Chapter 2

Spherical statistics and related directional issues

We commence this chapter by discussing various techniques for analysing trajectories and directions. We introduce statistics associated with directional data and discuss relevant spherical probability theory, which will assist us in investigating simulated and real animal movements.

2.1. Introduction to directional data and related issues

Spherical data takes the form of a set of directions in space (or equivalently, 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 are circular statistics, concerning data distributed on a circle.

Studies of directional data date back to the beginnings of statistical theory; Gauss developed the theory of errors to analyse directional measurements in astronomy. He was able to use linearity as an approximation. Early developments in the field concerned uniformly distributed random vectors. Daniel Bernoulli effectively considered points uniformly distributed on a sphere, when examining the orbital planes of the then known planets (Bernoulli 1734, Mardia 1972). Rayleigh considered the distribution of the resultant length of normal vectors to a plane and later, a uniform random walk on a sphere with approximations for large samples (Lord Rayleigh 1880, Lord Rayleigh 1919). Non-uniform distributions came to the attention of mathematicians and statisticians in the 20th century. Brownian motions on a circle and a sphere led to the appropriate wrapped Normal distributions. Distributions bearing the names of their authors, including von Mises, Langevin and Arnold, are well documented, see eg. Mardia (1972).

The analysis of spherical statistics essentially started with R.A. Fisher (1953), who developed a distribution for angular errors on a sphere and methods for inferences of mean directions and dispersions. From the mid-1950's, Watson further developed methodologies for spherical (and circular) statistics (Watson & Williams 1956, Watson 1960). Stephens was also responsible for furthering studies on statistical tests in the circular (Stephens 1962*a*) and the spherical cases (Stephens 1962*b*). Stephens (1967) examined tests for dispersion on a sphere resulting from the distribution suggested in Fisher (1953). N.I. Fisher (1985) investigated various properties of the spherical median and discussed equivalents for the sign test.

A systematic account of theory and methodology for circular and spherical data has been published in Mardia (1972). Batschelet (1981) summarised circular methodology in a text aimed primarily at biologists. Watson (1983*b*) published a set of lecture notes discussing a summary of various aspects of spherical statistics. More recently, the book by Fisher, Lewis & Embleton (1987) is devoted purely to analysis of spherical data. Recent times have seen an interest in adapting conventional linear statistical techniques to find analogues in the circular and spherical realms (Brunner 1994, Chang 1993, Fisher 1990, Mardia 2002).

Spherical data traditionally arises in the context of the earth sciences. Palaeomagnetism lends itself to the analysis of directional data using the orientations of remnant magnetism preserved in rocks (Fisher 1990). The seminal paper of R.A. Fisher (Fisher 1953) applied spherical methodologies to analyse paleomagnetic data. Directional data are found in seismology in the determination of earthquake fault planes (Storetvedt & Scheidegger 1992) and structural geology (Fisher 1989). Meteorology associates vectorial data with wind directions (Fisher 1987), oceanography with ocean current directions (Bowers, Morton & Mould 2000) and astrophysics with the arrival directions of cosmic ray showers (Briggs 1993). Preisler & Akers (1995) were concerned with movement of female bark beetles and used circular theory for the purposes of modelling. Their paper develops an autoregressive model to study the distribution of the angular response of bark beetles attracted to a pheromone emitting source at a particular time, given the past history of the beetle's movements.

In this chapter, we introduce appropriate statistics for analysing directional output from models in subsequent chapters, and tests of inference for these data.

The data in this thesis has been generated artificially. However, real data on animal movement is starting to become available. Riley, Greggers, Smith, Reynolds & Menzel (2005) discuss data obtained by attaching transmitters to the backs of honeybees and tracking the bees whilst they were flying. Data obtained from these trajectories were downloaded and stored on computer. Couzin & Franks (2003) constructed individual trajectories of army ants (*Eciton burchelli*) with the use of a digital video camera. We forsee that the methods discussed in this chapter could be used to analyse real information in a similar way to the analysis of simulated data presented in this thesis.

2.2. Descriptive statistics to aid analysing directional data

We need to define some descriptive statistics associated with directions, to facilitate our analysis. These statistics will be used to assess global properties of a sample of directions.

Suppose we have a group composed of N individuals. At a particular time t, each individual has an associated position $\mathbf{r}_i(t)$ (a column vector of Cartesian coordinates) and a unit direction vector $\mathbf{v}_i(t)$ (where $i = 1, \ldots, N$; and the time period is partitioned into sub-intervals such that $t = 1, \ldots, T$). This section is, in the main, concerned with statistics for an analysis of a set of directions at a particular point in time; hence for the remainder of this chapter, we shall drop time dependence where relevant and assume it is an implicit property herein. We assume all vectors are in their Cartesian coordinates, for the duration of this chapter.

One fundamental statistic that we need to define is the centre of the group. This is analogous to the centre of mass in a multiparticle physical system, where the mass of each individual component is identical. The group centre is calculated as the mean of all the individuals position vectors at time t:

$$\mathbf{r}_{group} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{r}_i \ . \tag{2.1}$$

We need a measure of the group direction. We assume that the group is heading towards a spherical mean direction $\boldsymbol{\mu}$ (where $\mathrm{E}(\mathbf{v}_i) = \boldsymbol{\mu}$) and we define the resultant vector as

$$\mathbf{S}_N = \sum_{i=1}^{N}\mathbf{v}_i. \tag{2.2}$$

The spherical mean of the N directions is estimated by:

$$\hat{\boldsymbol{\mu}} = \frac{\mathbf{S}_N}{|\mathbf{S}_N|}, \tag{2.3}$$

where $|\mathbf{S}_N|$ is the norm of the vector $\mathbf{S}_N$. The spherical mean is by no means a unique way to determine group direction. Spherical medians and other such measures exist (Fisher et al. 1987). However, the spherical mean (similar to the univariate mean) enjoys certain statistical properties, which we shall make use of in the next section.

The group polarisation (p_{group}) measures the degree of alignment amongst individuals within the group. A highly coherent (or aligned) group will have a high polarisation, while a more dispersed arrangement will be assigned a lower value. Conversely, angular momentum (m_{group}) is a measure of the degree of rotation of the group around the group centre. It is to be understood that in the context of this thesis, when we refer to momentum, we mean in the rotational sense and not the linear (unless specifically stated to the contrary). These two statistics are defined below, see eg. Couzin et al. (2002):

$$p_{group} = \frac{1}{N}\left|\sum_{i=1}^{N}\mathbf{v}_i\right| = \frac{|\mathbf{S}_N|}{N} \quad (2.4)$$

$$m_{group} = \frac{1}{N}\left|\sum_{i=1}^{N}\hat{\mathbf{r}}_{i,group}\times\mathbf{v}_i\right| \quad (2.5)$$

$$\text{where } \hat{\mathbf{r}}_{i,group} = \frac{\mathbf{r}_i - \mathbf{r}_{group}}{|\mathbf{r}_i - \mathbf{r}_{group}|}.$$

Analogues to the univariate variance in the spherical realm measure the dispersion of a set of vectors. Should the observed directions be 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, $|\mathbf{S}_N|$ is a measure of the concentration of the sample about the mean direction $\boldsymbol{\mu}$ (Fisher et al. 1987). Concentration refers to the degree of alignment of the sample of vectors, relative to one another (2.4). We can define a spherical variance, as an estimate of the dispersion of the group in the following way (Mardia 1972):

$$\hat{\sigma} = 1 - |\mathbf{S}_N|/N = 1 - p_{group}. \quad (2.6)$$

We can measure the relative size of the group with two measures, nearest neighbour distance (*NND*) and group expanse. The nearest neighbour distance varies for each individual within the group. It simply is the distance between a particular individual and its closest neighbour. See for example, Viscido, Parrish & Grünbaum (2004), Viscido, Parrish & Grünbaum (2005). The nearest neighbour distance for individual i is defined as

$$\mathrm{NND}_i = \min_{j\neq i}|\mathbf{r}_i - \mathbf{r}_j|. \quad (2.7)$$

The further away an individual is from its neighbours (and the group), the larger the nearest neighbour distance for that individual. The *NND* can indicate outlying individuals from the group.

The group expanse measures the size of the group, as an entity (Viscido et al. 2005, Huth & Wissel 1992). We calculate expanse as:

$$\text{Expanse} = \frac{1}{N}\sum_{i=1}^{N} |\mathbf{r}_i - \mathbf{r}_{group}|. \tag{2.8}$$

There are alternative measures for the expanse of the group; we could look at the maximum distance between individuals and the group centre, for example. Niwa (1998) is a study devoted entirely to the group size distribution of fish schools; the author examines the size of the group with an integrodifferential equation, based on the rates of amalgamation and fragmentation within the school. Viscido et al. (2005) use $\sum_{i=1}^{N}\sqrt{(|\mathbf{r}_i| - |\mathbf{r}_{group}|)^2}/N$, which measures the relative 'tightness' of the group. We use (2.8) as our measure of expanse, an 'average' distance between individuals and the group centre.

We want to be able to measure the curvature of the path that an individual takes over a certain time period. The net to gross displacement ratio (*NGDR*) for an individual is the ratio of the actual displacement of an individual's path (gross displacement) to the displacement of an individual assuming that the individual took a linear path between the two positions (net displacement). The *NGDR* for individual i over the interval (t_0, t_f) is:

$$\begin{aligned} \text{NGDR}_i &= \frac{|\mathbf{r}_i(t_f) - \mathbf{r}_i(t_0)|}{\int_{t_0}^{t_f} |\frac{d\mathbf{r}}{dt}| dt} \\ &= \frac{|\mathbf{r}_i(t_f) - \mathbf{r}_i(t_0)|}{\sum_{j=1}^{f} |\mathbf{r}_i(t_j) - \mathbf{r}_i(t_{j-1})|}. \end{aligned} \tag{2.9}$$

The interval (t_0, t_f) is partitioned by the points $t_0, t_1, \ldots, t_f$. The closer the value of *NGDR* to 1, the straighter the path of the individual. As the individual takes a more tortuous route, the *NGDR* value tends towards 0 (Parrish et al. 2002).

Finally, we need to be able to compare two sets of spherical data. Suppose we have two sets of directional data and we wish to calculate the degree of association between them. In a univariate situation, we could measure this association with the correlation coefficient. We want an equivalent for the directional case.

Let $\mathbf{x}$ and $\mathbf{y}$ be two random unit vectors in $\mathbb{R}^q$, where q is the dimension of the space we are interested in (usually $q = 2$ or 3). Let $\mathrm{E}(\mathbf{x}) = \boldsymbol{\mu}_x$ and $\mathrm{E}(\mathbf{y}) = \boldsymbol{\mu}_y$. We adjust the vector $\mathbf{x}$ by subtracting the mean, denote this new vector $\mathbf{x}'$. Let $(\mathbf{x}_j, \mathbf{y}_j)$ with $j = 1, \ldots, J$, be J random observations of the random vectors $\mathbf{x}$ and $\mathbf{y}$.

We define $\mathbf{x}$ and $\mathbf{y}$ to be perfectly positively associated if there exists an A such that $\mathbf{y} = A\mathbf{x}$ and $det(A) = 1$ (A is a $q \times q$ orthogonal matrix and is therefore a rotational matrix). The random vectors are perfectly negatively associated if $det(A) = -1$, meaning A is a reflection matrix (Stephens 1979).

We want to measure the degree of association between $\mathbf{x}$ and $\mathbf{y}$. We can do this by measuring the difference between the degree of positive and negative association (Fisher & Lee 1983). We can use the measure

$$\mathrm{E}\left(\min_{det(A)=-1}|\mathbf{y}' - A\mathbf{x}'|^2 - \min_{det(A)=1}|\mathbf{y}' - A\mathbf{x}'|^2\right), \qquad (2.10)$$

where $|\mathbf{y}' - A\mathbf{x}'|$ is the magnitude of the difference between $\mathbf{y}'$ and $A\mathbf{x}'$ (Fisher & Lee 1986). The choices of the orthogonal matrix A for the cases $det(A) = -1$ and $det(A) = 1$ are non-unique, we choose the orthogonal matrix A that minimises the distance between $\mathbf{y}'$ and $A\mathbf{x}'$, in each case.

This measure (2.10) is equivalent to $\lambda_1 \times \ldots \lambda_p \times sign\left(det\left(\mathrm{E}\left(\mathbf{x}'\mathbf{y}'^T\right)\right)\right)$, where the covariance matrix $\mathrm{E}\left(\mathbf{x}'\mathbf{y}'^T\right)$ contains information about the dependence of two vectors on one another and the λ_k $(k = 1, \ldots, p)$ terms refer to the singular values of $\mathrm{E}\left(\mathbf{x}'\mathbf{y}'^T\right)$ (Rivest 1982, Fisher & Lee 1983). This can be thought of as a normalising constant multiplied by the $det\left(\mathrm{E}\left(\mathbf{x}'\mathbf{y}'^T\right)\right)$ term (Fisher & Lee 1986). This leads to the spherical correlation coefficient (Fisher & Lee 1986)

$$\rho_q = \frac{det\left(\mathrm{E}\left(\mathbf{x}'\mathbf{y}'^T\right)\right)}{\sqrt{det\left(\mathrm{E}\left(\mathbf{x}'\mathbf{x}'^T\right)\right)det\left(\mathrm{E}\left(\mathbf{y}'\mathbf{y}'^T\right)\right)}}. \tag{2.11}$$

We can estimate ρ_q with (Fisher et al. 1987)

$$\widehat{\rho}_q = \frac{det\left(\sum_{j=1}^{J}\mathbf{x}_{j,-}\mathbf{y}_{j,-}^T\right)}{\sqrt{det\left(\sum_{j=1}^{J}\mathbf{x}_{j,-}\mathbf{x}_{j,-}^T\right)det\left(\sum_{j=1}^{J}\mathbf{y}_{j,-}\mathbf{y}_{j,-}^T\right)}}, \tag{2.12}$$

where $\mathbf{x}_{j,-}$ indicates we have deducted the estimated spherical mean (2.3) from the directional sample. We will use spherical correlation, in much the same way as classical correlation, to check independence between populations in the directional inference tests discussed later. We now have a set of descriptive statistical tools at our disposal, which will allow us to evaluate samples of directional data.

2.3. Randomisation tests

Before we discuss spherical theory and inference tests, we first introduce an alternative (albeit somewhat naive) approach for analysing a sample of directional data to see if the group conceivably has a common direction (we reconsider this later in Section 2.4.2).

When a model is investigated using a classical hypothesis test, we can regard it as alternative to a *null hypothesis* (an assumption about a population) of randomness. That is, the model under investigation suggests that there will be a tendency for a certain type of pattern to appear in the data. We look for evidence to support the model by determining the chance that the pattern occurs under a null hypothesis of randomness, that is it came about purely by chance.

The general practice in hypothesis testing is to compute a *test statistic*, using sample data. In a classical hypothesis test, we typically compare the observed test statistic with a theoretical value obtained from the distribution of the test statistic. A rejection region can be selected, based on the set of values of the test statistic that are contradictory to the null hypothesis. The probability that values of the test statistic lie in the rejection region is called the *significance level*. This procedure allows us to decide if the null hypothesis is credible, given the limitations of the data.

Alternatively and equivalently, we can use a *p-value* to assess our assumptions. We can test the validity of the null hypothesis, by assuming it to be true and calculating the probability of observing a value of the test statistic as extreme or more so than the one actually observed. This probability is the p-value. If the p-value is small, that is, if the p-value is less than a predetermined significance level (0.05 is used in this thesis), then either the null hypothesis is true and we have had the good fortune to observe a rare event or more likely, the null hypothesis is false.

Randomisation testing is a way of determining whether the null hypothesis is reasonable in situations where classical hypothesis tests cannot be used. A test statistic is selected to measure the extent to which the data shows the pattern in question (Manly 1997). An observed test statistic is calculated for the data. This data is then randomly reordered many times. Each time, a test statistic is calculated from the reordered data. This allows a distribution of randomised test statistics to be built up and the observed test statistic is compared with this distribution. If the null hypothesis is true then all possible orders of the data are equally likely to have occurred. The observed data order is one of the equally likely orderings and the test statistic from the observed data should appear as a typical value from the randomisation distribution of test statistics obtained by randomly reordering the data. If this is not the case, the test statistic for the observed data is 'significant'. The null hypothesis is discredited and the alternative hypothesis is considered more reasonable.

In the context of the self-organising model, randomisation tests have an advantage over standard statistical methods. With randomisation tests, it is relatively simple to take into account the peculiarities of the situation using non-standard test statistics.

This particular randomisation test is designed to test the idea that a small number of individuals with knowledge of the direction of a goal, can successfully guide a large number of naive individuals to the goal (see Chapter 4 for a detailed discussion). We aim to evaluate how our data has evolved in time and gain some indication as to whether the knowledgeable members have been able to influence the group during the time period of the simulation. We do this by looking at the directions of the naive group members.

The null hypothesis is that the sample of orientations of the ignorant individuals is random. The alternative hypothesis is that the ignorant individual's direction of travel coincides with the goal direction.

We define an appropriate test statistic to assess the null hypothesis. Once we have the sample of average group directions at each timestep, the mean direction of this sample of group directions is calculated (2.3). We define the angle between the group direction of the sample and the x-axis, as ψ. The angle ψ can be calculated using the scalar product of vectors (where $\psi \in [0, \pi]$). If the group of directions is relatively aligned with the direction of the goal, the angle ψ will be small. The observed angle ψ_{obs} is calculated directly from the sample. We define the test statistic, T, as:

$$T = \frac{\psi}{p_{group}}, \tag{2.13}$$

where p_{group} is the polarisation of the directional sample. A cohesive group ($p_{group} \to 1$) heading towards the direction of the x-axis ($\psi \to 0$) will lead to low values of the test statistic (2.13). As the group becomes more disorganised, the value of the test statistic will increase. By using p_{group}, we are taking into account the spread of the group.

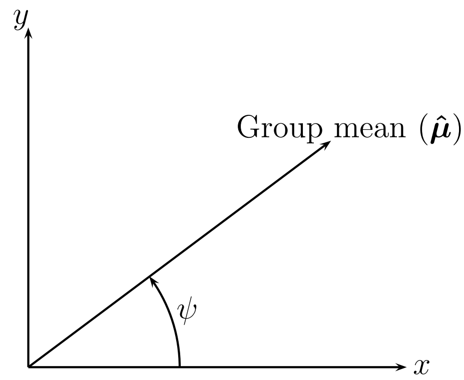

FIGURE 2.1. Simplified diagram of ψ (the angle between the spherical mean of the sample and the x-axis).

Random reorderings of the data are generated from the polar coordinates (θ_i and ϕ_i, $i = 1, ..., N$) of the sample of group directions. These polar coordinates are reordered 1000 times (θ_i and ϕ_i are reordered separately, as the two have different ranges). From these reordered samples, 1000 test statistics can be calculated to generate the empirical reference distribution. The observed value of T ($T_{obs} = \psi_{obs}/p_{group}$) is compared with this distribution, to decide if T is a typical value from the reference distribution. A p-value can be calculated, small p-values lead to the conclusion that the pattern in the data is unlikely to have arisen by chance alone. In this case, we have a one-sided test, as values of ψ close to zero and p_{group} close to 1 support the alternative alignment hypothesis that the naive individuals are moving towards the goal.

This randomisation test will be used in Chapter 4.

2.4. Inference tests for directional data

A statistical hypothesis is a claim or assumption about a population that can be tested by drawing a random sample from that population (Harraway 1997). The directional cases have analogies with the familiar univariate hypothesis tests (inference tests), which are taught in undergraduate courses.

In this section, we shall generalise to q dimensions ($\mathbb{R}^q$) and consider probability distributions on the surfaces of hyperspheres. The specific cases for the spherical ($q = 3$) and circular ($q = 2$) are therefore easily obtained.

We will discuss three important directional inference tests and show how these tests are derived. We discuss tests to assess whether a sample of directions is from a population with a particular average direction; whether several sets of directions can share a common mean direction; and whether two samples of directions share a common population polarisation. In order to this, we first discuss relevant background theory, before deriving these inference tests.

2.4.1. Background theory. We start with some essential background theory, which is sketched out in Watson (1983*b*). We commence by defining the central limit theorem for the multivariate case, along with key definitions.

Let the random vector $\mathbf{x} \in \mathbb{R}^q$, $\mathrm{E}(\mathbf{x}) = \boldsymbol{\mu}_p$ and $\mathrm{E}\left(\mathbf{x} - \boldsymbol{\mu}_p\right)\left(\mathbf{x} - \boldsymbol{\mu}_p\right)^T = \Sigma$, a $q \times q$ matrix. Generally, we will be interested in $q = 2, 3$, but we will present the results in a general format. Thus, $\boldsymbol{\mu}_p$ denotes the mean position and Σ the covariance matrix for the random coordinates. Note $\boldsymbol{\mu} = \boldsymbol{\mu}_p / |\boldsymbol{\mu}_p|$ is a unit vector giving the mean direction.

Let $\mathbf{x}_i$ ($i = 1, \ldots, N$) be a random sample of independent observations on the vector $\mathbf{x}$. Define the resultant $\mathbf{S}_N = \sum_{i=1}^{N} \mathbf{x}_i$ (2.2).

Theorem 1. *Multivariate central limit theorem (MCLT)*

As $N \to \infty$,

$$\sqrt{N}\left(\frac{\mathbf{S}_N}{N} - \boldsymbol{\mu}_p\right) \xrightarrow{D} N_q\left(\mathbf{0}, \Sigma\right)$$

where $N_q(\boldsymbol{\mu}_p, \Sigma)$ *denotes a Gaussian distribution in* $\mathbb{R}^q$ *with mean vector* $\boldsymbol{\mu}_p$ *and covariance matrix* Σ .

See eg. Anderson (1958) for proof.

We define a unit hypersphere as $\Omega_q = \{\mathbf{x} : \mathbf{x} \in \mathbb{R}^q, \quad |\mathbf{x}| = 1\}$ and suppose $\mathbf{x}, \boldsymbol{\mu} \in \Omega_q$. Note all vectors in Ω_q are of unit length. We will be using these vectors to indicate direction only. The spherical mean direction of the n observations $\mathbf{x}_i$ ($\mathbf{x}_i \in \Omega_q$) is $\hat{\boldsymbol{\mu}} = \mathbf{S}_N/|\mathbf{S}_N|$ (2.3).

Let the set $\{\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_q\}$ be an orthonormal basis for $\mathbb{R}^q$. Then any random vector $\mathbf{x}$ from Ω_q can be decomposed as

$$\mathbf{x} = c\boldsymbol{\epsilon}_q + (1 - c^2)^{\frac{1}{2}}\boldsymbol{\xi}_{q-1} \tag{2.14}$$

where c is a scalar random variable ($-1 \leq c \leq 1$) and $\boldsymbol{\xi}_{q-1}$ is a unit vector in the space spanned by $\{\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_{q-1}\}$. We assume that the random vector $\mathbf{x}$ is distributed symmetrically about the mean direction $\boldsymbol{\mu}$.

We adjust the basis so that $\boldsymbol{\epsilon}_q = \boldsymbol{\mu}$. Hence, (2.14) becomes

$$\mathbf{x} = c\boldsymbol{\mu} + (1 - c^2)^{\frac{1}{2}}\boldsymbol{\xi}_{q-1}. \tag{2.15}$$

where $c = \boldsymbol{\mu}^T\mathbf{x}$. This is illustrated in Figure 2.2.

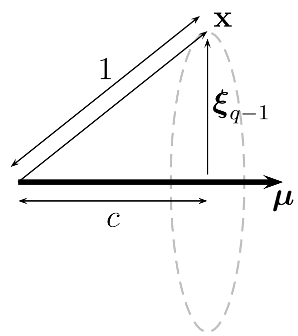

Figure 2.2. Diagram of Equation (2.15) in three-dimensional space ($q = 3$). The space $\{\mathbf{x} : \mathbf{x} \in \mathbb{R}^3, \quad |\mathbf{x}| = 1, \quad \mathbf{x} \perp \boldsymbol{\mu}\}$ is indicated by the ellipse and c represents the length of the projection of $\mathbf{x}$ onto $\boldsymbol{\mu}$.

From the assumption of the symmetrical distribution of $\mathbf{x}$, $\boldsymbol{\xi}_{q-1}$ is uniform on the surface $\{\mathbf{x} : \mathbf{x} \in \mathbb{R}^q, \quad |\mathbf{x}| = 1, \quad \mathbf{x} \perp \boldsymbol{\mu}\}$, where $\perp$

indicates $\mathbf{x}$ is orthogonal to $\boldsymbol{\mu}$. Now

$$\mathrm{E}\left(\mathbf{x}\right) = \boldsymbol{\mu}\mathrm{E}\left(c\right) \tag{2.16}$$

because of the rotational symmetry of $\mathbf{x}$. Thus $\mathrm{E}\left(c\right)$ is a measure of the concentration (degree of alignment or polarisation) of the distribution on Ω_q about $\boldsymbol{\mu}$ and can be estimated by $|\mathbf{S}_N|/N$ (2.4). 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. Hence, $0 \leq \mathrm{E}\left(c\right) \leq 1$.

The terms 'concentration' and 'polarisation' (2.4) refer to the same property. The key idea is to measure the spread of the directions around some average direction, which can also be though of as reflecting the degree of alignment of these vectors. In the spherical statistical literature, the term 'concentration' is used. We will use the term 'polarisation' consistently in this thesis, to avoid confusion.

Consequently,

$$\begin{aligned} \mathrm{E}\left(\mathbf{S}_N\right) &= \mathrm{E}\left(\sum_{i=1}^{N} \mathbf{x}_i\right) \\ &= N\boldsymbol{\mu}\mathrm{E}\left(c\right) \end{aligned} \tag{2.17}$$

and

$$\begin{aligned} \mathrm{Var}\left(\mathbf{x}\right) &= \mathrm{E}\left(\left(c - \mathrm{E}\left(c\right)\right)\boldsymbol{\mu} + (1-c^2)^{\frac{1}{2}}\boldsymbol{\xi}_{q-1}\right)\left(\left(c - \mathrm{E}\left(c\right)\right)\boldsymbol{\mu} + (1-c^2)^{\frac{1}{2}}\boldsymbol{\xi}_{q-1}\right)^T \\ &= \mathrm{E}\left(\boldsymbol{\mu}\boldsymbol{\mu}^T\left(c - \mathrm{E}\left(c\right)\right)^2 + (1-c^2)\boldsymbol{\xi}_{q-1}\boldsymbol{\xi}_{q-1}^T\right) \\ &= \boldsymbol{\mu}\boldsymbol{\mu}^T\mathrm{Var}\left(c\right) + \mathrm{E}\left(1-c^2\right)\mathrm{E}\left(\boldsymbol{\xi}_{q-1}\boldsymbol{\xi}_{q-1}^T\right) \end{aligned} \tag{2.18}$$

using the fact that $\boldsymbol{\mu} \perp \boldsymbol{\xi}_{q-1}$.

We need to determine $\mathrm{E}\left(\boldsymbol{\xi}_{q-1}\boldsymbol{\xi}_{q-1}^T\right)$. Let us assume that the random vector $\mathbf{y}$ is distributed uniformly on Ω_q. With a large sample size, we can approximate the distribution of $\sum_{i=1}^{N}\mathbf{y}_i$ using the MCLT (Theorem 1). We need to determine $\mathrm{E}\left(\mathbf{y}\right) = \boldsymbol{\delta}$ and $\mathrm{E}\left(\mathbf{y}-\boldsymbol{\delta}\right)\left(\mathbf{y}-\boldsymbol{\delta}\right)^T = \Sigma_y$. If A is any orthogonal matrix, then $A\mathbf{y}$ is also distributed uniformly on Ω_q.

Hence, $\boldsymbol{\delta} = A\boldsymbol{\delta}$ and $\Sigma_y = A\Sigma_y A^T$. The first condition implies that $\boldsymbol{\delta}$ must be $\mathbf{0}$. From the second, it is straightforward to see that $\Sigma_y = kI_q$ (see Proof 1 in Appendix B), where k is a constant and I_q is the $q \times q$ identity matrix. Now,

$$kq = \text{Trace}\,(\Sigma_y) = \text{Trace}\left(\text{E}\left(\mathbf{y}\mathbf{y}^T\right)\right)_{q\times q} = \text{E}\left(\text{Trace}\left(\mathbf{y}^T\mathbf{y}\right)\right) = 1. \quad (2.19)$$

Hence, $\Sigma = I_q/q$. Apply this to $\boldsymbol{\xi}_{q-1}$ and noting that the equivalent of the identity matrix in the space $\{\mathbf{x} : \mathbf{x} \in \mathbb{R}^q, \quad |\mathbf{x}| = 1, \quad \mathbf{x} \perp \boldsymbol{\mu}\}$ is $I_q - \boldsymbol{\mu}\boldsymbol{\mu}^T$. Hence, $\text{E}\left(\boldsymbol{\xi}_{q-1}\boldsymbol{\xi}_{q-1}^T\right) = \left(I_q - \boldsymbol{\mu}\boldsymbol{\mu}^T\right)/(q-1)$.

Using this, (2.18) becomes

$$\begin{aligned} \text{Var}\,(\mathbf{x}) &= \boldsymbol{\mu}\boldsymbol{\mu}^T\text{Var}\,(c) + \frac{1-\text{E}\,(c^2)}{q-1}\left(I_q - \boldsymbol{\mu}\boldsymbol{\mu}^T\right) \\ &= \Sigma. \end{aligned} \quad (2.20)$$

Thus from the MCLT, (2.17) and (2.20), we arrive at the following.

Lemma 2.

$$\frac{1}{\sqrt{N}}\left(\mathbf{S}_N - NE(c)\,\boldsymbol{\mu}\right) \xrightarrow{D} N_q\left(\mathbf{0},\, \Sigma\right)$$

as $N \to \infty$.

Now

$$\frac{1}{\sqrt{N}}\boldsymbol{\mu}^T\left(\mathbf{S}_N - N\text{E}\,(c)\,\boldsymbol{\mu}\right) = \frac{1}{\sqrt{N}}\left(\boldsymbol{\mu}^T\mathbf{S}_N - N\text{E}\,(c)\right), \quad (2.21)$$

and

$$\begin{aligned} \text{Var}\left(\frac{1}{\sqrt{N}}\left(\boldsymbol{\mu}^T\mathbf{S}_N - N\text{E}\,(c)\right)\right) &= \frac{1}{N}\text{Var}\left(\boldsymbol{\mu}^T\mathbf{S}_N\right) \\ &= \frac{1}{N}\left(\text{Var}\left(\sum_{i=1}^{N}\boldsymbol{\mu}^T\mathbf{x}_i\right)\right) \\ &= \text{Var}\,(c) \end{aligned} \quad (2.22)$$

since $\boldsymbol{\mu} \perp \boldsymbol{\xi}_{q-1}$. Combining Lemma 2 and (2.22), we obtain:

Lemma 3.

$$\frac{1}{\sqrt{N}}\left(\boldsymbol{\mu}^T\mathbf{S}_N - NE(c)\right) \xrightarrow{D} N\left(0, Var(c)\right)$$

as $N \to \infty$.

Let $\mathbf{S}_N^{\perp}$ be the component of $\mathbf{S}_N$ orthogonal to $\boldsymbol{\mu}$. Hence, $\mathbf{S}_N^{\perp} = \left(I_q - \boldsymbol{\mu}\boldsymbol{\mu}^T\right)\mathbf{S}_N$. For convenience, let $H = I_q - \boldsymbol{\mu}\boldsymbol{\mu}^T$. Now,

$$\begin{aligned} \mathrm{E}\left(\mathbf{S}_N^{\perp}\right) &= \mathrm{E}\left(H\mathbf{S}_N\right) \\ &= NH\boldsymbol{\mu}\mathrm{E}(c) \\ &= \mathbf{0}, \end{aligned} \tag{2.23}$$

as $H\boldsymbol{\mu} = \mathbf{0}$. The matrix H is symmetric and idempotent, so

$$\begin{aligned} \mathrm{Var}\left(\mathbf{S}_N^{\perp}\right) &= \mathrm{Var}\left(H\mathbf{S}_N\right) \\ &= NH\Sigma H^T \\ &= N\frac{(1-\mathrm{E}\left(c^2\right))}{q-1}\left(I_q - \boldsymbol{\mu}\boldsymbol{\mu}^T\right). \end{aligned} \tag{2.24}$$

Note that:

$$\begin{aligned} \sqrt{N}H\left(\frac{\mathbf{S}_N}{N} - \mathrm{E}(c)\boldsymbol{\mu}\right) &= \frac{1}{\sqrt{N}}\left(H\mathbf{S}_N - N\mathrm{E}(c)H\boldsymbol{\mu}\right) \\ &= \frac{1}{\sqrt{N}}\mathbf{S}_N^{\perp}. \end{aligned} \tag{2.25}$$

Using (2.23), (2.24) and (2.25) in conjunction with Lemma 2, we obtain Lemma 4.

Lemma 4.

$$\frac{1}{\sqrt{N}}\left(\mathbf{S}_N^{\perp}\right) \xrightarrow{D} N_q\left(\mathbf{0}, \frac{1-E\left(c^2\right)}{q-1}\left(I_q - \boldsymbol{\mu}\boldsymbol{\mu}^T\right)\right) \tag{2.26}$$

as $N \to \infty$.

The covariance matrix of $\mathbf{S}_N^{\perp}$ (which is $q \times q$) is of rank $(q-1)$, so an immediate consequence of Lemma 4 is the following:

Lemma 5. *For large* N, $\frac{1}{N}\left(|\mathbf{S}_N^{\perp}|^2\right)$ *is distributed approximately like* $\frac{1-E(c^2)}{q-1}\chi^2_{q-1}$.

These results can be used to obtain Lemma 6:

Lemma 6. *As* $N \to \infty$,

$$\sqrt{N}\left(\frac{|\mathbf{S}_N|}{N} - E(c)\right) \xrightarrow{D} N\left(0,\, Var(c)\right).$$

Proof. Using the Law of Large Numbers, we have the following statistics converging in probability,

$$\frac{\mathbf{S}_N}{N} \xrightarrow{p} \boldsymbol{\mu}\mathrm{E}(c), \quad \frac{|\mathbf{S}_N|}{N} \xrightarrow{p} \mathrm{E}(c) \quad \text{and} \quad \frac{\mathbf{S}_N}{|\mathbf{S}_N|} \xrightarrow{p} \boldsymbol{\mu}. \tag{2.27}$$

Write

$$\begin{aligned}
\sqrt{N}\left(\frac{|\mathbf{S}_N|}{N} + \mathrm{E}(c)\right)\left(\frac{|\mathbf{S}_N|}{N} - \mathrm{E}(c)\right) &= \sqrt{N}\left(\frac{|\mathbf{S}_N|^2}{N^2} - \mathrm{E}(c)^2\right) \\
&= \sqrt{N}\left(\left(\frac{\left(\boldsymbol{\mu}^T\mathbf{S}_N\right)^2}{N^2} - \mathrm{E}(c)^2\right) + \frac{|\mathbf{S}_N^{\perp}|^2}{N^2}\right) \\
&= \sqrt{N}\left(\frac{\left(\boldsymbol{\mu}^T\mathbf{S}_N\right)}{N} - \mathrm{E}(c)\right) \\
&\quad \times \left(\frac{\left(\boldsymbol{\mu}^T\mathbf{S}_N\right)}{N} + \mathrm{E}(c)\right) + \frac{|\mathbf{S}_N^{\perp}|^2}{N^{3/2}}. \quad (2.28)
\end{aligned}$$

Applying (2.27) as $N \to \infty$,

$$\begin{aligned}
\frac{|\mathbf{S}_N|}{N} + \mathrm{E}(c) &\xrightarrow{p} 2\mathrm{E}(c) \\
\text{and } \frac{\boldsymbol{\mu}^T\mathbf{S}_N}{N} + \mathrm{E}(c) &\xrightarrow{p} \boldsymbol{\mu}^T\boldsymbol{\mu}\mathrm{E}(c) + \mathrm{E}(c) = 2\mathrm{E}(c). \quad (2.29)
\end{aligned}$$

Moreover, using Lemma 4,

$$\frac{|\mathbf{S}_N^{\perp}|^2}{N^{3/2}} = \frac{1}{\sqrt{N}}\frac{|\mathbf{S}_N^{\perp}|^2}{N} \xrightarrow{p} 0 \text{ , as } N \to \infty. \tag{2.30}$$

Consequently, taking limits in (2.28) and applying (2.29) and (2.30), we can deduce that the limiting distribution of $\sqrt{N}\left(\frac{|\mathbf{S}_N|}{N} - \mathrm{E}(c)\right)$ is the same as that of $\sqrt{N}\left(\frac{(\boldsymbol{\mu}^T\mathbf{S}_N)}{N} - \mathrm{E}(c)\right)$. Thus Lemma 6 follows.

□

This is essential background material, where the details have been enlarged and expanded upon from Watson (1983*b*). We can now use this theory to derive hypothesis tests for directional data, which will be used in subsequent chapters to aid analysis of trajectories.

2.4.2. Derivation of directional hypothesis tests. We are now in a position to formulate hypothesis tests for directional data.

2.4.2.1. *Hypothesis test of a single direction.* From the multivariate central limit theorem, we can develop a test that a sample of directions (vectors) come from a population with spherical mean direction $\boldsymbol{\mu}_0$ ($\boldsymbol{\mu}_0 \in \Omega_q$). Practically speaking, if we suspect a sample of directions is pointing in a particular direction ($\boldsymbol{\mu}_0$), we can test the validity of this assumption with the following Lemma 7.

Given an independent sample of positions, $\mathbf{x}_1, \ldots, \mathbf{x}_N$ on Ω_q, let $\mathbf{S}_N = \sum_{i=1}^{N} \mathbf{x}_i$. The projection of $\mathbf{S}_N$ in the direction $\boldsymbol{\mu}_0$ is $c = \boldsymbol{\mu}_0^T \mathbf{S}_N$. Thus, the orthogonal component to this direction is $\mathbf{S}_N^{\perp\boldsymbol{\mu}_0} = \left(I_q - \boldsymbol{\mu}_0 \boldsymbol{\mu}_0^T\right) \mathbf{S}_N$. If the null hypothesis is true we can argue this orthogonal component will be small, as we expect $\mathbf{S}_N$ to point in the direction of the mean, $\boldsymbol{\mu}_0$.

From Lemma 5, it is clear that (for large N) $\frac{1}{N}\left(|\mathbf{S}_N^{\perp\boldsymbol{\mu}_0}|^2\right)$ is distributed approximately as $\frac{1-E\left(c^2\right)}{q-1}\chi^2_{q-1}$. We have a procedure for testing the null hypothesis.

Lemma 7. *Under the hypothesis* $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$,

$$\left(\frac{1 - E\left(c^2\right)}{q-1}\right)^{-1} \frac{1}{N}\left(|\mathbf{S}_N^{\perp\boldsymbol{\mu}_0}|^2\right) \xrightarrow{D} \chi^2_{q-1} \tag{2.31}$$

as $N \to \infty$ *(Watson 1983*b*).*

We can estimate $\mathrm{E}\left(c^2\right)$ in (2.31) by $\sum_{i=1}^{N}\left(\hat{\boldsymbol{\mu}}^T \mathbf{x}_i\right)^2 / N$, as in Watson (1983*b*). It is possible to invert the test in Lemma 7 to construct a confidence region for the mean direction.

2.4.2.2. *Hypothesis test of equality of directions of several groups.* If we have several independent samples of direction, we may wish to test that these come from populations with a common mean direction. This is the directional analogue of one-way analysis of variance (ANOVA) in standard univariate statistics. Recall in one-way ANOVA the test statistic is based on the variability of the sample means, that are the estimates for the common population mean. The test statistic for a hypothesis about the equality of population means is written in terms of variances (or measures of spread). Therefore, it is no surprise that the analogous statistic in the directional context turns out to be written in terms of estimates of concentration. We want to compare the concentration of the estimate of the mean directions with the concentration from within the directional samples.

Let $\mathbf{S}_{N_k}$ be the resultant vector from the k-th sample $(k = 1, \ldots, K)$. The concentration in the k-th sample is estimated by $|\mathbf{S}_{N_k}|/N_k$. The K estimates of the mean directions are $\hat{\boldsymbol{\mu}}_k = \mathbf{S}_{N_k}/|\mathbf{S}_{N_k}|$. Since these estimates are not based on samples of the same size, we need to consider a weighted version to estimate the concentration or alignment of these estimates, just as weighting for different samples sizes is used in ANOVA. The statistic used in directional statistics for testing for common mean directions across K populations is

$$T = \sum_{k=1}^{K} w_k |\mathbf{S}_{N_k}| - |\sum_{k=1}^{K} w_k \mathbf{S}_{N_k}|, \tag{2.32}$$

where the positive weights w_k reflect the normalising that is required for the distribution of the orthogonal component (to the common population direction, $\boldsymbol{\mu}$) of the resultant. We shall derive the form of w_k shortly.

Effectively, the first term in (2.32) is a comparison of the weighted average of the individual sample concentrations (similar to the within sum of squares (SS) in ANOVA) and the second term reflects the concentration of the estimates of the sample means (the between SS). Heuristically speaking, if the K samples have similar directions, then $\sum_{k=1}^{K} w_k|\mathbf{S}_{N_k}|$ should be approximately $|\sum_{k=1}^{K} w_k\mathbf{S}_{N_k}|$, as each vector $\mathbf{S}_{N_k}$ is pointing in a similar direction. In this case, the value of T will be small.

Using (2.28) from Section 2.4.1, we can express:

$$\frac{|\mathbf{S}_{N_k}|}{N_k} = \frac{\boldsymbol{\mu}^T\mathbf{S}_{N_k}}{N_k} + \frac{1}{2\mathrm{E}\left(c_k\right)}\frac{|\mathbf{S}_{N_k}^{\perp}|^2}{N_k^2} + O_p\left(N_k^{-3/2}\right), \tag{2.33}$$

where c_k is a scalar random variable ($c_k = \boldsymbol{\mu}^T\mathbf{x}_k$) from the k-th sample ($-1 \leq c_k \leq 1$, see (2.15)). This is not a trivial result, see Proof 2 in Appendix B for details. We can then rewrite the first component of (2.32), using (2.33):

$$\sum_{k=1}^{K} w_k|\mathbf{S}_{N_k}| = \boldsymbol{\mu}^T\sum_{k=1}^{K} w_k\mathbf{S}_{N_k} + \frac{1}{2}\sum_{k=1}^{K}\frac{w_k}{\mathrm{E}\left(c_k\right)}\frac{|\mathbf{S}_{N_k}^{\perp}|^2}{N_k} + R_1,$$

where R_1 is the remainder terms in the Taylor series expansion. A necessary condition for $R_1 \to 0$ is that $\min(N_k) \to \infty$ ($k = 1, \ldots, K$). If we assume $N_k = \alpha_k N$ (α_k are positive constants that sum to 1), then

$$|\sum_{k=1}^{K} w_k\mathbf{S}_{N_k}| = \boldsymbol{\mu}^T\sum_{k=1}^{K} w_k\mathbf{S}_{N_k} + \frac{1}{2}\frac{|\sum_{k=1}^{K} w_k\mathbf{S}_{N_k}^{\perp}|^2}{\sum_{k=1}^{K} w_k N_k\mathrm{E}\left(c_k\right)} + R_2, \tag{2.34}$$

where R_2 is the remainder terms in the Taylor series expansion. Again, we require $\min(N_k) \to \infty$. The proof of this result is shown in Proof 3 in Appendix B. Consequently, (2.32) becomes

$$2T \approx \sum_{k=1}^{K}\frac{w_k}{\mathrm{E}\left(c_k\right)}\frac{|\mathbf{S}_{N_k}^{\perp}|^2}{N_k} - \frac{|\sum_{k=1}^{K} w_k\mathbf{S}_{N_k}^{\perp}|^2}{\sum_{k=1}^{K} w_k N_k\mathrm{E}\left(c_k\right)}. \tag{2.35}$$

Using Lemma 4, if we set:

$$w_k = \frac{(q-1)\,\mathrm{E}\left(c_k\right)}{1-\mathrm{E}\left(c_k^2\right)},$$

then

$$\left(\frac{w_k}{N_k\mathrm{E}\left(c_k\right)}\right)^{\frac{1}{2}}\left(\mathbf{S}_{N_k}^{\perp}\right) \xrightarrow{D} N_q\left(\mathbf{0},\, I_q - \boldsymbol{\mu}\boldsymbol{\mu}^T\right).$$

Now define $\mathbf{z}_{N_k}$ as:

$$\mathbf{z}_{N_k} = \left(\frac{w_k}{N_k \mathrm{E}\left(c_k\right)}\right)^{\frac{1}{2}} \mathbf{S}_{N_k}^{\perp}. \tag{2.36}$$

Clearly,

$$|\mathbf{z}_{N_k}|^2 = \left(\frac{w_k}{N_k \mathrm{E}\left(c_k\right)}\right) |\mathbf{S}_{N_k}^{\perp}|^2.$$

To simplify matters, define:

$$\lambda_k = \frac{N_k w_k \mathrm{E}\left(c_k\right)}{\sum_{k=1}^{K} N_k w_k \mathrm{E}\left(c_k\right)}, \tag{2.37}$$

so that $\lambda_k > 0$ and $\sum_{k=1}^{K} \lambda_k = 1$. We can now write

$$2T \approx \sum_{k=1}^{K} |\mathbf{z}_{N_k}|^2 - |\sum_{k=1}^{K} \lambda_k^{\frac{1}{2}} \mathbf{z}_{N_k}|^2. \tag{2.38}$$

Let A (consisting of components a_{rs}, such that $r, s = 1, \ldots, K$) be a $K \times K$ orthogonal matrix, whose last row consists of $(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_K})$. Set $\mathbf{y}_r = \sum_{s=1}^{K} a_{rs} \mathbf{z}_{N_s}$, where $\mathbf{z}_{N_s}$ is defined in (2.36). Then $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_K$ are independent and furthermore, the $\mathbf{y}_r$'s are distributed asymptotically as $N_q\left(\mathbf{0}, I_q - \boldsymbol{\mu}\boldsymbol{\mu}^T\right)$. Therefore

$$\begin{aligned} \sum_{r=1}^{K-1} |\mathbf{y}_r|^2 &= \sum_{r=1}^{K} |\mathbf{z}_{N_r}|^2 - |\sum_{r=1}^{K} \lambda_r^{\frac{1}{2}} \mathbf{z}_{N_r}|^2 \\ &\xrightarrow{D} \chi^2_{(K-1)(q-1)}. \end{aligned}$$

More detail is supplied in Proof 4 in Appendix B. This result yields

Lemma 8. *Assume the null hypothesis of equality of mean directions across independent samples is true. Then for large N_k $(k = 1, \ldots, K)$, $2T$ is distributed approximately as $\chi^2_{(K-1)(q-1)}$ (Watson 1983*b*).*

2.4.2.3. *Hypothesis test of equality of polarisations of two groups.* We can compare the degree of alignment (polarisation) between two samples of directions, to see if the polarisations are equal. Recall we measure polarisation via $\mathrm{E}\,(c)$ defined in (2.15) and we estimate this with $|\mathbf{S}_N|/N$.

From Lemma 6 and the additivity properties of the Normal distribution, the following is derived.

Lemma 9. *Let* $\mathbf{S}_{N_1}$ *and* $\mathbf{S}_{N_2}$ *be two independent resultants from samples of sizes* N_1 *and* N_2*, respectively. If the concentrations from the two distributions are the same (i.e.* $E(c_1) = E(c_2)$*), then*

$$\left(\frac{Var(c_1)}{N_1} + \frac{Var(c_2)}{N_2}\right)^{-\frac{1}{2}} \left(\frac{|\mathbf{S}_{N_1}|}{N_1} - \frac{|\mathbf{S}_{N_2}|}{N_2}\right) \xrightarrow{D} N(0,1), \qquad (2.39)$$

as N_1 *and* N_2 *tend toward infinity.*

We can use the left hand side of (2.39) 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 recalling $\mathrm{Var}(c) = \mathrm{E}(c^2) - \mathrm{E}(c)^2$ and using

$$\widehat{\mathrm{Var}}(c_k) = \frac{1}{N_k}\sum_{i=1}^{N_k}\left(\hat{\boldsymbol{\mu}}_k^T \mathbf{x}_{k,i}\right)^2 - \left(\frac{|\mathbf{S}_{N_k}|}{N_k}\right)^2,$$

where $\hat{\boldsymbol{\mu}}_k = \mathbf{S}_{N_k}/|\mathbf{S}_{N_k}|$ and $\mathbf{x}_{k,i}$ denotes the i-th observation from sample k; $k = 1, 2$. Lemma 9 is extendable to more than two groups by considering the variance of the standardised polarisations.

2.5. Summary

During the passage of this chapter, we have introduced statistical tools with the intention of using them to analyse samples of directional data. The results of the mathematical models discussed in following chapters, are essentially samples of directional data. We shall use these statistical tools to evaluate the model results.

Key ideas we shall use to evaluate our results and their purposes include:

- polarisations (2.4) and momenta (2.5) to measure the overall degree of alignment and rotation of groups, respectively;
- spherical mean (2.3) to measure average direction of the group;

- nearest-neighbour distance (2.7) and expanse (2.8) to measure the size of groups;
- net to gross displacement ratio (2.9) to measure the curvature of the paths that group members take;
- Lemma 7 to assess whether the sample of directions (of the group members) is from a population with a particular average direction or not;
- Lemma 8 to assess whether several samples of directions (from several groups) can be regarded as sharing a common mean direction;
- Lemma 9 to assess whether two samples of directions (from two groups) can be considered as having the same population polarisation (concentration).

Part of the aim in introducing these tools is to raise awareness amongst the community of researchers interested in collective animal motion to the existence and efficiency of these techniques. In particular, the directional hypothesis tests discussed in Section 2.4 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 could be used for studying real animal data.

The recent paper of Riley et al. (2005) is the beginning of the advanced recording of animal movement. The authors use radar and attach harmonic transponders to honeybees and track individual flight paths of honeybees moving in space and time, 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 anticipate that given the three dimensional coordinates of flight paths, it would be straightforward to analyse the trajectories using the methods discussed in this chapter. These tests could help in assessing whether the foragers are heading towards the food source and with what degree of precision. Couzin & Franks (2003) filmed army ant (*Eciton burchelli*) trails with a digital video film camera, to explore traffic flow. From the recorded data on the centre of each ant, they were able to reconstruct individual trajectories. These inference tests presented in this chapter can be used in two-dimensions, to analyse such data.

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, possibly enhanced by appropriate weights). The sequence of means should orientate towards the direction indicated by the knowledgeable individuals, as time progresses. Alternatively, there have been developments in spline techniques for smoothing and interpolating directional data (Watson 1983*a*) that may be applicable.