How to reliably estimate the tortuosity of an animal’s path: straightness, sinuosity, or fractal dimension?

Simon Benhamou

Behavioural Ecology Group, Centre d’Ecologie Fonctionnelle et Évolutive, CNRS, 1919 Route de Mende, F-34293, Montpellier Cedex 5, France

Received 14 November 2003; received in revised form 19 March 2004; accepted 26 March 2004

Abstract

The tortuosity of an animal’s path is a key parameter in orientation and searching behaviours. The tortuosity of an oriented path is inversely related to the efficiency of the orientation mechanism involved, the best mechanism being assumed to allow the animal to reach its goal along a straight line movement. The tortuosity of a random search path controls the local searching intensity, allowing the animal to adjust its search effort to the local profitability of the environment. This paper shows that (1) the efficiency of an oriented path can be reliably estimated by a straightness index computed as the ratio between the distance from the starting point to the goal and the path length travelled to reach the goal, but such a simple index, ranging between 0 and 1, cannot be applied to random search paths; (2) the tortuosity of a random search path, ranging between straight line movement and Brownian motion, can be reliably estimated by a sinuosity index which combines the mean cosine of changes of direction with the mean step length; and (3) in the current state of the art, the fractal analysis of animals’ paths, which may appear as an alternative and promising way to measure the tortuosity of a random search path as a fractal dimension ranging between 1 (straight line movement) and 2 (Brownian motion), is only liable to generate artifactual results. This paper also provides some help for distinguishing between oriented and random search paths, and depicts a general, comprehensive framework for analysing individual animals’ paths in a two-dimensional space.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Fractal; Diffusion; Orientation; Random Search

1. Introduction

Animals’ movements are involved in a large spectrum of fundamental behavioural and ecological processes, such as navigation, migration, dispersal, space use and food searching. Although the animal’s speed can be a key factor, e.g. in prey detection (Gendron and Staddon, 1983; Knoppien and Reddingius, 1985) or in space use (Benhamou and Bovet, 1989), numerous questions arise at the purely spatial level corresponding to the path structure. The analysis of such a structure most often reduces to the main question of how to reliably measure the “tortuosity” (i.e. the convoluted aspect) of the path (e.g. Clausen et al., 1997).

The tortuosity of oriented paths is obviously linked to the efficiency of the orientation mechanism involved: the more the animal is efficient, the more its path should be close to the straight line segment linking the starting point to the goal. To measure this closeness, Batschelet (1981) promoted the use of a simple and intuitively appealing straightness index ranging between 0 and 1, computed as the ratio between the initial beeline distance to the goal and the path length travelled to reach it. To my knowledge, however, no theoretical study yet attempted to determine the reliability of this index as a measure of the orientation efficiency.

The tortuosity of random search paths is a key parameter in space use, dispersal or food searching. The index of sinuosity proposed by Bovet and Benhamou (1988) has proved valuable to tackle various theoretical questions such as optimal search paths for a central place forager (Bovet and Benhamou, 1991), optimal food searching in a patchy environment (Benhamou, 1992) and, maybe more surprisingly, gradient field-based orientation (Benhamou, 1989; Benhamou and...
Bovet, 1989). When applied to the analysis of actual paths, however, it is restricted to a partial range of tortuosity and requires a regular path sampling (see below).

Since it was introduced by Dicke and Burrough (1988), the fractal analysis of animals’ paths appeared to be an alternative and promising method for measuring the tortuosity of a random search path as a fractal dimension ranging between 1 (curvilinear path) and 2 (fully jagged and wiggly path). The critique made by Turchin (1996) has thrown some doubt, however, on the reliability of the fractal dimension for measuring an animal’s path tortuosity.

This paper aims at providing a comprehensive framework for analysing the structure of animals’ paths through the answers to the following four questions: Is the straightness index really a reliable estimator of the animal’s orientation efficiency? Can this easily computed index be used as a reliable measure of the tortuosity of a random search path as well? How can the sinuosity index be generalized to deal reliably with any level of tortuosity of an irregularly sampled path ranging between the straight line movement and the Brownian motion? Is the fractal dimension a reliable estimator of the tortuosity of an animal’s random search path? In the present paper, an estimator is considered to be reliable if it fulfils two conditions: (1) it is unbiased, preventing the occurrence of any over- or under-estimation on the average, and (2) its coefficient of variation (standard deviation/mean) decreases with the sample size, making it possible to get a better accuracy when considering longer paths.

Except in the particular case of an animal that naturally performs discrete steps (e.g. a bee moving from one flower to another; see Discussion), an animal is assumed to move along a curvilinear path, which is discretized as a sequence of fixes at the recording level. It must be kept in mind that fixes are only sampling points: the lengths and orientations of the steps (i.e. movements between successive fixes) have no biological meaning by themselves. This discrete step representation of curvilinear paths proves very useful, however, because it provides a mathematically tractable way for both analysing and modelling them. Many variables involved in this paper being of angular nature, let us briefly recall a key notion of circular statistics (see Batschelet, 1981; Fisher, 1993 or Mardia, 1972) before entering the core of the subject. Any sample of \(n\) angular values \(\zeta\) can be characterized by a mean vector, whose Cartesian coordinates are the mean cosine \(c(\zeta) = \sum_{i=1}^{n} \cos(\zeta_i)/n = r(\zeta) \cos(\phi(\zeta))\) and the mean sine \(s(\zeta) = \sum_{i=1}^{n} \sin(\zeta_i)/n = r(\zeta) \sin(\phi(\zeta))\). The mean vector orientation \(\phi(\zeta) = \arctan[s(\zeta), c(\zeta)]\) expresses the angular mean, and the mean vector length \(r(\zeta) = [c(\zeta)^2 + s(\zeta)^2]^{1/2}\) expresses the concentration of the distribution around its mean between 0 (uniform, fully dispersed distribution) and 1 (punctual distribution, all angles being equal). Note: \(\arctan_2(y/x)\) is defined as \(\arctan(y/x)\) for \(x > 0\) and \(\arctan(y/x) + \pi\) radians for \(x < 0\).

### 2. Oriented paths

At any location, the ratio between the magnitude of the movement component in the goal direction and the magnitude of movement actually performed is equal to the cosine of the angular difference, hereafter referred to as the directional error, between the movement direction (step orientation) \(\theta\) and the goal direction \(\gamma\). Consequently, the best measure of the orientation efficiency, and thereby of the (inversely related) path tortuosity, is provided by the expected value of the cosine of the directional errors, \(E[\cos(\theta - \gamma)]\). It ranges between 0, for an animal which tends, on the average, to move orthogonally to the goal direction, and 1, for an animal which moves in straight line towards the goal (negative values can be considered when an animal aims at moving away from a dangerous location rather than at approaching a goal). Two cases can be distinguished: (1) when the goal is located at infinity, so that the animal will never reach it, but only aims at maintaining its course in the correct direction (e.g. a given compass bearing for a migrating animal); (2) when the goal is located at a finite distance from the starting point, and the animal will reach it after having moved some variable path length.

#### 2.1. Goal located at infinity (Never reached)

The situation occurring with a goal located at infinity or at a very large distance is quite simple because the goal direction \(\gamma\) (expressed with respect to the \(X\)-axis) is constant, whatever the animal’s current location \((X_n, Y_n)\). The orientation efficiency \(E[\cos(\theta - \gamma)]\) of a path made of \(n\) steps with lengths \(l_i\) \((i = 1 \text{ to } n)\) can be reliably estimated as the \(l\)-weighted mean cosine of directional errors, \(c_n(\theta - \gamma) = \sum_{i=1}^{n} l_i \cos(\theta_i - \gamma)/L\), where \(L = \sum_{i=1}^{n} l_i\) is the total path length travelled. The global movement component in the goal direction, referred to as the drift \(G\), corresponds to the distance between the starting point \((X_0, Y_0)\) and the projection of the animal’s location after \(n\) steps \((X_n, Y_n)\) on an axis running through \((X_0, Y_0)\) with the orientation \(\gamma: G = (X_n - X_0)\cos(\gamma) + (Y_n - Y_0)\sin(\gamma)\). It can be expressed as the sum of the local (step by step) movement components in the goal direction: \(G = \sum_{i=1}^{n} l_i \cos(\theta_i - \gamma) = Lc_n(\theta - \gamma)\). The drift is therefore proportional to the path length, and the ratio \(G/L = c_n(\theta - \gamma)\) constitutes an unbiased estimator of the orientation efficiency \((E(G/L) = E[\cos(\theta - \gamma)]\). The ratio \(G/L\) being equal to a weighted mean, its variability obviously decreases.
when the sample size (i.e., the step number) increases. In the simplest case of a path with a constant step length $p$ and statistically independent step orientations $\theta_i$, it can be easily demonstrated that the variance of the drift $V(G)$ is proportional to the path length: $V(G) = LpV[\cos(\theta - \gamma)]$, with $L = np$. There usually exists some correlation between the successive movement directions, however, so that the expression of $V(G)$ is much more complex, but the main point (confirmed using computer simulations) is that the variance remains proportional to the path length. The variance of the ratio $G/L$ is therefore inversely proportional to the path length. This ratio thus clearly constitutes a reliable estimator of the orientation efficiency of a path oriented towards an infinitely distant goal: this is an unbiased estimator whose accuracy increases when considering longer paths.

2.2. Goal located at a finite distance (Reached after n steps)

When the goal is located at a finite distance $D$ from the starting point, and the animal reaches it after having travelled some variable path length $L$, the orientation efficiency is often empirically measured as the ratio $D/L$, referred to as the straightness index (Batschelet, 1981). This is a popular estimate because it is very easy to compute and results in an intuitively appealing value ranging between 0 and 1. Batschelet (1981) stated that the ratio $D/L$ is practically equal to the mean vector length of the step orientations $r(\theta)$. He concluded that both can be used to estimate the orientation efficiency. In fact, with $D = [(\sum_{i=1}^{n} l_i \cos(\theta_i))^2 + (\sum_{i=1}^{n} l_i \sin(\theta_i))^2]^{0.5}$ and $L = \sum_{i=1}^{n} l_i$, the ratio $D/L$ is exactly equal to the “l-weighted mean vector length” of the $\theta$ distribution, and therefore to the basic (non-weighted) mean vector length $r(\theta)$ if the step length is constant. This does not prove by itself, however, that the straightness index reliably estimates the orientation efficiency $E[\cos(\theta - \gamma)]$.

The theoretical approach of the situation occurring with a goal located at a finite distance is rather complex because the goal direction $\gamma$ is no more constant but depends on the animal’s current location: $\gamma_i = \arctan2[(Y_n - Y_i), (X_n - X_i)]$, the goal location corresponding to the last point of the path $(X_n, Y_n)$. As the beeline distance from the starting point to the goal can also be expressed as $D = \sum_{i=1}^{n} l_i \cos(\theta_i - \gamma_0)$, the ratio $D/L$ seems only to be an approximation of the $l$-weighted mean cosine of directional errors $c_w(\theta - \gamma) = \sum_{i=1}^{n} l_i \cos(\theta_i - \gamma_{i-1})/L$, where the goal direction is taken to be constant and equal to the initial goal direction ($\gamma_0$) whatever the animal’s current location. Contrary to what occurs with a goal located at infinity, however, $c_w(\theta - \gamma)$ appears itself to be a biased estimator of the orientation efficiency $E[\cos(\theta - \gamma)]$ when the goal is located at a finite distance, because the goal direction $\gamma$ is computed only step by step, whereas it varies continuously during the animal’s movement. The longer the path the larger is the bias, and the unbiased estimation is therefore provided by the limit value to which the mean cosine of directional errors converges when the step length becomes infinitesimal. This limit value is equal to the ratio $D/L$, as illustrated by the simple following example involving a constant step length $p$. To reach a goal located 100 m eastward, an animal moves 50 m northward, turns 90° clockwise, moves 100 m eastward, turns again 90° clockwise and eventually moves 50 m southward ($D/L = 0.5$). The mean cosine of directional errors is equal to $0.65$ for $p = 50$ m ($n = 4$), $0.533$ for $p = 10$ m ($n = 20$), and $0.503$ for $p = 1$ m ($n = 200$). More formally, let us call $v$ the path curvilinear abscissa, ranging from 0 to $L$. An infinitesimal movement $dv$ results in an infinitesimal variation in beeline distance from the animal’s current location to the goal $dD = -\cos(\theta - \gamma)dv$: The total beeline distance from the starting point to the goal is then equal to $D = - \int_0^L \cos(\theta - \gamma) dv$. As $dv/L$ represents the relative frequency of occurrence of the various values of $\cos(\theta - \gamma)$ obtained along the path, the straightness index $D/L$ therefore gives the exact measure of the orientation efficiency of the path in question

$$D/L = \int_0^L \frac{\cos(\theta - \gamma)}{L} dv = E_{(\text{path})} [\cos(\theta - \gamma)].$$

The basic mean vector length of step orientations $r(\theta)$ hence usually results in an approximate value of the orientation efficiency (it provides the exact value only when the step length is constant). Obviously, the question of orientation efficiency concerns the orientation process at work rather than any given particular path. Because of the random noise occurring in the environment and/or in the orientation process, the actual path (and thereby its length $L$) travelled by an animal to reach a goal located at a given distance $D$ is just a sample of the parent population of paths potentially generated by the orientation process used. The particular value of $D/L$ obtained for a given path (Eq. (1)) is an unbiased estimator of the orientation efficiency ($E(D/L) = E[\cos(\theta - \gamma)]$). Furthermore, using computer simulations, it can be shown that the variance of this index is also inversely proportional to $D$ (or $L$). The straightness index $D/L$ thus appears to be a reliable (and easy to compute) estimator of the efficiency of the orientation process used by an animal that orients itself towards and eventually reaches a goal.

2.3. Correcting the straightness index

The straightness index $D/L$ is primarily a measure of the discrepancy between the path actually followed by the animal and a perfectly oriented straight segment
linking the goal to the starting point. The ratio \( G/L \) can be considered as a straightness index as well: it plays exactly the same role for a path oriented towards a very distant, never reached goal. A general formula which applies to any situation (including the case of an animal whose track was lost before it reaches a goal located at a finite distance) is given by the ratio \( \Delta D/L \), where \( \Delta D \) is the difference between the initial and final beeline distances to the goal, and \( L \) is the path length travelled between the initial and final locations recorded.

Although the straightness index has been proven to be a reliable estimate of the orientation efficiency, the physical structure of the environment is liable to weaken the value obtained at the path level. For example, the environment may be cluttered with many obstacles, forcing the animal to perform undesired detours before reaching the goal. An animal may also choose to perform some detours to go preferentially through some safe places or to avoid going through some other dangerous places. The actual efficiency of the orientation mechanism involved should be then higher than the measure obtained at the path level. In contrast, the path length travelled may be minimized at the recording level. Similar to what happens in fractal analysis when the ruler length increases (see below), the path length measured as the sum of the step lengths tends to decrease with the recording frequency, causing the straightness index to take too high values if the recording frequency is low. As far as possible, these various factors should be taken into account to obtain a better estimate of the orientation mechanism efficiency than the crude value of the straightness index.

3. Random search paths

Wandering animals usually travel with both a persistence and a balance propensity: they tend to move each new step in a direction (\( \theta_{i+1} \)) correlated to the previous one (\( \theta_i \)), and to turn to the right and to the left at random, with the same magnitude and frequency. A random search path can therefore be modelled as a correlated random walk. The distribution of turning angles (\( \phi(z) = 0 \)) is assumed to have a null mean and to be symmetrical about it. The mean sine is then null (\( s(z) = 0 \)), and the mean cosine is equal to the mean vector length (\( c(z) = r(z) \), hereafter referred simply to as \( c \)) and represents the correlation between the orientations of successive steps. In contrast, the turning angles themselves are assumed to be statistically independent (this constraint is addressed in the Discussion). The extreme cases with \( c = 0 \) (uniform distribution) and \( c = 1 \) (punctual distribution) correspond to Brownian motion (non-correlated random walk) and straight line movement, respectively.

3.1. Relationship between the diffusion distance and the path length

Insofar as the beeline distance \( D \) between the first and last points of a random search path, usually referred to as the diffusion distance or the net displacement, is an increasing function of the path length travelled and a decreasing function of the path tortuosity, it is tempting to use the simple straightness index \( G/L \) to derive a tortuosity measure (e.g. \( 1 - D/L \)). Unfortunately, such an index has no meaning for random search paths because it tends to 0, whatever the path tortuosity, when the path length \( L \) increases. As we saw before, the ratio \( D/L \) is approximately (variable step length) or exactly (constant step length) equal to the mean vector length \( r(\theta) \). The successive step orientations \( \theta \) of a random search path being correlated, there exists an initial bias linked to the first step orientation. Without orientation mechanism, this directional bias progressively vanishes: the step orientations tend to be uniformly distributed and therefore \( D/L \approx r(\theta) \) tends to 0 when the step number increases.

The expected value of the squared diffusion distance of an \( n \)-step random search path with a directional correlation \( c \) (\( c < 1 \)) and a randomly variable step length \( l \) is equal to (Hall, 1977; Kareiva and Shigesada, 1983; Marsh and Jones, 1988)

\[
E(D^2) = nE(l^2) + 2E(l)^2 \frac{c}{1-c} \left( n - \frac{1-c^2}{1-c} \right).
\]

Eq. (2) can be rewritten as

\[
E(D^2) = p^2 \left[ n \left( 1 + \frac{c}{1-c} + b^2 \right) - \frac{2c(1-c^2)}{(1-c)^2} \right],
\]

where \( p \) and \( b \) are the expectation and the coefficient of variation of the step length (\( E(l) = p \), \( E(l^2) = p^2(1+b^2) \)). This formula was previously provided by Tchen (1952) and Skellam (1973) in the particular case of a constant step length (\( b = 0 \)). When \( n \) is large, Eq. (3) reduces to

\[
E(D^2) = Lp \left( \frac{1+c}{1-c} + b^2 \right)
\]

with \( L = np \). It is worth noting that, when \( n \) is large but the mean sine of turning angles \( s(z) \) is liable to take any value (noted \( s \)), the general but intricate expression provided by Kareiva and Shigesada (1983) can be reduced to a form just a little more complex than Eq. (4)

\[
E(D^2) = Lp \left( \frac{1-c^2-s^2}{(1-c)^2+s^2} + b^2 \right).
\]

Bovet and Benhamou (1988) demonstrated that, when \( n \) is large, the expected diffusion distance is equal to

\[
E(D) = \left( \frac{n}{4} E(D^2) \right)^{0.5}.
\]
According to Wu et al. (2000) and Byers (2001), Eq. (6) generates poor estimations of \( E(D) \) when \( n \) is small and \( c \) is close to 1. This holds true only if Eq. (6) is based on the approximate (Eq. (4)) rather than exact (Eq. (3)) value of \( E(D^2) \) (Benhamou, 2004). Hence, when \( n \) is large enough (depending on \( c \)), \( E(D) \) increases proportionally to \( \sqrt{L} \). This confirms that the ratio \( D/\sqrt{L} \) tends to decrease when \( L \) increases (and reaches 0 when the path is infinitely long). In contrast, the ratio \( D/\sqrt{L} \) constitutes an (inversely related) unbiased estimator of the tortuosity of a random search path. It cannot be considered as a reliable estimator, however, because of its intrinsic variability. Indeed, the variance \( V(D) = (4/\pi - 1)E(D^2) \) involves a rather high coefficient of variation (0.52), whatever the path length travelled. Nevertheless, the form of this ratio shows that, contrary to the straightness index of an oriented path, which is dimensionless, the tortuosity of a random search path has a dimension which is clearly related to the unit length.

### 3.2. Definition of the sinuosity

Bovet and Benhamou (1988) modelled random search paths as correlated random walks with a constant step length \( p \) (\( b = 0 \)). We assumed that the turning angles were drawn, independently from each other, in a normal distribution wrapped on the circle with a null mean and a standard deviation \( \sigma \) (\(< 1.2 \text{ rad}\)), and looked at the turning angle distribution obtained after rediscretization with a constant step length \( q \) (this consists in replacing the original sequence of data points by a new sequence such that the distance between any two successive new points is equal to \( q \); see appendix in Bovet and Benhamou, 1988). Using computer simulations, we showed that the resulting turning angles are normally distributed with a null mean and a standard deviation \( \sigma_q = 0.85\sigma q/p^{0.5} \) (for \( \sigma_q < 1.2 \text{ rad} \)). Segregating the parameters initially introduced in the model (\( p \) and \( \sigma \)) from the variables resulting from the rediscretization (\( q \) and \( \sigma_q \)), we defined the sinuosity \( S \) of a random search path as:

\[
S = \sigma/\sqrt{p} = 1.18\sigma_q/\sqrt{q}.
\]  

(7)

The first part of Eq. (7) can be used to simulate random search paths with a given sinuosity by choosing appropriate values of \( p \) and \( \sigma \), and the second part to compute the sinuosity of an actual path after rediscretization with a step length \( q \). However, Eq. (7) presents two shortcomings. The first comes from the fact that it rests on linear statistics. Drawing the turning angles from a 0-centred wrapped normal distribution makes it possible to generate any path ranging from a straight line (\( \sigma = 0 \)) to Brownian motion (\( \sigma > 4 \text{ rad} \)). At the practical level of path analysis, however, the standard deviation is computed on angular values which have been “linearized” by expressing them strictly between \(-\pi\) and \(\pi\) rad. The standard deviation of a linear uniform distribution between \(-\pi\) and \(\pi\) is equal to \(\pi/\sqrt{3} = 1.81\). Hence, the standard deviation of the linearized distribution tends to 1.81 rad when the initial standard deviation \( \sigma \) tends to infinity (\( \sigma > 4 \text{ rad in practice} \)). The two values diverge markedly for \( \sigma > 1.2 \text{ rad} \). Eq. (7) can therefore be applied only to paths characterized by a lower value of \( \sigma \) or \( \sigma_q \).

The second shortcoming comes from that Eq. (7) is based on the rediscretization of constant step correlated random walks whereas actual paths, after recording, occurred as discrete sequences of fixes, with a randomly variable distance between successive fixes due to changes in the animal’s speed and possibly in the recording frequency.

How to define a general sinuosity index able to deal with any level of turning angle dispersion and with a randomly variable step length? Let us recall that step lengths and turning angles are useful tools for representing a path but have no meaning by themselves: this is the way they are combined which makes it possible to define a biologically meaningful sinuosity index. It appears retrospectively that the particular combination occurring in Eq. (7) comes from the relation taking place between the expected squared diffusion distance and the path length. Indeed, with a constant step length \( p \) (\( b = 0 \)), Eq. (4) reduces to \( E(D^2) = Lp(1+c)/(1-c) \). Turning angles being drawn from a zero-centred wrapped normal distribution with a standard deviation \( \sigma \), the value of the mean cosine is \( c = \exp(-\sigma^2/2) \) (Mardia, 1972). In the restricted range, where the linear definition of the sinuosity is valid (\( \sigma < 1.2 \text{ rad, i.e. } c > 0.5 \)), the ratio \( (1+c)/(1-c) \) is approximately equal to \( 4/\sigma^2 \). The sinuosity index can therefore be written as:

\[
S = 2[E(D^2)/L]^{-0.5}.
\]  

(8)

Other formulations are possible. For example, a “random search straightness index” may be defined as \( R = 4/S^2 = p[(1+c)/(1-c) + b^2] \), so that one gets \( R = E(D^2)/L \) (when \( n \) is large).

To verify whether \( S \) constitutes a suitable combination of the three parameters involved (\( c, p, \) and \( b \)), I looked at the evolution of the mean cosine of turning angles obtained after rediscretization, \( c_q \), as a function of the rediscretization step length \( q \). For this purpose, I simulated two very long (\( 3 \times 10^6 \) steps, to allow statistical convergence) correlated random walks, one with a constant simulation step length \( p \) (\( b = 0 \)), and the other with a randomly variable simulation step length \( l \) with mean \( p \) (\( l = p z^2 \), where \( z^2 \) is a random variable following a \( \chi^2 \) law with one degree of freedom: \( E(l) = p, b^2 = 2 \)), for each of the following 13 values of \( c \): 0, 0.1, 0.2, …, 0.8, 0.9, 0.95, 0.99, and 1. Turning angles were
drawn, independently from each other, from a wrapped normal distribution with a null mean and standard deviation \( \sigma = \sqrt{-2 \ln(\epsilon)} \). These paths were rediscretized with 33 values of \( q \) \((0.6p, 0.7p, \ldots, 1.9p, 2p, 3p, \ldots, 20p) \), and the mean cosine \( c_q \) of the resultant turning angle distributions was computed. The results (Fig. 1) are suitably approximated by the function \( c_q = 1 - 2/\pi \arctan(qS^2/2) \), with \( S \) as defined by Eq. (8). Whatever the arbitrary rediscretization step length \( q \) used, the sinuosity can be computed as

\[
S = \left[ \frac{2}{q} \tan \left( \frac{\pi}{2} (1 - c_q) \right) \right]^{0.5}.
\]

Eq. (8) constitutes an appropriate combination of \( c, p \), and \( b \). Additional computer simulations involving four values of \( b \) \((0, 1, \sqrt{2}, \text{and } 2) \) showed that the coefficient of variation of \( S \) is approximately equal to \( (1 + b^2/14)/(2n)^{0.5} \), whatever the current values of \( p \) and \( e \) \((e < 1) \). \( S \) thus constitutes a reliable sinuosity index. It is inversely related to the ratio \( E(D)/\sqrt{L} \) expected for a large step number. Being based on the whole path structure, its accuracy increases with the path length or the sampling frequency, contrary to the simple \( D/\sqrt{L} \) ratio, which is based only on the relative locations of the first and last points of the path.

Sometimes, turning angles may be noticeably not balanced. For example, thrushes searching for earthworms in high-density patches tend to turn systematically to the same side (Smith, 1974a, b), involving a mean sine of turning angles \( s \) markedly different from 0. An even more general sinuosity index can be defined theoretically based on Eq. (5) as

\[
S = 2 \left[ p \left( \frac{1 - c^2 - s^2}{(1 - c^2) + s^2 + b^2} \right) \right]^{-0.5}.
\]

Eqs. (8) and (10) can be used to simulate correlated random walks with any given sinuosity by choosing appropriate values of \( c, p, \) and \( s \) \((b \text{ may be set to } 0; \text{ Fig. 2a}) \). In the case where turning angles are drawn from a zero-centred wrapped normal distribution with \( c > 0.5 \) \((\sigma < 1.2 \text{ rad}) \), the simple linear expression of the sinuosity (first part of Eq. (7)) can be applied as well.

The practical computation of the sinuosity of an actual random search path, based on Eqs. (8), (9) or (10) is detailed in the Discussion.

### 3.3 Fractal analysis

The fractal dimension \( F \) of an object measures its ability to fill the Euclidean space in which it is embedded (Mandelbrot, 1982). In particular, one gets \( F = 1 \) for a straight line, and \( F = 2 \) for Brownian motion, which covers the whole plane when the step number increases indefinitely. Between these two extremes, random search paths might be characterized by intermediates \( F \) values, reflecting various values of tortuosity. The dimension of a fractal line can be measured by the divider method, which consists in looking at how the line length varies with the ruler length (formally equivalent to a rediscretization step length) used to measure it (see Sugihara and May, 1990). The logarithm of a fractal line length is a linear function of the logarithm of the ruler length, with a slope equal to \( 1 - F \).

I simulated a correlated random walk made of \( 3 \times 10^6 \) steps for each of the following six values of mean cosine of turning angles \( c \): 0 (Brownian motion), 0.4, 0.6, 0.8, 0.9 and 1 (straight line), with a constant step length serving as unit length. The path lengths \( L \) were measured with 109 ruler lengths: 1.1, 1.2, ..., 1.9, 2, 2.5, 3, 4, ..., 99, and 100. The results (Fig. 3) confirm those got by Turchin (1996): animal’s random search paths cannot be considered as fractal (i.e., indefinitely
The changes in search mode appear as smooth transitions involving a step number equal to the windows width.

The tortuosity of an animal’s path can be reliably measured either as a straightness index or a sinuosity index, depending on whether the animal performed an oriented movement or a random search movement. Before using the appropriate index, straightness or sinuosity, it is therefore required to determine which type of movement the animal performed. In homing experiments, animals are assumed to orient, more or less efficiently, towards their home (e.g. Papi, 1992). In some other experiments, they are assumed to wander freely in the environment, searching at random for items which cannot be detected at distance (e.g. Kareiva and Shigesada, 1983). In many other experiments, however, this is not so clear. An animal reaching a given target may have performed a random search path and encountered the target by chance, or it may have actively oriented itself towards the target. A very straight path suggests an oriented movement, but a slightly more tortuous path may result from a less efficient orientation mechanism as well as from a low sinuosity random search.

A complex statistical method based on angular distribution analyses was developed by Benhamou and Bovet (1992) to determine whether or not an orientation mechanism is involved, and the type (scalar of vectorial) of orientation information used. It requires a large amount of precise data, however, to be effective (e.g., Dejean and Benhamou, 1993). A simpler method consists in computing the difference between the square of the mean vector length of step orientations, \( r^2(\theta) \), and the square of the mean vector length of turning angles, \( r^2(\alpha) \approx c^2 \). This quantity is, on the average, positive for a movement oriented in a given constant direction (goal located at infinity) and negative for a random search movement (Marsh and Jones, 1988). The variances involved are however so high that distinguishing
between the two types of movement is not statistically possible when the step number \( n \) is low. In addition, the case of a goal reached after \( n \) steps is not considered.

An alternative method rests on the way the beeline distance \( D \) (or the drift \( G \) for a goal located at infinity) is related to the path length \( L \). Indeed, on the average, \( L \) is proportional to \( D \) for an oriented path, whereas for a random search path, \( D \) is proportional to \( \sqrt{L} \) (when \( n \) is large or \( c \) is not too close to 1; otherwise Eq. (3) cannot be suitably approximated by Eq. (4), and \( D \) is then almost proportional to \( L \)). Using simply the final values of \( D \) and \( L \) is rather ineffective because of the very high variability involved (Marsh and Jones, 1988). Looking at the whole relation between \( D \) and \( L \) in a backward procedure (starting at the penultimate point of the path and moving backward along the path down to the first point) is more informative. Considering an animal that reaches a target after \( n \) steps, this procedure consists in:

1. computing the current backward beeline distance \( D_i = [(X_n - X_{n-i})^2 + (Y_n - Y_{n-i})^2]^{0.5} \)
2. plotting \( D_i \) as a function of \( L_i \) and as a function of \( \sqrt{L_i} \) (and 3) for each plot, computing the sum of the linear regression residual squares with a null intercept (\( D_0 = L_0 = 0 \), whatever the type of path). The type of path involved is likely to correspond to the minimum sum. Additional computer simulations involving random search movements showed however that the probability of obtaining a relationship closer to the linear line than the square root curve, corresponding to the probability of erroneously rejecting the null hypothesis, is rather high at the individual path level: \( e \approx 1 - 1.45 \arctan((1-c)(n-1)/(3c))/\pi \) for a constant step length (a randomly variable step length leads to a slightly higher risk). The minimum risk, obtained with large values of \( n \), is therefore \( e = 0.275 \). Hence, a reliable statistical decision can be taken only at the population level, considering several paths reaching a target (with \( N \) paths characterized by similar values of \( n \) and \( c \), the statistical test rests on a binomial law \( B(N, c) \)).

4.2. Practical computation of the sinuosity of actual random search paths

One of the basic assumptions of the correlated random walk model, on which rests the definition of the sinuosity index \( S \), is that turning angles are statistically independent. Depending on the recording frequency used (from 25 fixes per second in videotracking to a few fixes per day in Argos satellite tracking), the sequence of fixes obtained constitutes a more or less accurate representation of the original path. Recording a path with a high frequency results in a “quasi-continuous” path representation preserving a maximum of details (provided the accuracy of the fixes is good enough; otherwise the data need filtering before analysis). Contrary to what occurs for a path recorded with a low frequency, successive turning angles measured at the fix level are then positively correlated, because of motion constraints: the long-term average of turning angles remains null, but the probability of occurrence of a (anti)clockwise turning angle is greater when the previous turning angle was (anti)clockwise. To get rid of this auto-correlation, it is necessary to
resample the path at a larger scale, either by filtering out a fraction of the fixes (subsampling), or by rediscretizing the path with a constant step length. As the accuracy of the sinuosity measure increases with the step number, it is preferable to use the minimum resampling required to respect the non-correlation constraint. The recorded path should therefore be resampled several times with an increasing (filtration or rediscretization) step length, starting with a length close to the average distance between fixes, until the turning auto-correlation becomes statistically non-significant.

Insofar as the measure of the sinuosity of an actual random search path depends on the way the path was discretized at the recording level, and possibly rediscretized afterwards, it is important to decipher the effects due to these processes. Three effects appear to be involved: “scattering”, “smoothing”, and “grazing”. The first two effects correspond to an increasing dispersion of the turning angles, and to an increasing loss of details, respectively, when the step length increases. As long as the path does not degenerate into Brownian motion (involving a complete loss of essential information due to a too low recording frequency and/or a too long rediscretization step), the way the mean cosine of turning angles varies with the step length is predictable, making it possible to compute the sinuosity index. In addition, the rediscretization process tends to plane the sharpest asperities of the path, resulting in a turning angle distribution more concentrated than expected. This grazing effect is best exemplified when a constant step length non-correlated random walk is rediscretized with a rediscretization step length close to the simulation step length ($q/p \approx 1$ in Fig. 1a): the mean cosine of turning angles increases from $c = 0$ to $c_p = 0.28$. Similarly, the 1.18 factor in Eq. (7) counterbalances this effect. As its magnitude appears to be inversely related to $c$, this effect should not occur with the rediscretization of the quasi-continuous tracks obtained with a high recording frequency, which are usually characterized by small and positively correlated turning angles. Similarly, the initial discretization (at the recording level) of an actual curvilinear path should not involve any grazing effect.

The practical computation of the sinuosity of an actual random search path hence depends on whether a grazing effect is involved or not. The second case occurs when the path was recorded either with a low frequency, so that the turning angles measured at the fix level were statistically independent, or with a high frequency (quasi-continuous track) and was then resampled to get statistically independent turning angles. In this case, the sinuosity value can be computed based on Eq. (8) or (10), depending on whether the mean of turning angles does not or does differ significantly from 0, respectively (with $p$ set to $q$ and $b$ set to 0 for quasi-continuous tracks rediscretized with a constant step length $q$). The first case occurs when the path was recorded with an intermediate frequency, and was then rediscretized with a constant step longer than the mean distance between fixes to get statistically independent turning angles. In this case, the sinuosity value should be computed based on Eq. (9) which counterbalances the grazing effect, provided the mean of turning angles does not differ significantly from 0. Otherwise, the sinuosity cannot be computed because the way the mean cosine $c$ and the mean sine $s$ are altered by the rediscretization is unknown (and beyond the scope of this paper). In any case, the sinuosity cannot be computed if there is a significant cross-correlation between the magnitude of turning angles and the step lengths. The path rediscretization with a constant step length is a useful mean to get rid of such a correlation. Note that, to simply compare different random search paths which were rediscretized with the same constant step length $q$, a simplified sinuosity index may be computed as $S_c = 1 - c$ (provided the mean of turning angles remains close to 0).

A worth noting point is that the sinuosity index $S$, contrary to straightness index $D/L$, is not dimensionless but related to the unit length. Random search paths are often used to find prey items whose density is also related to a unit length. Expressing the sinuosity with the same unit length as the one used to express the prey item density ensures the results to be consistent (Benhamou 1992, 1994). The choice of the unit length may be critical when comparing sinuosities of paths made by animals of very different sizes, e.g. an ant and an albatross (if such a comparison proved relevant). The two movement patterns may look similar, because of close mean cosine values $c$, whereas the actual mean step lengths $p$ are very different, say, $1$ cm for the ant and $1$ hm for the albatross. If the sinuosity is expressed with respect to a common metric unit, e.g. $1$ m, the values obtained will be extremely different. However, the biological meaning of a $1$-m distance for the two species is also very different in terms of both prey density and perceptual range detection. Hence, for interspecific comparative purpose, the sinuosity values have to be “rescaled” to biologically relevant unit lengths (e.g. in rad/cm$^{0.5}$ for the ant and rad/hm$^{0.5}$ for the albatross).

4.3. Analysing composite paths

Paths recorded for a long time are likely to encompass both oriented and random search sections. A global analysis, in terms of correlated random walk (Bergman et al., 2000) or fractal dimension (Fritz et al., 2003), is therefore liable to provide dubious results. The correct procedure requires first to identify the various parts corresponding to oriented stages and to random search stages (Dejean and Benhamou, 1993; Girard et al., 2004), and possibly to the simultaneous use of the two
kinds of process (see below). In addition, during a random search stage, an animal may alternate between extensive and intensive search modes, involving a separate computation of the sinuosity for each mode. The best situation occurs when one can a priori predict where the animal should perform intensive searching and extensive searching (and then test the hypothesis that the animal used two modes), based on additional information about the habitat structure or behavioural events such as prey captures (Smith 1974a, b; Mellgren and Roper, 1996). When such an information is missing, the areas concerned may be deduced from the movement data only (involving a less powerful analysis). For this purpose, one can look at the mean cosine and mean sine of the turning angle distribution measured on a sliding window. The width of the window should be large enough to ensure statistical stability (the narrower is the window, the noisier are the results). On the average, the mean cosine is lower and/or the mean sine is closer to 0 when the animal performs extensive searching than when it performs intensive searching (see Fig. 2b). If the shift from intensive to extensive searching is progressive, the observed movement pattern has to be analysed in the framework of a modified model where the sinuosity decreases with the path length (Morales and Ellner, 2002), because the mere correlated random walk model is unable to account for the data.

Some animals perform naturally discrete steps starting and ending at specific places, e.g. a bee moving from one flower to another. This results in a composite path which usually requires a two-level analysis (Benhamou, 1990). The first, global level of analysis focuses on the general movement pattern from place to place, the moves between two successively visited places being schematically represented as straight segments. The global movement pattern may correspond to an oriented path, the places visited serving as beacons (piloting behaviour). It can then be quantified using a straightness index \( \frac{D_g}{L_g} \), where \( D_g \) is the beeline distance between the first and last places visited, and \( L_g \) is the sum of the move lengths between places. If the global movement pattern involves a random search path, Eq. (8) or (10) should result in a reliable sinuosity value, provided there are no significant auto-correlations between successive turning angles and cross-correlation with move lengths. The second, local level of analysis focuses on the oriented moves performed between successively visited places: the orientation efficiency of each move can be quantified as a local straightness index \( \frac{D_i}{L_i} \), and the ratio \( \sum D_i / \sum L_i \) provides a mean local efficiency value for the whole path. Many animals also perform “saltatory searching” (O’Brien et al., 1990), which consists in naturally discretized movements at the temporal level. The stop phases of these stop-and-go movements do not occur at obvious places (such as flowers for a bee), so that any go phase may correspond to a random search for favourable stopping places as well as to an oriented move in a direction liable to change at each stop. If the animal tends to perform the successive moves in a given preferred direction, the global pattern is directionally biased and can no more be accounted for by a simple correlated random walk model (Samu et al., 2003).

An animal may search for prey at random, and at the same time, tends globally to move in a given direction. This results in second type of composite path where the two types of processes are used simultaneously rather than alternately. In such biased random walks, the diffusion distance tends to increase faster than predicted by the correlated random walk model (Firle et al., 1998; Mårell et al., 2002; Samu et al., 2003). A useful means to model a biased random walk involves differential klinokinesis (Benhamou and Bovet, 1989), i.e. a random search path whose local sinuosity \( S_i \) depends on the previous movement orientation \( \theta_{i-1} \) with respect to a given preferred direction \( \gamma \), e.g. \( S_i = S_0 [1 - k \cos(\theta_{i-1} - \gamma)] \), where \( S_0 \) is the basic sinuosity, which controls the overall searching intensity, and \( k \) (\( 0 < k < 1 \)) is the klinokinetic factor, which controls the variations in sinuosity, and thereby the movement bias (drift \( G \)) in the preferred direction (\( k \approx G/L \)). The analysis of this kind of movement pattern then involves to specify the values of both \( S_0 \) and \( k \) (Benhamou and Bovet, 1992). Alternatively, the diffusion distance may increase slower than predicted (e.g. Fortin, 2003). This occurs in particular when an animal tends to perform search loops, e.g. a mammal which simultaneously searches for prey items and restricts its movement within the given area corresponding to its home range, or an insect searching for its nest hole after a homing path. Search loops involve simultaneously a centripetal orientation component and a random search component. Analysis of such patterns obviously require specific models (Benhamou, 1989, 1994), because neither an analysis in pure terms of orientation efficiency nor an analysis in pure terms of sinuosity can provide reliable results: as the beeline distance \( D \) is repeatedly reset to a value close to 0, the straightness index \( D/L \) has no meaning, and the global movement pattern clearly does not fit a simple correlated random walk model.

4.4. Weakness of the fractal analysis of animals’ paths

Rediscretizing or measuring a path with an increasing step or ruler length involves some loss of information, because the longer is the step or ruler, the larger is the number of smaller scale details ignored. The main information about the structure of the path appears to be roughly conserved, however, because of some intrinsic self-similarity, suggesting that the tortuosity of random search paths may be characterized in the form of a fractal dimension (Dicke and Burrough,
There is in fact no self-similarity: looking at a random search path at different scales does not result in the same pattern because the directional correlation (i.e., the mean cosine of turning angles) varies with the scale. In some way, the variation of the former counter-balances the variation of the latter, as long as the path has not degenerated into Brownian Motion (the distribution can then no more flatten to counterbalance the increase in step length). When recorded with a high frequency, actual animals’ paths look like curvilinear lines, which are theoretically differentiable at any location. It is therefore hard to think about them as fractal lines (i.e., continuous but non-differentiable lines; see Sugihara and May, 1990). Turchin (1996) and the present paper provide clear evidence that animals’ random search paths, at least when they are modelled as correlated random walks, are not fractal. Random walks are able to cover the whole plane, whatever the correlation value $c<1$. This means that the true fractal dimension of these walks is equal to 2, whatever the apparent $F$ value computed from the local slope of the log(path length)-log(ruler length) relation.

Halley et al. (2004) warned against measuring fractal dimensions of non-fractal objects which may seem to be fractal at some spatial scales (apparent fractality). Beyond the logical argument that measuring the fractal dimension of a non-fractal object has no meaning and is therefore only liable to generate artifactual results, a strong mathematical argument against the use of the apparent fractal dimension $F$ (as computed from the local slope of the log–log relation) to measure the path tortuosity was provided paradoxally by Nams (1996) in a paper advocating the opposite point of view. Indeed, he showed that $F$ is approximately equal to $2/(1 + \log_2(1 + c))$. In other terms, $F$ is no more than a monotonously decreasing function of the mean cosine of turning angles $c$ (with $2 > F > 1$ for $0 < c < 1$). The decrease of the local slope (equal to $1 - F$, from 0 to $-1$; see Fig. 3) is eventually the simple reflect of the decrease (from 1 to 0) of the mean cosine of turning angles (see Fig. 1). As we have seen, there are two key parameters in random search path analysis: the mean cosine of turning angles and the mean step length. In a seminal paper, Siniff and Jessen (1969) considered the distributions of turning angles and of step lengths, but only separately. In most cases, turning angles and step lengths are just useful tools for representing a curvilinear path, but have no biological meaning by themselves: both distributions depend on the arbitrary (re)discretization used at the level of movement recording and analysis. The sinuosity index $S$, since its introduction by Bovet and Benhamou (1988) to the generalized definition proposed in the present paper, combines the two key parameters explicitly to obtain a measure of tortuosity which is independent of the arbitrary (re)discretization used. This is why the sinuosity $S$ is a reliable measure of the path tortuosity of a random search path whereas the apparent fractal dimension $F$, which is just an analogue of the mean cosine of turning angles and fully ignores the mean step length, is not.

The related concepts of chaos, strange attractor and fractal have been largely used in biology (Briggs and Peat, 1989; Hastings et al., 1993; Sugihara and May, 1990; Halley et al., 2004). Applying the fractal approach to the study of random search paths might be fruitful, but first requires building up a fractal movement model as an alternative to the correlated random walk model. A fractal movement model was proposed by Dicke and Burrough (1988; see also Sugihara and May, 1990), based on a fractal Brownian function. To my knowledge, however, it has never been proved yet that this type of model may constitute a more suitable representation of actual animals’ random search paths than the correlated random walk model. The way to further theoretical research on this topic is wide open.

Acknowledgements

I thank Pierre Bovet, Laurent Dagorn, Charlotte Girard and an anonymous referee for stimulating discussions and/or helpful comments and suggestions. The Monte-Carlo simulations and rediscretization programs run thanks to FreePascal compiler (freely available at www.freepascal.org).

References


