

Chapter 4

Statistical Analysis of Cellular Directional Movement: Application for Research of Single Cell Movement



Masahiro Kanai, Kazuo Tonami, and Hideto Tozawa

4.1 Single Cell Movement

Cell movement plays a fundamental role in physiological collective phenomena, especially in self-organization. In earlier studies, it has been mainly discussed whether cell movement may be considered as random walk or not. We note that by *random walk* or *Brownian motion*, we denote the random motion of small particles driven by external force although these two terms have a different mathematical definition.

Due to Einstein's theory of Brownian motion, the mean squared displacement (MSD), which we can estimate from experimental data, is endorsed as a criterion for the motility to be random walk. The MSD of a cell trajectory $x(t)$, the position of a single cell at time t , is defined as the following equation:

$$\text{MSD}(t) = \langle |x(t) - x(0)|^2 \rangle,$$

where $|x|$ and $\langle \dots \rangle$ denote, respectively, the norm of vector x and the expectation. The MSD of a Brownian motion is proportional to time: $\text{MSD}(t) \propto t$, whereas that

M. Kanai (✉)
Kurume Institute of Technology, Kurume, Fukuoka, Japan
e-mail: kanai@kurume-it.ac.jp

K. Tonami
Graduate School of Medicine and Faculty of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo, Japan
e-mail: kazutonami@g.ecc.u-tokyo.ac.jp

H. Tozawa
Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Bunkyo-ku, Tokyo, Japan
e-mail: htozawa@g.ecc.u-tokyo.ac.jp

of linear movement with a constant velocity is proportional to the square of time: $\text{MSD}(t) \propto t^2$. These two cases are called *diffusive* and *ballistic*, respectively, and the others are anomalous.

Recent studies [1, 2, 4, 6–9], based on experimental data of high accuracy, have tended to claim that cells move with a directional trend and a persistent random walk is suitable for modeling cell movement. The persistence parameter included in those models appears in the MSD and reveals a crossover from persistent movement to random. We will cover the persistent movement in the next section.

On the other hand, we consider another 2-dimensional model in which a cell moves with a constant speed prescribed and decides his direction entirely randomly at each time step. This model requires another method for analysis, namely circular statistics. In the latter section, as well as the MSD, circular statistics will be applied for the sample of cell movement.

4.1.1 Persistent Random Walk

One of the simplest ways to introduce *the persistent random walk* is to adopt, as equation of motion for cells, a Langevin-type equation in 2 dimension:

$$\tau \frac{dv}{dt} = -v + \sqrt{2D}\xi, \quad (4.1)$$

where $v = v(t)$ denotes the velocity vector of a cell at time t , and a vector $\xi = \xi(t) = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$ represents the white noise, i.e., a random force which obeys the following properties:

$$\langle \xi \rangle = 0, \quad \langle \xi_i(s)\xi_j(t) \rangle = \begin{cases} \delta(s-t) & (i=j) \\ 0 & (i \neq j) \end{cases},$$

where $\langle \dots \rangle$ denotes the expectation. Here, $\delta(t)$ denotes Dirac's delta function which takes the value of 0 unless $t = 0$. The two components of the random force vector ξ are independent, and each has no auto-correlation. Note that one can arrange the coefficient of the friction term $-v$ to be 1 without loss of generality.

The parameter τ , taking positive values, does no longer represent the inertial mass because a cell can move by itself without any external force, and hence the relationship between the acceleration and the force is not retained in general. It is, however, still true that if τ is large the acceleration remains small. This means that the velocity will not change rapidly. Hence, τ should be regarded as a criterion for persistent movement and is called the persistence time since it has dimension of time.

The parameter D controls the strength of the random force and hence shows the motility of cells. From a macroscopic point of view, it corresponds exactly to the

so-called diffusion coefficient originally introduced from Fick's law. Actually, in Fick's second law, the diffusion coefficient manifests itself in a diffusion equation.

4.1.2 Mean Squared Displacement (MSD) and the Fürth Formula

We calculate the MSD for the persistent random walk introduced in (4.1).

The velocity auto-correlation function (VACF) is obtained by integrating (4.1) as

$$\langle v(t) \cdot v(0) \rangle = \frac{2D}{\tau} e^{-|t|/\tau}, \quad (4.2)$$

where $v_1 \cdot v_2$ denotes the inner product of vectors v_1 and v_2 , and $\langle \dots \rangle$ denotes the expectation. The coefficient $\frac{2D}{\tau}$ is determined from the Green–Kubo formula:

$$D = \frac{1}{2} \int_0^\infty \langle v(t) \cdot v(0) \rangle dt.$$

In general, one can obtain the MSD from the VACF as follows. Since

$$v = \frac{dx}{dt}$$

by definition, we have

$$\begin{aligned} \langle x(t)^2 \rangle &= \left\langle \left(\int_0^t v(s) ds \right)^2 \right\rangle \\ &= \left\langle \int_0^t v(s) ds \int_0^t v(s') ds' \right\rangle \\ &= \int_0^t \int_0^t \langle v(s) \cdot v(s') \rangle ds ds' \\ &= \int_0^t \int_0^t \langle v(s-s') \cdot v(0) \rangle ds ds'. \end{aligned} \quad (4.3)$$

Note that we can let $x(0) = (0, 0)$ without loss of generality. From (4.2), we calculate the integral and thus obtain the MSD referred to as the Fürth formula [3]:

$$\langle x(t)^2 \rangle = 4D(t - \tau(1 - e^{-t/\tau})). \quad (4.4)$$

We remark that, as already noted in [1, 8], the Fürth formula can be derived from other models than the Langevin equation.

In the short time region, i.e. for $t \ll \tau$, (4.4) is approximated as

$$\langle x(t)^2 \rangle \simeq \frac{2D}{\tau} t^2.$$

We note that from (4.2), the coefficient $\frac{2D}{\tau}$ coincides with $\langle v(0) \cdot v(0) \rangle$. This leads to $\sqrt{\langle x(t)^2 \rangle} \simeq \sqrt{\langle v(0) \cdot v(0) \rangle} t$, and hence this time region is called ballistic.

In the large time region, i.e. for $t \gg \tau$, (4.4) is approximated as

$$\langle x(t)^2 \rangle \simeq 4Dt.$$

This time region is called diffusive. In this region, the persistence time τ approximately vanishes in the MSD. This suggests that the mobility of cells may be considered as random walk in the long run.

4.2 Circular Statistics

In this section, we introduce statistical methods for analysis of *circular data*: numeric data measured in the form of angles or two-dimensional orientations.

Now, we have a wide variety of circular data as a set of vectors/axes, e.g., wind directions, circadian rhythms, cell division axis, directional movement of animals, and so on. In contrast, most scientists may not be familiar with the methods to deal with those data. The latter part of this section will be hence devoted to cell migration data measured as planar vectors; however, we do not refer to the general ways how these data should be recorded.

This section is concerned with basic methods for statistical analysis of a single sample of circular data $\{\theta_1, \theta_2, \dots, \theta_n\}$ including methods for displaying and summarizing the sample. You may consider that the sample should present angles.

4.2.1 Raw Data Plot

First we consider the advantage of descriptive methods for statistical data. These enable us to gain an initial idea of the important characteristics of the sample: whether the sample does appear from a uniform distribution, from a unimodal distribution, or from a multimodal distribution. Moreover the distribution may be regarded as one of the fundamental distributions.

Here we refer to the *distribution* for a sample as the counts of sample points distributed over all possible values. Raw data plot is the first important step to

analyzing the data because it implies that we will have the *next* measurement in the range in which the major part of the data concentrate.

Example: Wind Direction

Table 4.1 gives the sample of wind directions at *Kurume City, Fukuoka, Japan* observed on March 8th, 2018. The angles, figured in degree, are measured clockwise from the north direction. Figure 4.1 shows an angular plot for the raw data given in Table 4.1. This diagram enables one to recognize the whole aspect of the data.

4.2.2 Histograms

The next step is to exploit histograms. *Histograms* are constructed as a type of bar plot for numeric data that group the data into bins, i.e. a series of intervals dividing the entire range of values. First, there are two types of histograms/diagrams: linear and angular, and then some variations in the angular diagram. A table of unprocessed numerics is often referred to as raw data, compared to processed data (Fig. 4.2).

1. A simple angular histogram is obtained by plotting bars each of which is centered at the midpoint of its grouping interval, with the length of the bar proportional to the relative frequency in the group.
2. The *rose diagram* is more commonly used for angular data than the angular histogram above is, in which each group is displayed as a sector. The radius of each sector is taken so as to be proportional to the square root of the relative frequency of the group; the area of the sector is thus proportional to the group frequency (Fig. 4.3).

Example: Wind Direction Statistics (Rose Diagram)

Figure 4.4 shows a rose diagram for the data in Table 4.1. As immediately seen, the rose diagram is named for its shape.

3. The *stem-and-leaf* diagram is a histogram retaining the raw data values; in particular, their mode is immediately found from the sample. Each stem, as a column sequence of integers, consists of the data falling in a 10° interval with the leaves being the data values in that interval, sorted in an increasing order. The stem-and-leaf diagram is superior in displaying data without detracting the individual measurements.

Example: Wind Direction Statistics (Stem–Leaf Diagram)

Figure 4.3 gives a stem–leaf diagram for the data in Table 4.1. In order to make this diagram, we process the raw data in Table 4.1 by dividing all the numerics by 30° . Accordingly, the stem consists of 12 items, and stems 0, 1, 2, 3, ... present actual angles 0° , 30° , 60° , 90° , ..., respectively. Then, for example, leaf 2 | 3 present actual angle 67.5° . We remark that this processing of the raw data changes numerics superficially but does not detract intrinsic statistics at all.

Table 4.1 Wind directions at *Kurume City, Fukuoka, Japan* on March 8, 2018 [5]

Time	Direction	Time	Direction	Time	Direction
0:10	45°	8:10	0°	16:10	22.5°
0:20	45°	8:20	337.5°	16:20	22.5°
0:30	45°	8:30	0°	16:30	0°
0:40	45°	8:40	337.5°	16:40	45°
0:50	45°	8:50	22.5°	16:50	22.5°
1:00	45°	9:00	0°	17:00	22.5°
1:10	22.5°	9:10	22.5°	17:10	22.5°
1:20	22.5°	9:20	0°	17:20	0°
1:30	45°	9:30	292.5°	17:30	22.5°
1:40	22.5°	9:40	45°	17:40	22.5°
1:50	22.5°	9:50	180°	17:50	22.5°
2:00	45°	10:00	225°	18:00	45°
2:10	45°	10:10	225°	18:10	22.5°
2:20	22.5°	10:20	157.5°	18:20	22.5°
2:30	22.5°	10:30	157.5°	18:30	0°
2:40	22.5°	10:40	157.5°	18:40	0°
2:50	22.5°	10:50	180°	18:50	45°
3:00	45°	11:00	135°	19:00	0°
3:10	45°	11:10	180°	19:10	0°
3:20	45°	11:20	180°	19:20	22.5°
3:30	45°	11:30	225°	19:30	22.5°
3:40	45°	11:40	202.5°	19:40	0°
3:50	45°	11:50	225°	19:50	0°
4:00	45°	12:00	247.5°	20:00	0°
4:10	45°	12:10	247.5°	20:10	22.5°
4:20	45°	12:20	247.5°	20:20	22.5°
4:30	45°	12:30	292.5°	20:30	22.5°
4:40	67.5°	12:40	270°	20:40	22.5°
4:50	22.5°	12:50	247.5°	20:50	0°
5:00	45°	13:00	247.5°	21:00	0°
5:10	22.5°	13:10	247.5°	21:10	22.5°
5:20	45°	13:20	247.5°	21:20	0°
5:30	45°	13:30	225°	21:30	0°
5:40	45°	13:40	225°	21:40	0°
5:50	0°	13:50	225°	21:50	22.5°
6:00	45°	14:00	225°	22:00	45°
6:10	22.5°	14:10	225°	22:10	22.5°
6:20	45°	14:20	225°	22:20	22.5°
6:30	45°	14:30	247.5°	22:30	22.5°
6:40	45°	14:40	225°	22:40	22.5°
6:50	45°	14:50	292.5°	22:50	22.5°

(continued)

Table 4.1 (continued)

Time	Direction	Time	Direction	Time	Direction
7:00	67.5°	15:00	337.5°	23:00	22.5°
7:10	45°	15:10	0°	23:10	22.5°
7:20	67.5°	15:20	90°	23:20	337.5°
7:30	45°	15:30	22.5°	23:30	0°
7:40	22.5°	15:40	0°	23:40	0°
7:50	0°	15:50	0°	23:50	0°
8:00	22.5°	16:00	22.5°	0:00	0°

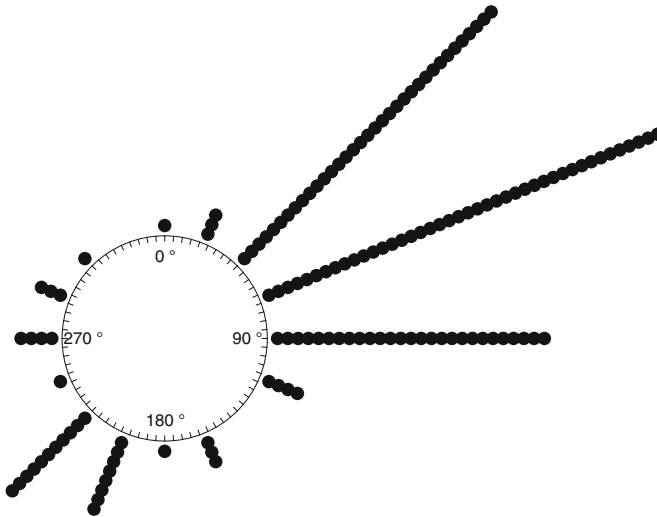


Fig. 4.1 An angular plot for the data on the raw data is given in Table 4.1. The angles are measured clockwise from the north direction

4. We introduce a *kernel density estimation*, which is one of nonparametric methods to estimate the distribution for a sample. This method is different from the histograms described above in that the contribution of each data point θ_k is defined as

$$\frac{1}{nh}w\left(\frac{\theta - \theta_k}{h}\right)d\theta,$$

where $w(\theta)$ presents a smooth function of θ taking the form of a bump, $d\theta$ denotes an infinitely small interval, and h controls the amount of smoothing effect. Then, the kernel density estimation $\hat{f}(\theta)$ is given by the sum of all the contributions:

$$\hat{f}(\theta)d\theta = \frac{1}{nh} \sum_{k=1}^n w\left(\frac{\theta - \theta_k}{h}\right)d\theta.$$

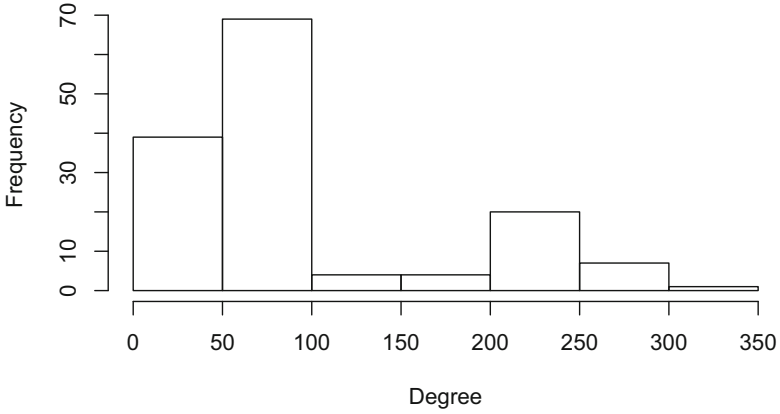


Fig. 4.2 A linear histogram for the data on the wind directions at Kurume City, Fukuoka, Japan. The binwidth is 50°. The raw data is given in Table 4.1

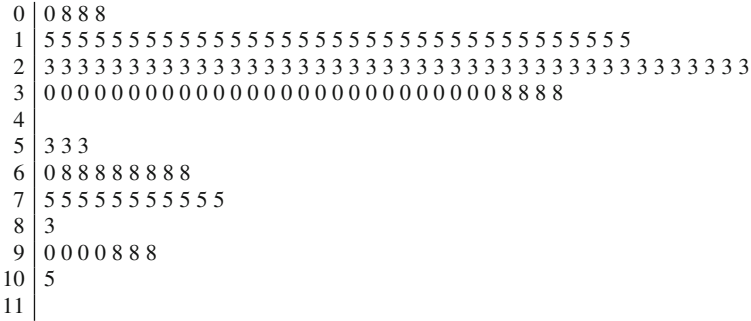


Fig. 4.3 A stem-leaf diagram for the wind directions at Kurume City, Fukuoka, Japan on March 8, 2018 [5]

Function $w(\theta)$ called the kernel function should be chosen so that its support D , a region where the function takes nonzero values, is small enough for $w(\theta)$ to ensure the condition

$$\int_D w(\theta)d\theta = 1.$$

However, the choice is not crucial to the density estimation but the magnitude of h is because it determines the amount of smoothing effect, namely, the larger h is, the more the density function gets blurred. It is, however, another issue how you determine the value of h .

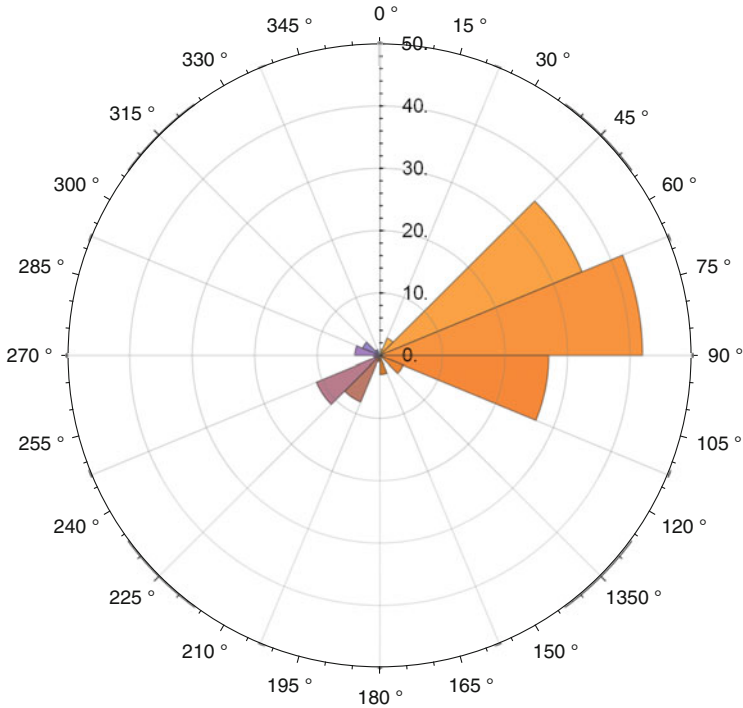


Fig. 4.4 A rose diagram for the data on the wind directions at *Kurume City, Fukuoka, Japan*. The raw data is given in [Table 4.1](#)

We note some important points in plotting histograms: Some choice of group boundaries can give rise to a serious distortion of the information about the modal groups observed in the sample; as well, the choice of the bin width should be carefully made so that we can anticipate the shape of the underlying distribution of the sample. It is, however, another issue whether your choice is correct or not.

4.2.3 Summary Statistics

The diagrams described above suggest the existence of the population from which the sample was drawn. Here we introduce basic quantities which describe important features of the sample distribution.

(A) Sample Circular Mean and Sample Circular Standard Deviation

The most popular as well as important statistic is the sample mean, which is ordinarily given by the arithmetic mean:

$$\frac{1}{n} \sum_{j=1}^n \theta_j.$$

Is it correct even for a sample of circular data? No, and it will be obvious if one considers, for example, the sample $\{10^\circ, 350^\circ\}$. We have the arithmetic mean 180° , but this cannot be acceptable.

The natural way is to consider that each circular data corresponds to a point located on the unit circle, and thereby the sample of circular data directly transforms into a sample of unit vectors and the sample mean for circular data is to be figured out from the mean vector. The mean vector in terms of vector addition is carried out as

$$C = \frac{1}{n} \sum_{j=1}^n \cos \theta_j, \quad S = \frac{1}{n} \sum_{j=1}^n \sin \theta_j, \quad R = \sqrt{C^2 + S^2},$$

where R should be chosen to be positive. Then, the mean direction $\bar{\theta}$ of the sample of unit vectors is obtained from

$$\cos \bar{\theta} = \frac{C}{R}, \quad \sin \bar{\theta} = \frac{S}{R}.$$

This mean direction $\bar{\theta}$ corresponds to the mean of the sample of circular data θ_i s.

On the other hand, R defined above presents the length of the resultant vector and is no longer a unit vector. R takes the value in the range $[0, 1]$; if $R = 1$, it means that all the vectors are in the same direction and hence all data θ_i s are coincident. By contrast, $R = 0$ does not always mean that all the unit vectors distribute in uniformly random directions. A simple counterexample is the sample data $\{10^\circ, 180^\circ, 350^\circ\}$. This point will become more clear when we consider statistics for axial data. We hence note that R cannot be a useful measure of deviation.

However, R may present the variance as

$$V = 1 - R$$

in the case that the sample shows a single modal distribution. We note that circular data is restricted to a finite range and so is the variance. Again, $V = 1$ does not immediately imply that the sample has a dispersed distribution.

The sample circular standard deviation is also defined as

$$v = \sqrt{-2 \log(1 - V)}.$$

(We may have other definitions for the standard deviation.) If V is nearly equal to 1, v can be well approximated by $\sqrt{2V}$.

(B) Advanced Circular Summary Statistics

In order to define advanced statistics, we need some mathematics in complex numbers. From the mean direction $\bar{\theta}$ and the resultant vector length R , we define the first trigonometric moment

$$m_1 = R(\cos \bar{\theta} + i \sin \bar{\theta}),$$

where $i = \sqrt{-1}$. In an analogous manner as the mean direction, the p th trigonometric moment is then defined as

$$m_p = R_p(C_p + iS_p),$$

where

$$C_p = \frac{1}{n} \sum_{i=1}^n \cos p\theta_i, \quad S_p = \frac{1}{n} \sum_{i=1}^n \sin p\theta_i, \quad R_p = \sqrt{C_p^2 + S_p^2}.$$

Note that $R_1 = R$.

Using Euler's formula in complex analysis, $e^{i\theta} = \cos \theta + i \sin \theta$, we have a simple expression:

$$C_p + iS_p = \frac{1}{n} \sum_{j=1}^n e^{ip\theta_j}.$$

Then, we introduce argument μ_p as

$$R_p e^{i\mu_p} = \frac{1}{n} \sum_{j=1}^n e^{ip\theta_j}.$$

In particular, $\mu_1 = \bar{\theta}$, but however $\mu_p \neq p\bar{\theta}$ for $p \geq 2$ in general.

The centered sample trigonometric moments are also defined as

$$m'_p = R'_p(C'_p + iS'_p),$$

where

$$C'_p = \frac{1}{n} \sum_{j=1}^n \cos p(\theta_j - \bar{\theta}), \quad S'_p = \frac{1}{n} \sum_{j=1}^n \sin p(\theta_j - \bar{\theta}),$$

and

$$R'_p = \sqrt{(C'_p)^2 + (S'_p)^2}.$$

Since a little calculation leads to $C'_1 = R$ and $S'_1 = 0$, we have $m'_1 = R$.

Using Euler's formula, we have

$$R'_p e^{i\mu'_p} = \frac{1}{n} \sum_{j=1}^n e^{ip(\theta_j - \bar{\theta})} = \frac{e^{-ip\bar{\theta}}}{n} \sum_{j=1}^n e^{ip\theta_j} = R_p e^{i(\mu_p - p\bar{\theta})},$$

and therefore $\mu'_p = \mu_p - p\bar{\theta}$, and $R'_p = R_p$.

As for a unimodal distribution of the sample, we exploit the first and second centered trigonometric moments, defining advanced statistics: the sample circular dispersion

$$\delta = \frac{1 - R'_2}{2(R'_1)^2},$$

the sample circular skewness

$$s = \frac{R'_2 \sin(\mu'_2 - 2\mu'_1)}{\sqrt[3]{1 - R'_1}},$$

and the sample circular kurtosis

$$k = \frac{R'_2 \cos(\mu'_2 - 2\mu'_1) - (R'_1)^4}{(1 - R'_1)^2}.$$

The circular dispersion δ is concerned with confidence interval for the sample mean direction. The skewness s presents the asymmetry of the sample distribution. The kurtosis k presents the peakedness of the sample distribution, and so it is also called peakedness.

Example: Wind Direction Statistics

We have summary quantities for the wind directions given in Table 4.1 as follows: the mean direction $\bar{\theta} = 71.7^\circ$, the sample circular variance $V = 0.43$, and the circular standard deviation $v = 1.1$. See Figs. 4.2 and 4.4, and verify these results.

4.2.4 Probability Models

Statistical analysis, especially statistical inference, of a sample data is based on the probability of it being obtained. Probability in statistics takes a role in formulating uncertainty of the data, i.e., the data we obtain is drawn from an underlying population. Each sample of the population contains the probability that one will draw it, and we call the set of the probability the probability distribution or distribution simply.

Probability models provide with a distribution for the data a priori. Some probability models have parameters to be inferred from the sample, and other models called nonparametric do not. One usually applies the normal distribution for linear data, then obtaining acceptable results. For circular data, some models appear as a counterpart of the normal distribution for linear data, and the von Mises distribution described below is one of them.

Example: von Mises Distribution

Figure 4.5 shows a von Mises distribution. This distribution is a continuous probability distribution on the circle and was introduced as a circular analogue of the normal distribution. The probability density function is defined by

$$f(\theta) = \frac{e^{\kappa \cos(\theta - \mu)}}{2\pi I_0(\kappa)} \quad (0 \leq \theta < 2\pi, 0 \leq \kappa < \infty),$$

where the parameters μ and $1/\kappa$ correspond, respectively, to the mean value and the variance. We note that the normalization is given by a modified Bessel function $I_0(\kappa)$. We have some analytic expressions for the function as

$$I_0(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} e^{\kappa \cos(\theta - \mu)} d\theta,$$

and

$$I_0(\kappa) = \sum_{r=0}^{\infty} (r!)^{-2} \left(\frac{\kappa}{2}\right)^{2r}.$$

However, these may not be useful for numerical calculation. The use of a software on computer is practical.

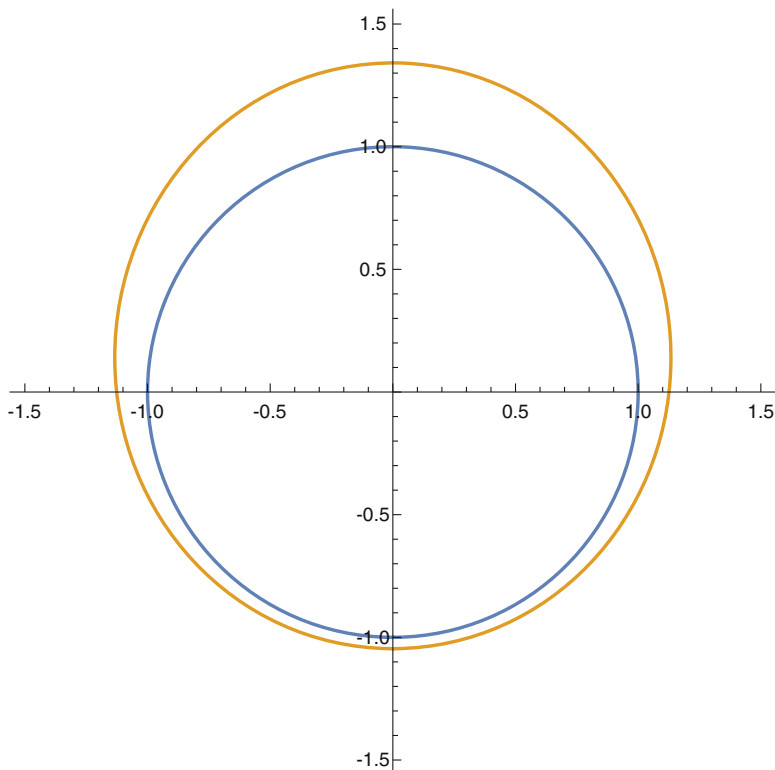


Fig. 4.5 The von Mises distribution. We illustrate a von Mises distribution in polar plot; the outer curve presents the value of probability increasing upon the unit circle (the inner curve)

For example, we apply the von Mises distribution to the wind direction data given in Table 4.1. As already calculated above, the circular mean $\mu = 71.7^\circ$ and the circular variance $1/\kappa = 0.43$. Figure 4.6 shows the probability density function in polar plot.

4.3 Application for Single Cell Movement

In this final section, we apply the circular statistical analysis introduced above for single cell movement. We obtain the sample of a single 3T3 cell in vitro. The 3T3 cell line is a spontaneously immortalized mouse fibroblast cell line established from mouse embryonic tissue.

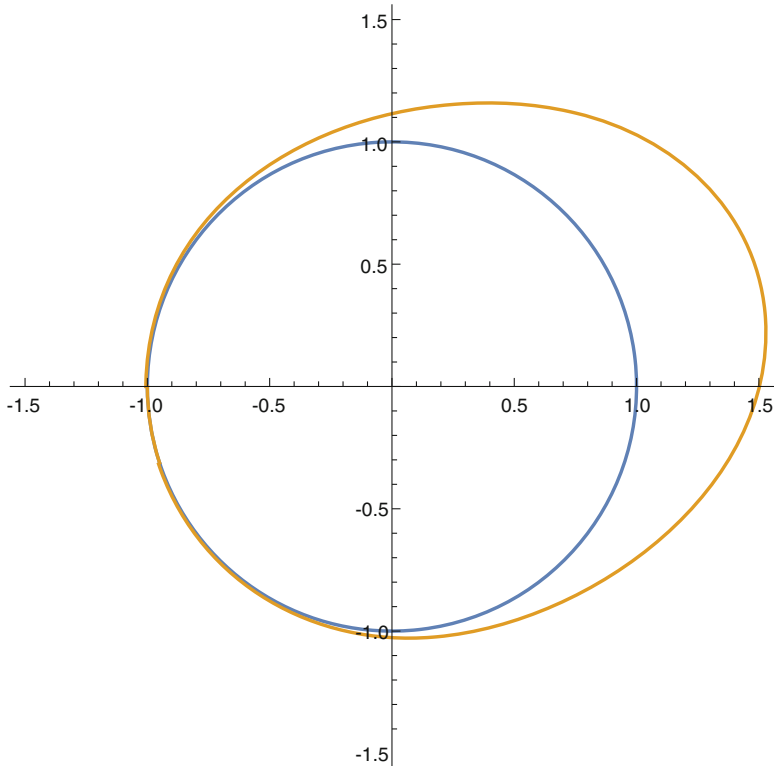


Fig. 4.6 The von Mises distribution applied for the wind direction data given in Table 4.1. Parameters are chosen so that the mean $\mu = 71.7^\circ$ and the variance $1/\kappa = 0.43$. We find that this fits with the rose diagram given in Fig. 4.4. To be precise, we illustrate the graph of polar equation $\rho = f(\theta) + 1$ with $\mu = 1.25$ and the variance $\kappa = 2.33$

Figure 4.7 shows the sample trajectory of a single 3T3 cell moving on a dish freely. We plot the data points by time lapse imaging of the cell movement and then connect them with line segments in time sequence.

Figure 4.8 shows the directions of motion obtained from the sample trajectory in both circular histogram and rose diagram. From the data, we have the mean circular mean direction 121° and the circular variance 0.93. Since the circular variance takes a value in between 0 and 1 and the greater it is the more random (irregular) the movement is, we consider from the present result that the 3T3 cell chooses the direction of movement with almost equal probability (Fig. 4.9).

Furthermore, we calculate the mean squared displacement from the sample trajectory data, so we test that the movement observed is a random walk or not. We use a log–log plot for the figure, and hence the line of the linear minimum mean squared error estimator gives the exponent λ for the MSD with respect to time: $\text{MSD} \propto t^\lambda$. The slope of the line is read as 1.0, i.e. $\lambda = 1.0$, and accordingly the result suggests that the present cell movement should be considered as diffusive. Hence

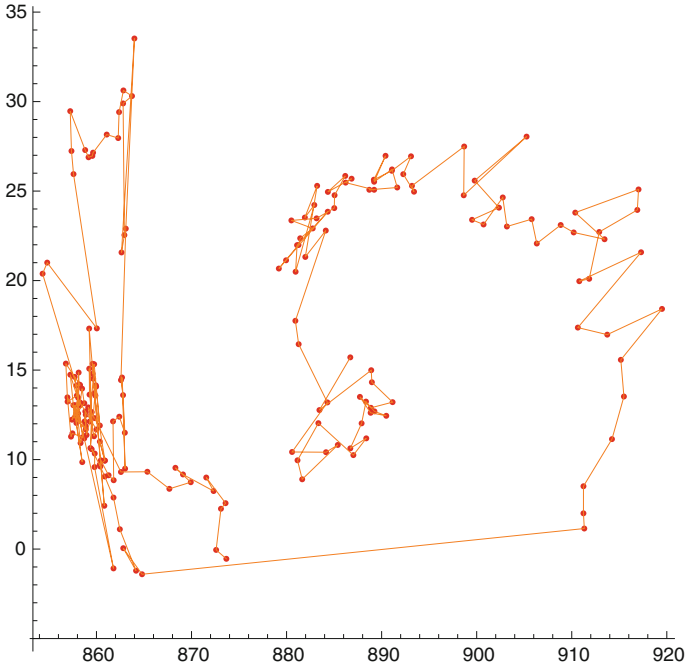


Fig. 4.7 The sample of a single 3T3 cell moving freely on a dish

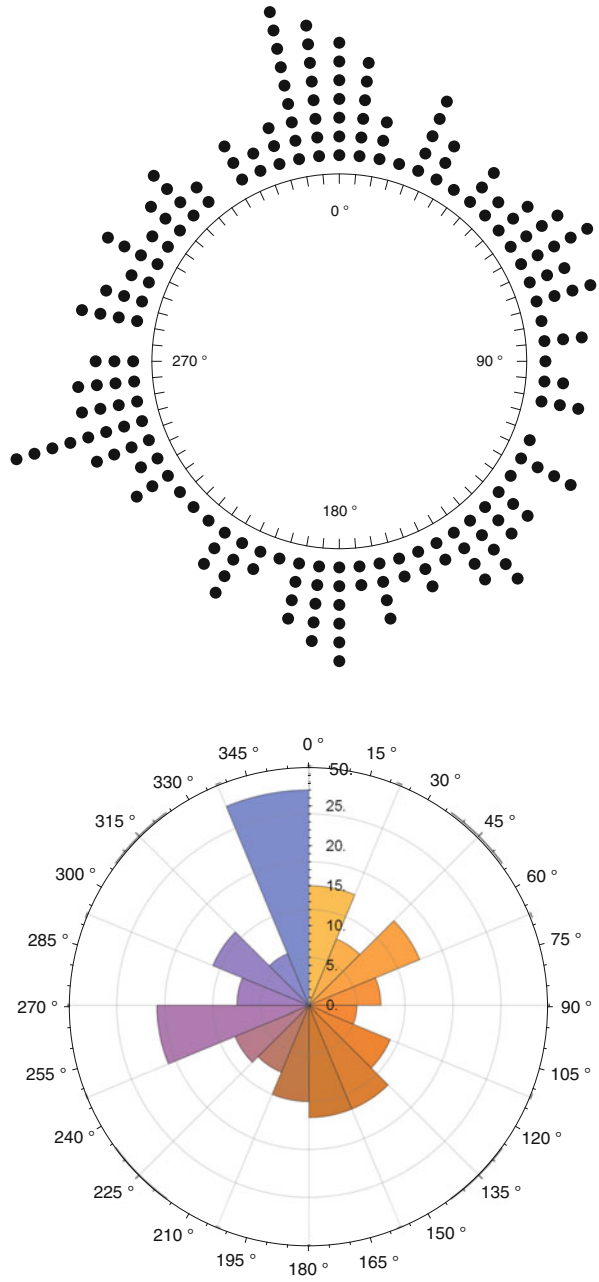
we conclude from the sample data given in Fig. 4.7 that the 3T3 cell moves as a random walk.

We finally make a remark on the calculation of the MSD from the sample data before closing the chapter. From a sample trajectory containing $N + 1$ points of position $(x(t), y(t))$ at time t , we actually compute the MSD as

$$\begin{aligned}
 \text{MSD}(\tau) &= \langle (x(t + \tau) - x(t))^2 + (y(t + \tau) - y(t))^2 \rangle \\
 &= \frac{1}{N - n + 1} \sum_{i=0}^{N-n} \left[(x((i + n)\Delta t) - x(i\Delta t))^2 \right. \\
 &\quad \left. + (y((i + n)\Delta t) - y(i\Delta t))^2 \right], \tag{4.5}
 \end{aligned}$$

where $\tau = n\Delta t$ ($n = 0, 1, 2, \dots, N$).

Fig. 4.8 The directions of motion obtained from the sample trajectory



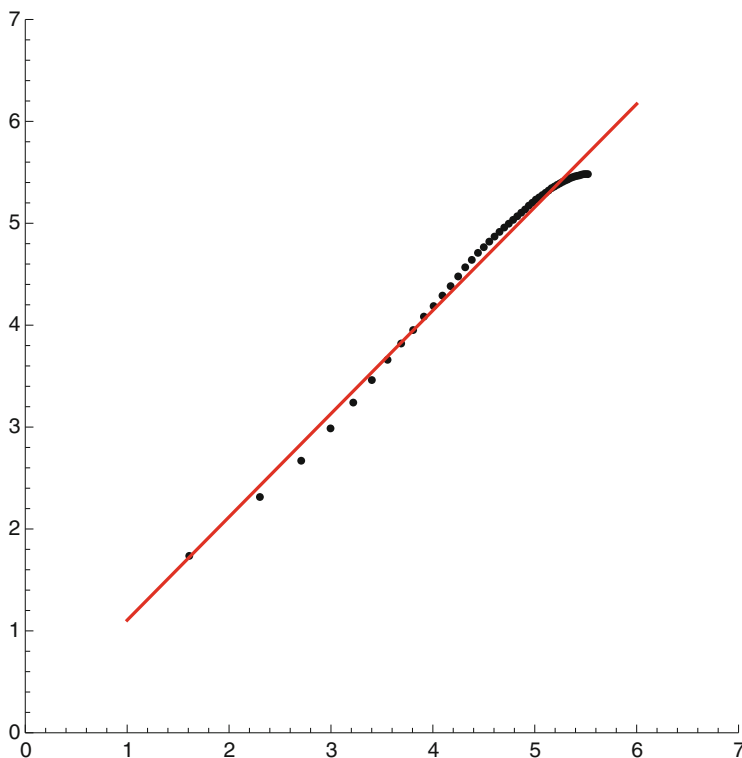


Fig. 4.9 The mean squared displacement for the sample trajectory

References

1. D. Campos, Méndez, V., Llopis, I.: Persistent random motion: uncovering cell migration dynamics. *J. Theor. Biol.* **267**, 526–534 (2010)
2. Dieterich, P., Klages, R., Preuss, R., Schwab, A.: Anomalous dynamics of cell migration. *Natl. Acad. Sci.* **105**, 459–463 (2008)
3. Fürth, R.: Die brownische bewegung bei berücksichtigung einer persistenz der bewegungsrichtung. mit anwendungen auf die bewegung lebender infusorien [in German]. *Z. Phys.* **2**, 244–256 (1920)
4. Gorelik, R., Gautreau, A.: Quantitative and unbiased analysis of directional persistence in cell migration. *Nat. Protocols* **9**, 1931–1943 (2014)
5. The Japan meteorological agency data service website: <https://www.jma.go.jp/jma/menu/menureport.html>. (in Japanese)
6. Masuzzo, P., Troys, M.V., Ampe, C., Martens, L.: Taking aim at moving targets in computational cell migration. *Trends Cell Biol.* **26**, 88–110 (2016)
7. Peruani, F., Morelli, L.G.: Self-propelled particles with fluctuating speed and direction of motion in two dimension. *Phys. Rev. Lett.* **99**, 010602 (2007)
8. Selmecki, D., Li, L., Pedersen, L., Nrelykke, S.F., Hagedorn, P.H., Mosler, S., Larsen, N.B., Cox, E.C., Flyvbjerg, H.: Cell motility as random motion: a review. *Eur. Phys. J. Topics* **157**, 1–15 (2008)
9. Wu, P.H., Giri, A., Wirtz, D.: Statistical analysis of cell migration in 3d using the anisotropic persistent random walk model. *Nat. Protocols* **10**, 517–525 (2015)