We conclude that either of these two configurations are effective for guiding the ignorant mass. The configurations with knowledgeable individuals flying on one side of the group (configuration 4 and 5) is not effective for inducing the ignorant group to follow. In configuration 4, the ignorant individuals are able to escape the influence of the informed members. During the simulation of configuration 5, the ignorant mass consistently elude the knowledgeable members by eventually flying around the informed fixed flight paths and follow their own course. We observed during the simulation that the informed individuals are able to influence the group in the short- and medium-term. Applications of Lemma 1 to the directions of the group during this time period bear this out. For example, for the directions of the group members at the half-way point of the simulation, the test statistic from Lemma 1 has the value of 5.74 and a p-value of 0.06. At this point, the group direction is conceivably aligned with the $x$-axis. As time progresses, the group moves beyond the reach of the knowledgeable members. This does not suggest that the proposition of
[4] is incorrect in the long-run; to properly investigate this idea the simulation needs to allow the knowledgeable members the flexibility to adjust their flight paths with the movement of the group.

4 Conclusion

We have discussed a number of results relating to directional data, stemming from the multivariate central limit theorem, for data distributed on a hypersphere. These techniques can be used to analyse the outcomes of models of directed movement of groups comprising of identical individuals. We argue that the data from these models are vectors and should be correctly treated as directional data. We use spherical correlation [22] as a measure of independence of samples. From [21], we introduce tests for a single prescribed direction of data, equality of directions between samples and equality of concentration between samples.

We have illustrated these techniques with the analysis of a self-organising model used to explore the idea of guidance of groups by homogeneous individuals endowed with prior knowledge of the directions of some goal. Using these methods, we have shown that the informed individuals have influenced the uninformed group to turn away from an arbitrary direction, to one aligned with the goal as pointed by the knowledgeable members. We have also used the hypothesis tests to compare several different patterns that the informed members can adopt to distribute themselves amongst the group. The tests allow us to differentiate between effective and ineffective methods to organise the group.

These directional tests prove tractable and are readily accessible to researchers wanting to test assumptions about the samples of data that have an element of direction associated, in particular, for the analysis of the output of models related to animal group movement. With more sophisticated methods becoming available for tracking animal movements, these methods can be used for studying real animal data. [1] use radar and attach harmonic transponders to honeybees and track individual flight paths (in space and time) of honeybees moving towards a food source. The honeybees flight path coordinates were recorded in three second intervals on a desktop computer. With minor modifications to their experiment to study swarm guidance, we envisage that, given the three dimensional coordinates of flight paths, it would be straightforward to analyse the trajectories using the methods discussed in this article.

Given that data of aggregative movement (real or simulated) consists of vectorial observations in time, a natural extension may be to consider a time series analysis (taking into account the geometry of the sphere). A simple approach
would be to smooth the data using a moving average process, to obtain a
time sequence of mean directions (using some predefined time window, possi-
bly enhanced by appropriate weights). The sequence of means should orientate
towards the direction indicated by the knowledgeable individuals, as time pro-
gresses. Alternatively, there have been developments in spline techniques for
smoothing and interpolating directional data [37] that may be applicable.

Acknowledgements

This work was made possible by a University Postgraduate Award to AM from
the University of Sydney. The authors are grateful to Madeleine Beekman and
David Easdown for helpful discussions.

5 Figures and Tables
Fig. 1. Normalised angular deviation (as used by [5], denoted NAD) and p-values of a random Fisher sample, as a function of the log concentration parameter ($\log_{10} \kappa$). The polarisation of the sample increases with increasing $\kappa$. The dashed line indicates the significance level of 0.05.
Fig. 2. Configurations for informed individuals path within the self-organising group. The grey circle represents the cross-section of the approximate mass of ignorant group members and the dots indicate the positions of the knowledgeable members. Figure (a) is the concentrated configuration (configuration 1), (b) the configuration where individuals are dispersed along the perimeter of the group (configuration 2), (c) the configuration where the informed are dispersed within the group (configuration 3), (d) the informed are dispersed on one side of the group (configuration 4) and (e) is the same as configuration D with a reflecting boundary (configuration 5).
<table>
<thead>
<tr>
<th>Null hypothesis</th>
<th>Test statistic</th>
<th>p-value</th>
<th>$\hat{\mu}$</th>
<th>$\hat{E}(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 group means all equal</td>
<td>5963.264</td>
<td>small</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_1 = (1,0,0)$</td>
<td>1.9461</td>
<td>0.3779</td>
<td>(.999,.022,.029)</td>
<td>0.902</td>
</tr>
<tr>
<td>$\mu_2 = (1,0,0)$</td>
<td>389.6532</td>
<td>small</td>
<td>(.924,.295,.242)</td>
<td>0.983</td>
</tr>
<tr>
<td>$\mu_3 = (1,0,0)$</td>
<td>3.4913</td>
<td>0.1745</td>
<td>(.999,.028,.023)</td>
<td>0.939</td>
</tr>
<tr>
<td>$\mu_4 = (1,0,0)$</td>
<td>497.9486</td>
<td>small</td>
<td>(-.571,.783,.244)</td>
<td>0.863</td>
</tr>
<tr>
<td>$\mu_5 = (1,0,0)$ (at end of simulation)</td>
<td>887.3989</td>
<td>small</td>
<td>(.030,.127,.991)</td>
<td>0.885</td>
</tr>
<tr>
<td>$\mu_5 = (1,0,0)$ (at half-way point)</td>
<td>5.71849</td>
<td>0.0573</td>
<td>(.986,.151,.067)</td>
<td>0.954</td>
</tr>
<tr>
<td>$\mu_1 = \mu_3$</td>
<td>5.7987</td>
<td>0.0551</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1
Summary statistics for hypothesis tests of different configurations of informed individuals. The test statistic column refers to the value of the test statistic from the appropriate hypothesis test. The p-value column gives the corresponding p-value. The $\hat{\mu}$ column lists the average direction of the sample and $\hat{E}(t)$ quotes the estimated concentration of the sample.
References


[12] Lord Rayleigh, On the resultant of a large number of vibrations of the same pitch and or arbitrary phase, Philosophical Magazine. 10 (1880) 73-78.


