Università degli Studi di Padova
DIPARTIMENTO DI FISICA E ASTRONOMIA
Corso di Studi in Fisica

Phenomenological Modeling of the Motility of Self-Propelled Microorganisms

Tesi di Laurea Magistrale
Anno Accademico 2013 – 2014

Laureanda
Silvia Zaoli

Relatore
Prof. Amos Maritan

Correlatori
Marco Formentin
Andrea Giometto

Contents

Introduction

If we could look at what happens in a puddle with a microscope we would probably see a myriad of different microorganisms moving around. Bacteria and protists of different species, which size spans from 1μm to 100μm and which move at speeds of less than 1mm per second. Most of these living creatures do not just diffuse passively, but actively swim around looking for nutrients or better living conditions. Their movement has many interesting features and can be studied from various points of view, each focusing on different aspects and seeking a different level of description.

The advancement of imaging techniques allowed in the last decades a better comprehension of the molecular and biochemical basis of motion peculiar to each organism (bacteria, ciliates, amoeba,...), leading to a branch of study focusing on the physiology of the propelling means, as for example the rotation of flagella by motor proteins in bacteria or the structure of the cilia of eukaryotic cells [1]. An alternative approach pursues the understanding of the hydrodynamics of the motion, whose main feature is that viscous forces dominate over inertia, requiring swimming techniques different from those seen in larger organisms [2].
Finally, at a further level of abstraction, one may completely neglect the above aspects and analyse the statistical properties of motion, treating the organisms as point-like and non-interacting objects [3]. This approach, which may appear like an overly simplified description of a complex phenomenon, could be an appropriate starting point to shed light on the key features of microbial behavior. Once a common ground for the description is established, the desired level of specificity can be added *a posteriori*. A legitimate question is, although, whether this general description overlooking specific details is possible or whether the differences in the mechanisms of motion of the various creatures influence deeply the statistical properties of the motion itself. Microbial motility is an open problem in the study of active matter, and a complete understanding has not been reached yet.

The point of view chosen in this thesis is this last, statistical one, and in particular the focus is on active, or self-propelled, motion in absence of external signals (in contrast with chemotactic motion, i.e. motion influenced by the presence of a gradient of certain substances, like nutrients, which orient the motion of the organisms). Thus, in the original model of motion proposed in this thesis [4], which will be presented in chapter 3, we do not take in consideration whether the organisms swim with cilia, flagella, or how do they use their appendages to propel themselves.
All the same, since it is instructive and fascinating to explore these issues, the first section of chapter 1 reviews (a very little part of) what is known about the swimming of microorganisms. In this first chapter we will examine the motion of organisms which are very different: from bacteria to motile epithelial cells. They differ in dimension and typical speed, belong to different domains (eukaryotes, prokaryotes), some are multicellular and some unicellular. Also the reason of their motion is different: some, as bacteria, protists, and small animals,

move to look for nutrients, while the motility of tissue cells seems to be connected to immune response and tumour spreading [5]. The goal is to investigate the traits of the motion that are common among them, to understand to what extent such common traits exist and finally to have a look at how the issue of modelling this motion has been dealt with. Section 1.2 presents the statistical features of the motion which have been observed experimentally in the past, the most peculiar of which being non-Gaussianity of $P(v)$, while section 1.3 contains a review of previous models which tried to explain these experimental results.
The original part of the thesis is exposed in chapter 3 and consists in the proposal of a new microscopic model for microbial motility, prompted by the observation of experimental trajectories of the ciliate *Colpidium* sp. (in this sense the model is "phenomenological"). The novelty of the proposed model is that of using a non-white and non-Gaussian process to model the propelling force exerted by the microorganisms, in contrast with all previous models. This hypothesis allows to find a non-Gaussian $P(v)$ which is in good agreement with the experimental one.
Chapter 2 exhibits the analytical solution a one-dimensional stochastic process subject to a Markovian dichotomous noise, which will serve as an analytical point of reference for the model in chapter 3, in which a 2D generalization of this process is employed to describe the motion of the microorganisms in a two-dimensional landscape.

Chapter 1

Active motility of microorganisms

This chapter contains a review on the motility of microorganisms. The subject is vast, so a selection had to be made: there is no claim of completeness. In section 1.1 mechanisms of motion of microorganisms are explained, with particular emphasis on bacteria and protists. In section 1.2 we look at statistical features of motion, exploring past experimental findings. Finally, section 1.3 deals with the modelling of the motion, comparing some of the models which have been proposed to explain the experimental results.

1.1 Mechanisms of motion

When we swim in water we use our arms and legs to push the water backwards, and consequently get an acceleration forward. Thanks to that thrust, we are able to coast for some meters before the drag stops us. For a microorganism some micrometers large things are different. At that size, and at a speed of less than a millimeter per second, viscous forces in water are dominating over inertial forces, so the latter cannot be exploited for propulsion. We can see this computing the Reynolds number, which measures the ratio between inertial and viscous forces. Taking as a typical speed $v \sim 10^{-4}\mathrm{m/s}$[1], as a typical length $L \sim 10\mu\mathrm{m} = 10^{-5}\mathrm{m}$ [2], and using the density d and viscosity η of water, we get

$$Re = \frac{vLd}{\eta} = \frac{10^{-4} \cdot 10^{-5} \cdot 10^{3}}{10^{-3}} = 10^{-3} \ll 1 \tag{1.1}$$

For bacteria, which are typically one order of magnitude smaller, Re lowers to 10^{-5}. (For comparison, a fish has Re $\approx 10^{5}$.) Both motions, of protists and of bacteria, may be classified as *motion at low Reynolds number* with all the resulting consequences. For instance, it is interesting to compute how far a bacterium will coast if it stops swimming [6]. At low Re we are in the laminar regime so the viscous force is linear with v. Then we may approximate it using Stokes Law, which would be valid for a smooth rigid sphere, obtaining an equation of motion

$$-m\frac{\mathrm{d}v}{\mathrm{d}t} = 6\pi\eta a v \tag{1.2}$$

for a bacterium of radius a which is not exerting any propelling force. The speed will then decay as

$$v(t) = v_0 e^{-\frac{t}{\tau}} \tag{1.3}$$

[1] This speed is typical for protists. Bacteria, like *E. coli*, move at $v \sim 30$ μ m/s [6], while other microorganisms, like amoeba, are much slower ($v \sim 10$ μm/min [7])

[2] This is the order of magnitude of the size of an eukaryote, as protists. Bacteria are usually an order of magnitude smaller.

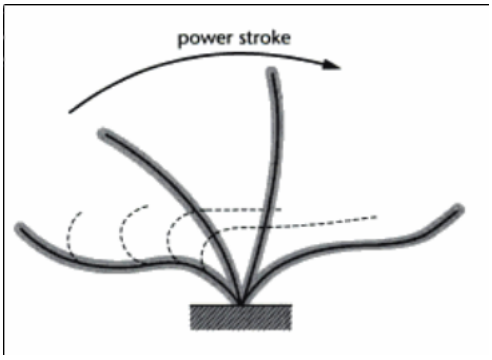


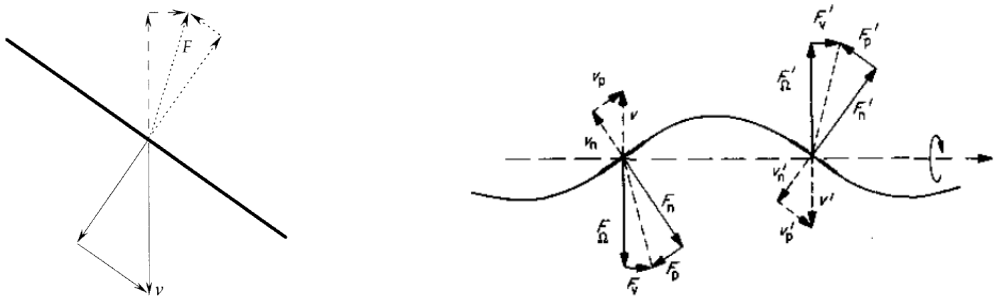
Figure 1.1: The beating of a cilium: each cilium performs a cycle of movement consisting in a fast power stroke and a slow recovery stroke. In solid lines the power-stroke, going from ahead to behind, with the cilium fully extended. In dashed lines the recovery stroke, during which the cilium curls back to its initial position [1].

with $\tau = \frac{m}{6\pi\eta a} = \frac{2a^2\rho_m}{9\eta}$, being ρ_m the density of the microorganism, which can be assumed equal to that of water. The distance covered by the bacterium before stopping is then

$$s = \int_0^\infty v(t)\mathrm{d}t = v_0\tau = 10^{-5}\frac{m}{s} \cdot 2 \cdot 10^{-7}s = 2 \cdot 10^{-12}m \tag{1.4}$$

which is smaller then the radius of an hydrogen atom, and also many order of magnitude smaller than the bacterium itself. The bacterium will stop moving immediately if no force is pushing it! This gives an idea of how different the life of microorganisms is from what we experience.

Let us now get an overview of the mechanisms of propulsion which may be effective in these conditions. The first point to make is that the reversible nature of the hydrodynamic equation for low Reynold numbers (the limit of the Navier-Stokes equations for Re $\to$0) makes useless those mechanisms of motion which are *reciprocal*, i.e. which consist in a deformation of the body to a certain shape, and then in a return to the initial shape by reversal of the sequence. This was first explained by Purcell in his article [8] under the name of *Scallop Theorem*. The name refers to the way scallops move, opening and closing their two half shells, which being reciprocal wouldn't work if they swam at low Re (but does work in water for them).
Purcell proposed various swimming mechanisms that avoid this problem, among which flexible oars. This is the solution employed by eukaryotic organisms, like protists, which exploit the flexibility of cilia or flagella, cylindrical appendages made of microtubules, which they move in a way similar to breast stroke (see fig. 1.1). Each cilium performs a repetitive cycle of movements consisting in a fast power stroke, during which the cilium extends in the fluid, and a slow recovery stroke during which the cilium curls back to the initial position. While flagella refers to a little number of long appendages, cilia refers to a large number of small ones, usually covering the whole body and possibly organized in rows beating with a cyclical pattern.
Prokaryotic organisms, like bacteria, also have flagella but with different features: in contrast with the flagella of protists, theirs are semi-rigid helical filaments rotated by a protein motor situated in the body, at the root of the flagellum. Some species have a unique filament which can rotate both counter-clockwise and clockwise, one sense providing forward motion and the other backward. Other species, as *E. coli*, a bacterium that has been extensively studied [2, 9], have more filaments that form a unique bundle when rotating in one sense (forward motion), but disband when rotating in the other, causing a change of direction. This mechanism of propulsion gives rise to the motion known as *run and tumble*: the bacterium alternates runs, when it moves mainly straight (unless some deviation due to rotational diffusion), and tumbles, when the directional changes occur thanks to the reorientation of the flagella. Duration of runs and tumbles are exponentially distributed, and tumbles display a mean



(a) **(b)**

Figure 1.2: (a) Forces on a segment of thin wire falling slantwise in a viscous medium at velocity v. This velocity can be decomposed into components perpendicular to the wire and parallel to the wire, as shown below the wire. The drag due to the perpendicular component is twice as great per unit velocity as the drag due to the parallel component, as shown by the dotted lines above the wire. The net drag is F. It is not vertical but is tilted to the right, so it has a horizontal as well as a vertical component, as shown by the dashed lines. The horizontal component tends to move the wire to the right. If the wire were a segment of a rotating helix, this component would provide thrust. The vertical component opposes v, and thus determines the power required to move the filament. If the wire were a segment of a rotating helix, this component would contribute to the torque required to rotate the helix [2]. (b) Mechanisms of generation of a forward thrust by rotating an helical filament in a viscous fluid. Two segments of the helix are considered, moving with opposite velocity: one upwards, $\vec{v}$, and one downwards, $\vec{v}'$. The same decomposition indicated in (a) can be done, and with the same reasoning one finds the components of the drag force. As indicated in the figure, the vertical components F_Ω and F'_Ω are equal and opposite, so they contribute to the torque on the helix, but not to the net force. The horizontal components F_v and $F_{v'}$ act in the same direction, and contribute to thrust [6].

duration an order of magnitude smaller than runs (0.1 s for tumbles, 1 s for runs) [2].
The generation of a forward thrust by rotating an helix in a viscous medium is due to the viscous drag on a thin filament being larger when the filaments moves sideways than when it moves end-on. The mechanisms is explained very clearly in [2], the book by Berg on the motion of *E.coli*. Consider a segment of thin wire falling in a viscous medium vertically, and another falling horizontally. The second is subject to a viscous drag which is about twice as large at that of the first due to the different outline it offers to the fluid. If now we drop a wire tilted downwards to the right, like in figure 1.2(a), thanks to this difference in the viscous resistance in the two directions (parallel and perpendicular to the wire), it will fall slantwise to the right.
If we apply the same reasoning to each segment of a rotating helix, as seen in figure 1.2(b), there will be a net force in the forward direction (right in the figure) given by the sum of the forces parallel to the axis on each segment of helix. The components of the viscous force perpendicular to the axis contribute to the torque, which is counterbalanced by the torque of the body of the bacterium which rotates in the opposite sense. If the helix does not have and integer number of turns, the net force perpendicular to the axis is not zero, and gives rise to a conical wobble.

Other kinds of motion may be observed on solid or semi-solid media: organisms like amoeba move by means of changing the shape of their body, with a mechanism known as *crawling*. They cyclically create protruding extensions of their body, called *pseudopods*, using which they push themselves forward. More pseudopods can be present at the same time on the body, and their position determines the direction of motion. Each pseudopod follows a typ-

ical life cycle: a new one is extended every $\sim$ 15s and it expands for $\sim$ 12s. Well after it has ceased to expand, it remains recognizable as a side bump on the body. It was suggested that these extensions may be used for gliding on smooth surfaces, to walk on surfaces with obstacles and also to swim [10]. In [11] it was found that pseudopodia can be of two types, *split* and *de novo*, according to whether they grow out of an existing pseudopod or not. Series of the first kind give rise to a motion with directional persistence, while the interruption of such a series by a pseudopod of the second kind gives rise to a turn.

To conclude this overview, the motion of microorganism is influenced by many factors. The medium in which they move determines which technique of swimming is necessary. The size of the organism is also crucial, not only because it determines the condition of low Reynolds number in which it swims, but also because it seems there is a minimum size for active motion, below which it is not advantageous to actively look for nutrients (i.e. Brownian diffusion is sufficient) [12]. Furthermore, the shape of the microorganism itself determines the viscous drag it experiences, causing an ellipsoidal shape to be convenient. In fact, it can be shown [13] that a spherical shape is the most efficient for diffusion, while an elongated one is the best for swimming; the result is that the optimal ratio of the two axes is 2:1 (for many bacteria a similar ratio of 3:1 is found). As a last remark, the density of organisms is also a factor to consider, since at high concentrations collective motion was observed [14, 15]. Consequently, to observe individual motion diluted conditions are required.
Much more could be said on mechanisms of motion: the variety of organisms to study is great and the literature on the subject is vast. Since this is not the central issue of this thesis, which is instead focused on modelling the statistical properties of motion, we stop here, hoping that the material reviewed so far is be sufficient for the reader to form an idea of the kind of phenomenon we are going to study.

1.2 Statistical features of motion

From a statistical point of view, motion of particles can be characterized by certain quantities, e.g. the velocities distribution and autocorrelation, which can be extracted from the trajectories of these particles.
For microorganism methods of tracking trajectories by means of sequential photographs appeared in the '60 and '70, and since then a great variety of species have been observed. More recently the high degree of automatization and the quality of microscopes available enhanced greatly the quality and quantity of data, with trajectories followed up to hundreds of body-lengths[3].
This section presents some of the latest experimental observations, some drawn from the book by Méndez et al. [3], which reviews the results obtained in several recent works, and some directly from articles. All results presented are relative to spontaneous movement in an isotropic and homogeneous medium, in the absence of external cues.

Velocity distribution

In the case of isotropic motion what one usually looks at is the distribution of the modulus of $\vec{v}$, P(v). Let us remark that if $\vec{v}$ was a Gaussian variable, the distribution of its modulus would be in the 2D case the so-called Rayleigh distribution $P(v) \propto ve^{-v^2}$ and in the 3D case the Maxwell distribution $P(v) \propto v^2 e^{-v^2}$. This because in changing variables one has to multiply by the Jacobian of the transformation.

[3]Depending on the organism, this may span from seconds to hours of tracking

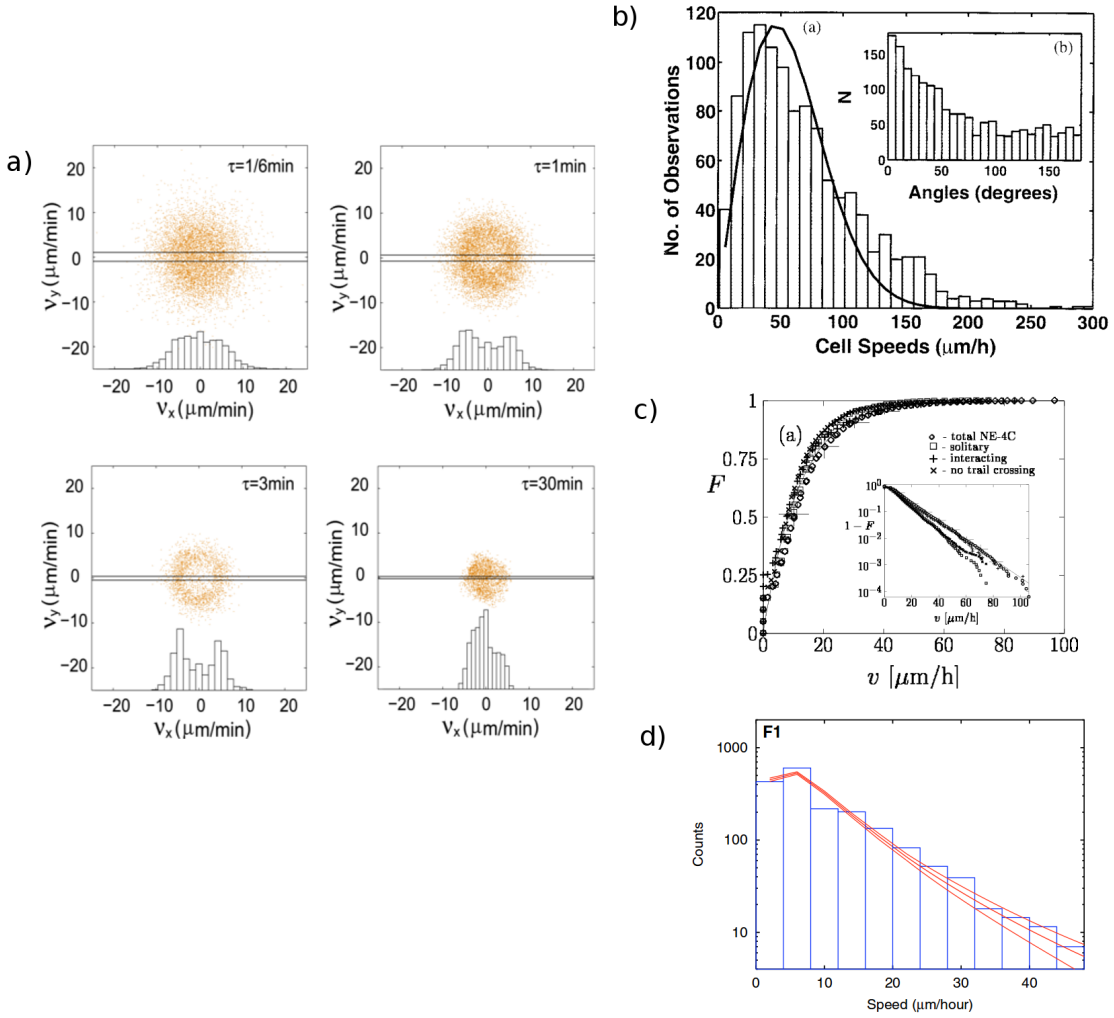


Figure 1.3: Speed distribution observed in experiments on different species. a) $P(v_x, v_y)$ for *Dictyostelium discoideum*, taken from [17], for different values of the time lapse between two measurements of position. b) $P(v)$ for *Hydra* cells, taken from [16]. The solid line is the fit with a Rayleigh distribution. c) CDF of speed for various cells, among which human fibrocytes, found in [5]. Inset: graphic of $1 - \mathrm{CDF}(v)$ on lin-log scale, showing that the curve is an exponential. d) $P(v)$ for HaCaT cells, from [18]. The red lines are fits of the distribution with a model that will be discussed in section 1.3.

Experimental velocity distributions do not seem to have this form. Various different distributions have been observed, some of the results are shown in figure 1.3. In [16], observing the trajectories of *Hydra* cells[4], they find a $P(v)$ with a heavier tail than a a Rayleigh distribution (see fig. 1.3(b)), while in [17] they look at $P(v_x, v_y)$ for *Dictyostelium discoideum*, an amoeboid cell moving with the crawling mechanisms explained in section 1.1, and find it is very different from a Gaussian distribution, showing a lower probability for speeds close to zero (see fig. 1.3(a)). Note that they observe a variation of the distribution if they vary the interval of time Δt between two measurements of the position used to compute $\vec{v}$ as $\vec{v} = \frac{\vec{x}(t+\Delta t)-\vec{x}(t)}{\Delta t}$. For small Δt they find a volcano-shaped distribution, while as Δt gets larger the "hole" fills-up and the distribution is peaked at the center. In fact, for Δt much larger than the typical persistence time of the trajectory, what we measure is not the ballistic velocity of a straight segment of trajectory, but a velocity that is the average on many segments of trajectory, with different orientations. Since the orientations are isotropic, this averaged velocity will be close to zero. This gives us the opportunity to remark that an appropriate resolution (Δt smaller than the typical persistence time) is essential to observe the right P(v).
In [5] they observe a distribution that decays exponentially, i.e. its Cumulative Probability Distribution (CDF) is well fitted by $CDF(v) = 1 - e^{-\alpha v}$ (see fig. 1.3(c)). The cell species on which they observe this behaviour are various: mouse fibroblasts[5], human glioblastoma and others, and authors claim this could be a general behaviour. Note that this distribution implies that P($\vec{v}$) is divergent at the origin. This is not impossible, but in [18] (fig. 1.3(d)), from the observation of HaCaT cells[6], an exponential distribution is seen only for velocities above a certain value, while below there is a different trend. If we looked at P($\vec{v}$), it would have an enhanced probability of small velocities with respect to a Gaussian.
What emerges from this review of experimental findings is that there is a great non uniformity of results, but a common feature is non-Gaussianity of P($\vec{v}$). We must stress that the cells observed in these experiments have very different features, from the dimensions (from $\sim$ 1μm of *Dictyostelium discoideum* to $\sim$ 50-60μm of human fibroblasts, to $\sim$1mm of *Hydra*), to the mechanisms of motion (crawling, swimming,...), to the average speed (of the order of μm/min in the case of *Dictyostelium discoideum*, while for fibroblasts only $\sim$ μm/h). It is certainly a possibility that this differences influence deeply the statistical features of motion. However, as is made clear from fig. 1.3(a), a small difference in the experimental resolution con qualitatively change the observed P(v). Thus, precision of tracking methods could bias results, making it difficult to compare results of different experiments. This could also be the cause of the contradiction between [5] and [18], which report different results for similar cells.

Velocity autocorrelation

Another feature characterizing the motion is the autocorrelation of velocity, which measures the persistence of motion, i.e. how long direction of motion is maintained.
It is defined as

$$C(\tau) = \langle \vec{v}(t) \cdot \vec{v}(t+\tau) \rangle \tag{1.5}$$

In figure 1.4 some examples of this quantity, also from different experiments with different species. In all cases a non zero autocorrelation time is found, indicatince persistence, but the trend varies: for *Hydra* cells and ephithelial cells a power-law decay is found (1.4(a) and (b)),

[4] *Hydra* is a small animal of $\sim$1mm, which moves moves by a crawling mechanism.
[5] Fibroblasts are the most common cells of connective tissue in animals
[6] HaCaT are a kind of human keratinocytes, motile cells of the epidermis which migrate to heal wounds.

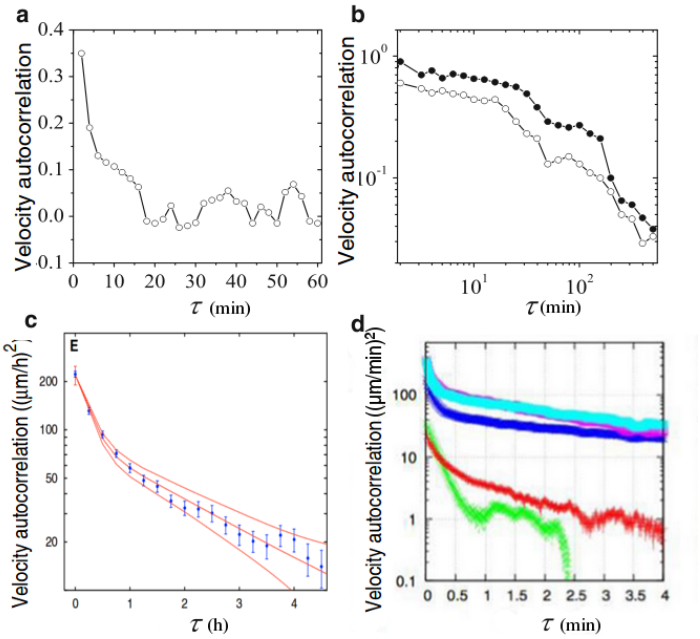


Figure 1.4: Velocity autocorrelation observed in experiments on different species: **a**) *Hydra* cells, from [16]. **b**) Epithelial Madin-Darby canine kidney cells of wild-type (full circles) and NHE-deficient (open circles), from [19] **c**) *Dictyostelium discoideum* cells, from [20] **d**) *Dictyostelium discoideum* cells during different states of their life, from [21].

while *Dictyostelium discoideum* displays a double-exponential (fig. 1.4(c) and (d)). More insight on why a double exponential could emerge will be given in section 1.3.
Note that in figure 1.4(a) an oscillatory pattern is superimposed to the decay. This can be explained by the zig-zag motion of cells which introduces a periodic oscillation of the velocity, and consequently a periodicity in the velocity autocorrelation, as that found sometimes in crawling amoeba [7, 17] or as the conical wobble described in section 1.1. An oscillating autocorrelation is found also in [7], see figure 1.14 and subsequent discussion in section 1.3.

Mean Square Displacement

Defined as

$$MSD(t) = \langle \vec{x}^2 \rangle \tag{1.6}$$

the Mean Square Displacement gives a measure of the dispersal capacity of the particle. Before looking at experimental findings, let us recall that if the MSD scales as t^2 it is sign of a ballistic motion, while a scaling with t means the particle is in the diffusing regime. If MSD$\sim t^\alpha$ with $\alpha \neq 1, 2$, the motion is called *anomalous diffusion*.

The graphics in figure 1.5 show that data seem to present anomalous diffusion with $1 < \alpha < 2$, at least in a certain transient of time. This can be interpreted in two ways: either the motion must be explained by anomalous diffusion, or the MSD has a form that implies a transient time between ballistic regime and diffusing regime during which the scaling resembles anomalous diffusion. This is for instance the case of an MSD of the form

$$MSD(t) = A(C_1 t - 1 + e^{-C_1 t}) + B(C_2 t - 1 + e^{-C_2 t}) \tag{1.7}$$

where A, B, C_1, C_2 are positive constants. Figure 1.6 shows the behaviour of the MSD of eq. (1.7) and the role of the two time-scales C_1 and C_2, compared with that of the MSD

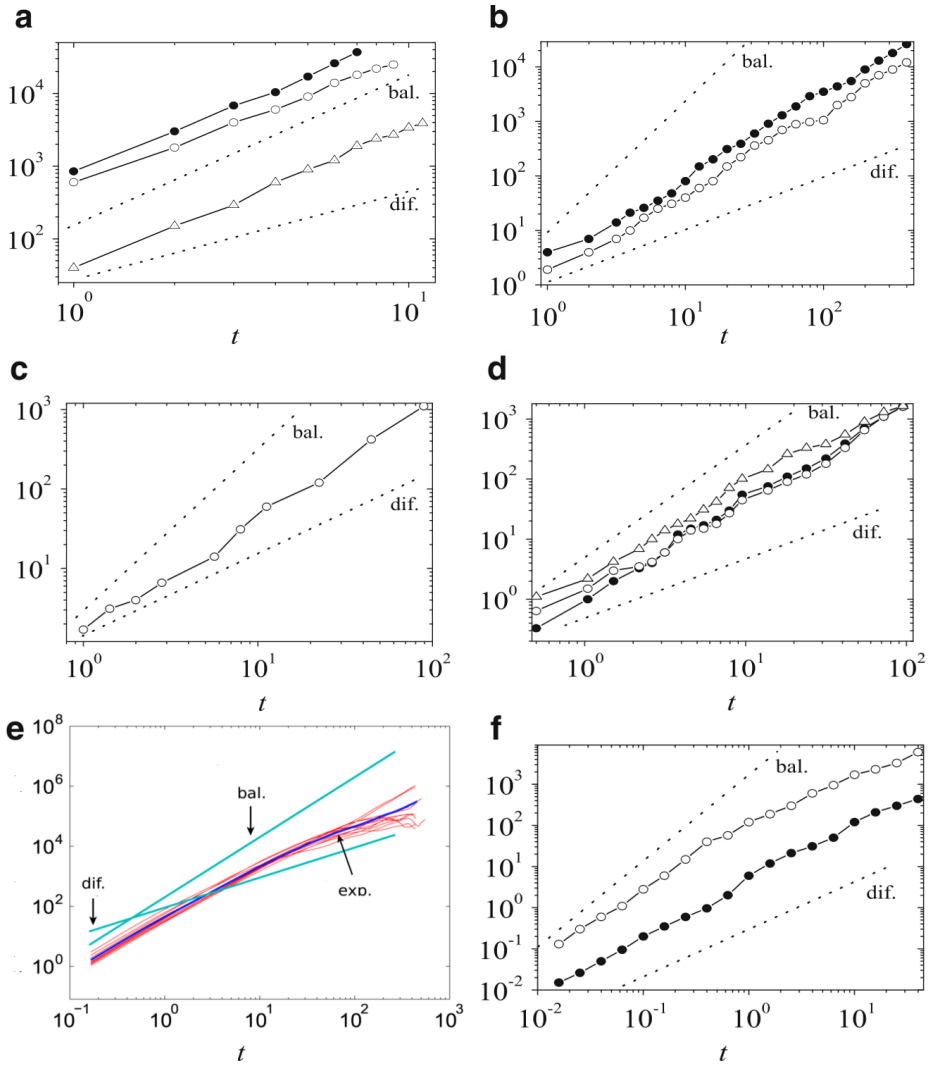


Figure 1.5: Mean Square Displacement observed in experiments on different species. All graphics are taken from [3], but based on data from different articles, as follows: **a**) Cells of amoebas *Tetramitus rostatus* (full circles), *Naegleria gruberi* (open circles) an *Acanthamoeba castellanii* (triangles) [23] **b**) Epithelial Madin-Darby canine kidney cells of wild-type (full circles) and NHE-deficient (open circles) [19] **c**)neuN human mammary epithelial cells [24] **d**)MCF-10A human mammary epithelial cells expressing pBABE (full circles), neuN (open circles) and neuT (triangles), [25] **e**) Strain AX4 of *Dictyostelium discoideum* cells [17] (red lines are for singular trajectories, blu line is the average) **f**) Strain AX2 of *Dictyostelium discoideum* cells at the vegetative state (full circles) and after 5.5 h starving (open circles) [21].
In all graphics, t is measured in minutes, and the MSD is in μm^2. The graphics are in log-log scale, and for comparison the lines corresponding to ballistic scaling and to diffusive scaling are shown. In almost all cases, except e), the slope of the experimental MSD does not seem to match the one expected for diffusive scaling, at least on the observed time scale. See the text for possible explanations of the reason.

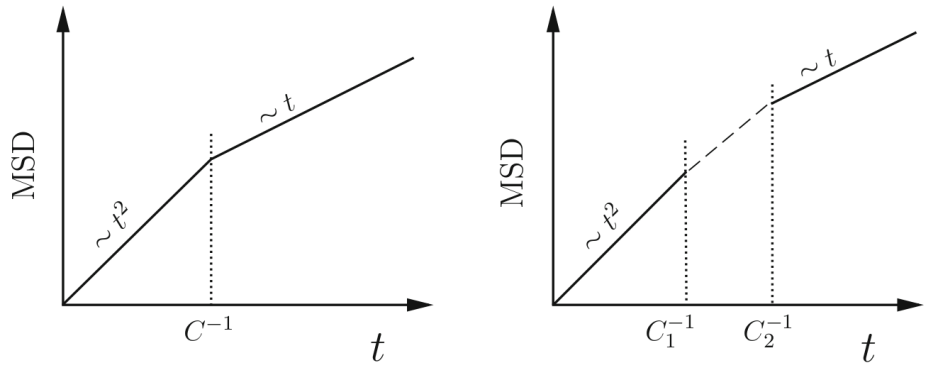


Figure 1.6: Comparison of the behaviour of the MSD predicted by an Ornstein-Uhlenbeck (1.8) process with that of equation (1.7) [3]. The first has only one time scale, and passes from ballistic to diffusive scaling. the second, instead, has two time scales, C_1^{-1} and C_2^{-1}, and displays a transient behaviour of duration $C_2^{-1} - C_1^{-1}$ between ballistic and diffusive regimes, which could look like anomalous scaling.

predicted by an Ornstein-Uhlenbeck process

$$MSD(t) = Ct - 1 + e^{-Ct} \tag{1.8}$$

where only one time-scale is present, and the scaling goes directly from ballistic to diffusive.
An MSD of the form (1.7) is predicted by some models, among which the one proposed in [22] that will be explained in section 1.3.
An experimental support to the hypothesis that anomalous scaling is only a transient behaviour is given by [7], where diffusive behaviour is observed for the MSD (see fig 1.5(e)), and authors comment that it is because they succeed in tracking the motion longer than other experiments, and consequently they see the end of the transient, culminating in diffusive scaling (while others stop before that time[7]).

Kurtosis

Kurtosis is defined as

$$K(t) = \frac{\langle \vec{x}^4 \rangle}{\langle \vec{x}^2 \rangle^2} \tag{1.9}$$

and it gives information about how much the observed process is different from diffusion since for a diffusion process this ratio has the value of 3. In fact, the position $\vec{x}$ of a diffusion process is a Gaussian variable with zero mean. Since for a Gaussian variable with mean μ and standard deviation σ

$$\langle \vec{x}^4 \rangle = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4 \qquad \langle \vec{x}^2 \rangle = \mu^2 + \sigma^2 \tag{1.10}$$

if $\mu = 0$ the kurtosis is equal to 3.
For the Ornstein-Uhlenbeck process (see sec. 1.3 and appendix A), $\vec{x}$ is also a Gaussian variable, but with mean $\mu = \frac{v(0)}{\rho}(1 - e^{-\rho t})$. Since for large t its standard deviation scales as t, while the mean remains constant, the kurtosis approaches the value 3 growing monotonically. As can be seen in figure 1.7 the kurtosis found experimentally is monotonically decreasing,

[7] What one must compare are not the tracking times, but the ratios of the tracking time and the persistence time, because the long-run trend of the MSD is visible after many persistence times have passed.

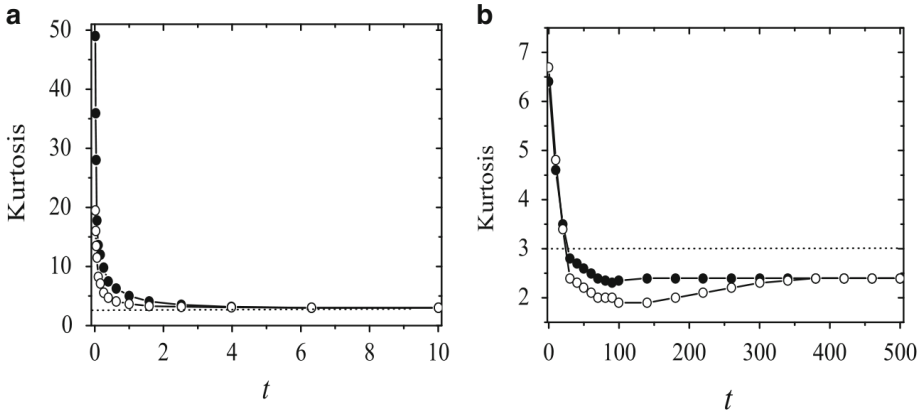


Figure 1.7: Kurtosis observed in experiments on different species. Graphics are taken from [3], but based on data from various articles, as follows: **a**) Strain AX2 of *Dictyostelium discoideum* cells at the vegetative state (full circles) and after 5.5 h starving (open circles) [21] **b**) Epithelial Madin-Darby canine kidney cells of wild-type (full circles) and NHE-deficient (open circles) [19].
In both graphics, t is in minutes.

and it does not always approach 3. This means that nor plain diffusion nor the Ornstein-Uhlenbeck process correctly describe the motion observed.

The observed behaviour of kurtosis, together with that of the MSD, hints that the motion process is such that the position is not Gaussian. This means that either the increments are non-Gaussian, or they are not independent. Consequently, to correctly describe the motion we will have to keep this in mind.

Turn angles distribution

The distribution of turn angles also gives us precious information on the motion, telling us whether the microorganism changes its direction gradually (distribution centred at small angles) or makes neat turns. Before looking at experimental results, let us establish a point: many microorganisms, due to their specific mechanisms of motion, have a noisy zig-zagging pattern[8]. In this case, we must distinguish between the instantaneous angle, which resents of the noise, and the "average" angle, which indicates the average direction of motion (see figure 1.9). A detailed study of the problem was made in [17] for *Dictyostelium discoideum* cells, and the distribution found for the average angle displays a bimodal pattern (fig. 1.8(a)). A similar distribution is also found in [26]. Such a distribution is characteristic of a motion similar to the*run and tumble* motion of *E.coli*: the particle stops and chooses a new direction which can differ substantially from the previous one. The choice of the turn angle seems to be biased in the forward direction, at least for some species. For *E. coli* the measure has been taken with result $\langle\theta\rangle$=68° [6] (if the probability were uniform, the mean turn angle would be 90°). In [27], although, a distribution peaked at small angles is found, indicating that the persistence of the motion lasts also across several "runs" (fig. 1.8). The studied specie is the same in all three cases, so the different results pose doubts on the compatibility of experimental conditions among different experiments.

[8]It is the case of crawling amoeba, like *Dictyostelium Discoideum*, and also of some ciliates like the *Colpidium* sp. which motion we will examine in chapter 3

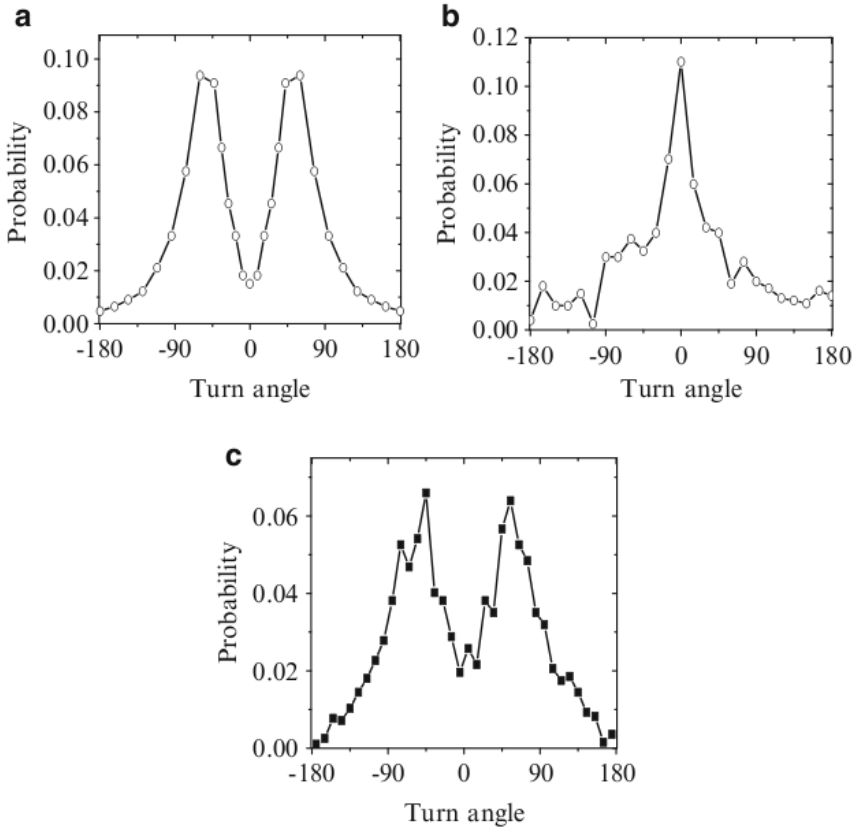


Figure 1.8: Turn angle distribution observed in experiments on different species. Graphics are taken from [3], but based on data from various articles, as follows: **a**) Strain AX4 of *Dictyostelium discoideum* cells [17] **b**) *Dictyostelium discoideum* cells [27] **c**) Strain AX3 of *Dictyostelium discoideum* cells [26].

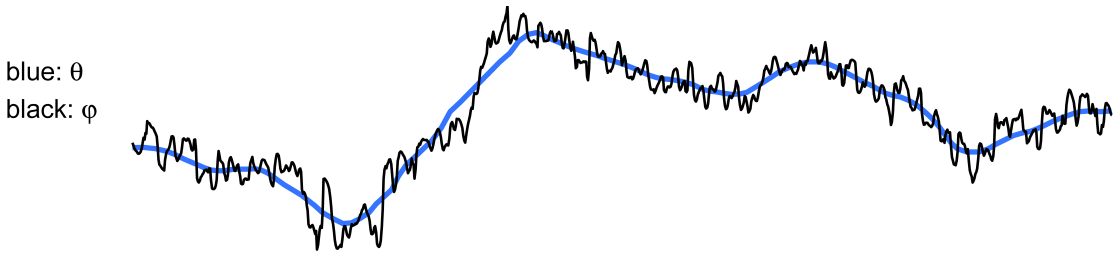


Figure 1.9: Example of a trajectory where a noisy pattern (in black, described by the instantaneous angle ϕ) is superimposed on an average motion (in blue, described by the "averaged" angle θ) [17].

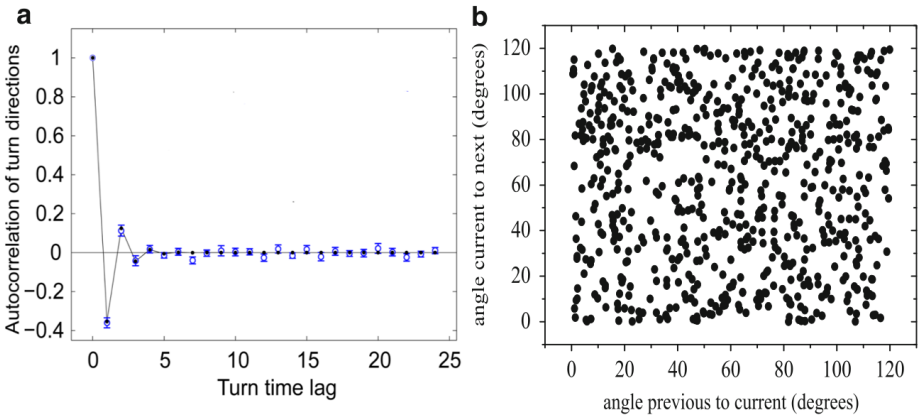


Figure 1.10: Different representations of turn angle correlations observed in experiments on different species. **a**) Strain AX4 of *Dictyostelium discoideum* cells [17]. Black points are a fit with the simple model for turns explained in the text. **b**) Strain AX3 of *Dictyostelium discoideum* cells [26].

Angles correlation

Another quantity that gives us information about the type of motion is the correlation of consecutive angles or turn angles. If the motion starts with direction α, after one turn will the new direction be similar to the previous one, or not? And after two turns? Provided that we succeed in determining turning events, which can be not so simple given the oscillations of the direction, this information can be represented in two ways: one solution is to look at the turning direction autocorrelation function, i.e. the correlation between the starting direction and the direction after n turns, with n varying. The other solution is to plot the current turn angle versus the previous turn angle.
Figure 1.10(a) shows the turn direction autocorrelation function relative to *Dictyostelium discoideum*: the plot shows that correlation is negative for consecutive turn directions, and then goes quickly to zero[9]. Authors of [17] argue that this is explained by supposing that the turning process is a markov process, with probability p to turn in the same direction of the previous turn (the black dots in figure 1.10(a) correspond to the autocorrelation of this process, which fits closely the experimental one). Also figure 1.10(b), where the current turn angle versus the previous turn angle is plotted, suggests there is no correlation between consecutive changes of direction.

Run and Tumble

From the observation of trajectories it was discovered that many microorganisms, including for instance *E. coli*, alternate directional displacement (runs) and angle reorientations (tumbles). This motion was then called *run and tumble*, or also *bimodal motion*. To understand whether a certain organism moves in this way visual inspection of trajectories can provide a first intuition, but a quantitative evidence comes from the correlation between the turn angle rate and the velocity. In fact in run and tumble motion changes of the angles occur during pauses of the motion, so a negative correlation is expected (see figure 1.11).

[9]Turns refer to the "noise averaged" trajectory.

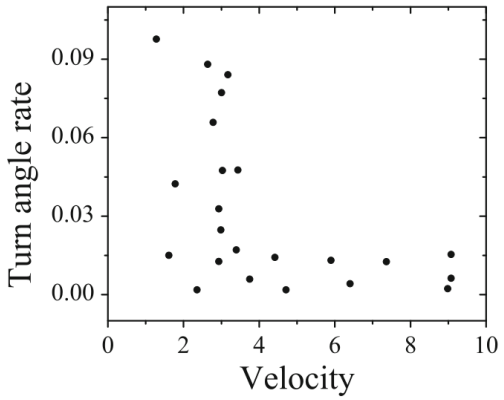


Figure 1.11: Correlations between the turn rate and the velocity for the AX3 strain of *Dictyostelium discoideum*, showing evidence of run and tumble motion [28]

1.3 Stochastic models of motility

After having examined the statistical features of the motion of microorganisms, we are only partially satisfied. We know how they behave, but why do they behave like that? What are the laws controlling their motion, and giving rise to those particular features we observe?
The process of modelling motility is then interesting not only for predictive purposes, but also because it gives us an insight of the origins of what we see. A good model should, in fact, not only give good predictions but also offer a realistic and enlightening description of the process. Modelling self-propelled motility poses a particularly stimulating challenge, since it requires that we make hypotheses on the propelling force exerted by the microorganisms themselves, which makes the problem quite different from the one of modelling the motion of non-living particles.

Modelling can be dealt with on different levels, depending on the degree of detail of the information we have on the motion and on the degree of detail we want to reach. If the tracking of trajectories is available, we may pursue a microscopic description, based on a Langevin equation, which predicts the probability of finding a single individual at position $\vec{x}$ at a certain time t.
On the other hand, if we only have averaged data for the motion, there is no use in going further than a mesoscopic description, which produces as a output the density of individuals expected in average at position $\vec{x}$ at a certain time t. With such a description, although, we can still keep in consideration aspects like memory effects, which are still deducible from the averaged information on trajectories [3].
Finally, if all we can measure are effective parameters (like diffusion coefficients), the appropriate description is a macroscopic one, also dealing with probability distributions for the average behaviour of the population but, differently from the mesoscopic description, not accounting anymore for memory effects or other such subtle aspects of the process.
In the following we will focus on microscopic description, since the original model proposed in chapter 3 of this thesis is a microscopic one. This kind of model, although more demanding from the mathematical point of view, gives the most accurate and exhaustive description of the process, and consequently is the most useful to characterize motility.
This kind of approach is more common in the study of the motion of cells and microorganisms than in that of bigger organisms (like animals), since tracking of trajectories is easier.

In this section we will try to retrace the attempts that have been made to model active dispersal of microorganisms, dwelling on some models of particular interest. Again, the discussion follows that in the book by Méndez et al. ([3]).
When naming dispersal one cannot but think of diffusion, which was first introduced to describe the dispersal of molecules in water. Microscopically, the equations of motion for a particle in a viscous fluid are given by the Langevin equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v \tag{1.11}$$

$$m\frac{\mathrm{d}v}{\mathrm{d}t} = -\rho v + \sigma\xi(t) \tag{1.12}$$

where $\xi(t)$ is a Gaussian white noise, i.e. a stochastic process with the following properties:

1. $\xi(t)$ is Gaussian
2. $\langle\xi(t)\rangle = 0$
3. $\langle\xi(t)\xi(t')\rangle = \delta(t-t')$

However, for a microorganism in water, as we have seen in sec. 1.1, the inertial forces are negligible with respect to the viscous drag, which means $\mathrm{m}|\frac{\mathrm{dv}}{\mathrm{dt}}| \ll |\rho\mathrm{v}|$, so equation (1.11) reduces to

$$\rho\frac{\mathrm{d}x}{\mathrm{d}t} = \xi(t) \tag{1.13}$$

meaning that x_t is a Wiener process, and v_t is a purely fluctuating variable.This does not seem a realistic description of the process, though. In fact, it does not describe the observed persistence of direction: at each step the direction chosen for the motion does not depend on the previous history of the trajectory, since v_t is a white process, and changes in an erratic way. On the other hand, the velocity of microorganisms changes smoothly.
We need then to find a different process to describe the motion, which gives a more realistic description. It turns out that the Ornstein-Uhlenbeck (O-U) process, which description is given by eq. (1.11), does a much better job. In fact if we keep the inertial term that we had neglected due to its being much smaller than the friction term we obtain that the value of v at a certain time depends on its previous values, and consequently the particle tends to maintain its velocity (and thus its direction) for a certain time, a measure of which is ρ^{-1} (*persistence time).* The resulting trajectory of the velocity is smooth: see in figure 1.12 a comparison between the fluctuating velocity found in the strong friction limit and the velocity of the O-U process.

It is important to remark that the term $\frac{\mathrm{d}v}{\mathrm{d}t}$ that we are now keeping does not represent true, physical inertia (since that is negligible, as said above). We are giving a statistical description of the motion, or better a phenomenological description: we observe that the motion shows persistence, and consequently describe it by means of a process which predicts this feature. This means also that ρ does not have the meaning of friction coefficient any more: we can only attribute to it the statistical meaning of persistence parameter.

Now we are going to examine some of the two dimensional models of motion based on the O-U process. In appendix A the one-dimensional O-U process is defined and the main results, which will be employed in the following, are derived.

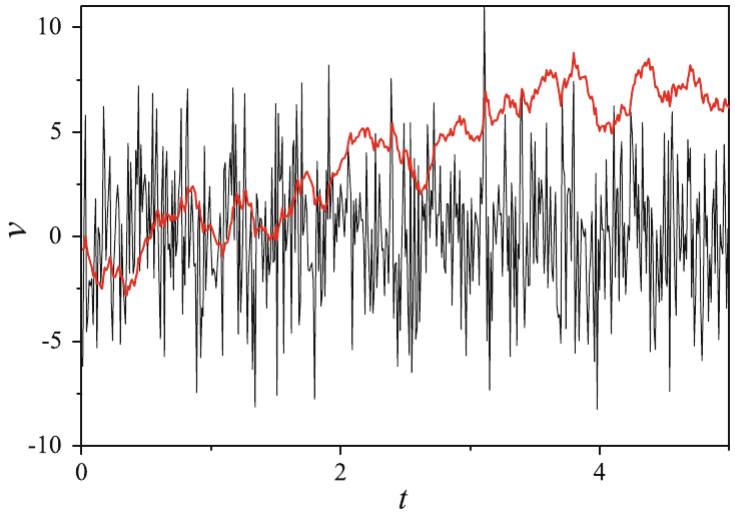


Figure 1.12: Comparison of the evolution of velocity described by a Gaussian white noise (strong friction limit) and by an O-U process [3].

Models based on the Ornstein-Uhlenbeck process

Since the motion observed experimentally is two or three-dimensional, we must find an appropriate extension of the 1D O-U process analysed in the previous section. Let us discuss the 2D case, both for simplicity and because the original model presented in chapter 3 is 2D.

The first idea that comes to mind is to take two independent O-U process in the two directions, as follows

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v_x \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = v_y \tag{1.14}$$

$$\frac{\mathrm{d}v_x}{\mathrm{d}t} = -\rho v_x + \sigma\xi_x \qquad \frac{\mathrm{d}v_y}{\mathrm{d}t} = -\rho v_y + \sigma\xi_y \tag{1.15}$$

where $\xi_x(t)$ and $\xi_y(t)$ are two independent Gaussian white noise terms. A model of this form was employed for instance in [29].
A likable feature of this equations is that they predict anticorrelation between the velocity and the turn angle rate, meaning that when the velocity is large the direction of motion does not change much, and viceversa. This can be seen passing to polar coordinates (v,θ), for which the equations become [30]

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -\rho v + \frac{1}{2}\frac{\sigma^2}{v} + \sigma\xi_v \tag{1.16}$$

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{\sigma}{v}\xi_\theta \tag{1.17}$$

where $\xi_v(t)$ and $\xi_\theta(t)$ are again two independent Gaussian white noise terms.
Such a dependence between the dynamic of v and θ is often found experimentally for motions of the *run and tumble* type (see fig. 1.11 of section 1.2).

However, this aside, it clearly has some flaws. First of all it seems quite unrealistic for the motion in the x direction to be totally independent from that in the y direction. In particular,

if the noise terms must represent the stochastic propelling force they should be the projection on the two axis of a certain vectorial force exerted by the microorganism, and so they would not be independent. Moreover, the statistical features it predicts for the motion ($P(v)$, velocity autocorrelation, MSD,..) do not seem to describe those observed experimentally (of which we had an overview in section 1.2).
Let us see what its predictions are. Employing the results for the 1D case (see appendix A), we can straightforwardly derive those for the 2D case, where we have two identical and independent copies of the 1D process.
To begin with, $\vec{v}(t)$ is again Gaussian, since its two components are Gaussian and $P(\vec{v},t) = P(v_x,t)\cdot P(v_y,t)$.
The autocorrelation is

$$\begin{aligned}\langle \vec{v}(t)\vec{v}(t+\tau)\rangle &= \langle v_x(t)v_x(t+\tau)\rangle + \langle v_y(t)v_y(t+\tau)\rangle = \\ &= \Big[v_x(0)^2 + v_y(0)^2 - \frac{\sigma^2}{\rho}\Big]e^{-\rho(2t+\tau)} + \frac{\sigma^2}{\rho}e^{-\rho\tau}\end{aligned} \tag{1.18}$$

while for the MSD we find

$$\begin{aligned}\langle \vec{x}(t)^2\rangle = \langle x(t)^2\rangle + \langle y(t)^2\rangle = \frac{v_x(0)^2+v_y(0)^2}{\rho^2}(1-e^{-\rho t})^2 + 2\frac{\sigma^2}{\rho^2}\Big[t - \frac{2}{\rho}(1-e^{-\rho t}) + \\ + \frac{1}{2\rho}(1-e^{-2\rho t})\Big]\end{aligned} \tag{1.19}$$

If we furthermore assume that we start at stationariety, i.e.

$$v_x^2(0) = v_y^2(0) = \frac{\sigma^2}{2\rho}$$

the above results simplify to

$$\langle \vec{v}(t)\vec{v}(t+\tau)\rangle = \frac{\sigma^2}{\rho}e^{-\rho\tau} \tag{1.20}$$

$$\langle \vec{x}(t)^2\rangle = \frac{\sigma^2}{\rho^3}\Big(2\rho t - 1 + e^{-2\rho t}\Big) \tag{1.21}$$

where again we can identify the two regimes, *ballistic* and *diffusive*.
Finally, the kurtosis for one of the two components is [3]

$$\frac{\langle x^4\rangle}{\langle x^2\rangle^2} = \frac{(1-e^{-2\rho t})^4 + 3(4\rho t - 1 + e^{-4\rho t})(4\rho t - 3 + 4e^{-2\rho t} - e^{-4\rho t})}{4(2\rho t - 1 + e^{-2\rho t})} \tag{1.22}$$

which for large t approaches 3 growing monotonically.

To summarize, according to the O-U process in 2D

- $\vec{v}$ has a Gaussian probability distribution
- the velocity autocorrelation has an exponential decay
- the MSD shows a transition between ballistic and diffusive behaviour
- the kurtosis increases monotonically and approaches 3.

Looking at the findings listed in section 1.2, it seems that the experiments show quite different features: non-Gaussianity, anomalous diffusion, a decreasing kurtosis which does not always approach 3.

However we must keep the O-U process in consideration as a general framework for the description of the motion, since an equation of that form correctly describes the motion persistence and this is a key factor in the motility of living organisms. For this reason other models made various modifications to the process we just described, in order to try to solve its flaws.

In [19] the authors propose to employ a modified form of the O-U Fokker-Plack equation, with the introduction of fractional derivatives (see [31], appendix) which account for long-range time correlations. This predicts anomalous diffusion and a power law decay of the autocorrelation of speeds, explaining the behaviour they observe for migrating epithelial cells.

In [7] and [18] an interesting solution is suggested to reproduce a form of the velocity autocorrelation that is not a simple exponential like the O-U process would predict. In both articles, in fact, a more complex behaviour is found for the autocorrelation. In [18] two types of motile cells are observed, fibroblasts and keratinocytes, and it is found that the autocorrelation has a multiexponential form

$$C(t) = C_1 e^{-\frac{t}{\tau_1}} + C_2 e^{-\frac{t}{\tau_2}} \tag{1.23}$$

In [7], instead, the microorganism observed is *Dictyostelium discoideum*. Due to the nature of its mechanism of motion (crawling, explained in sec. 1.1), the microorganisms moves with a zig-zagging pattern, so that the trajectories show a persistent motion overlaid with a noisy periodic motion (figure 1.13). The velocity autocorrelation they find is on the whole decaying as a double exponential , but superimposed to this trend they see oscillations of amplitude that remain constant in time (see figure 1.14). This latter feature is due to the zig-zagging motion.

In both cases a data-driven model is proposed that predicts the observed form of the autocorrelation. The question is, what generalization of an O-U process predicts a certain autocorrelation $C(t)$? The fact that it is not simply an exponential means that the rate of change of velocity $\frac{\mathrm{d}v}{\mathrm{d}t}$ does not depend only on the current velocity like in the O-U equation, but there must also be a dependence on previous velocity, a memory. Thus the proposed form for the process is

$$\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = -\int_{-0}^{t} K(t-\tau)\vec{v}(\tau)\mathrm{d}\tau + \sigma\vec{\xi}(t) \tag{1.24}$$

where $K(t)$ is a *memory kernel* whose form is to be deduced directly from the observed autocorrelation (in this sense the approach is data-driven), and $\vec{\xi}(t)$ is the usual vector of two independent Gaussian white noises. The appropriate form of $K(t)$ is found by means of comparing the power-spectrum of the experimental autocorrelation with the power-spectrum of the autocorrelation for a solution $\vec{v}(t)$ of eq. (1.24).
Let us consider a formal solution of the form

$$\vec{v}(t) = \int_{0}^{t} g(t-\tau)\sigma\vec{\xi}(\tau)\mathrm{d}\tau \tag{1.25}$$

where $g(t)$ is an arbitrary function satisfying $g(t) = 0$ for $t < 0$ to account for the fact that the velocity cannot depend on future values of the noise, for causality.
Which form must $g(t)$ have for equation (1.25) to be a solution of (1.24)? We can discover

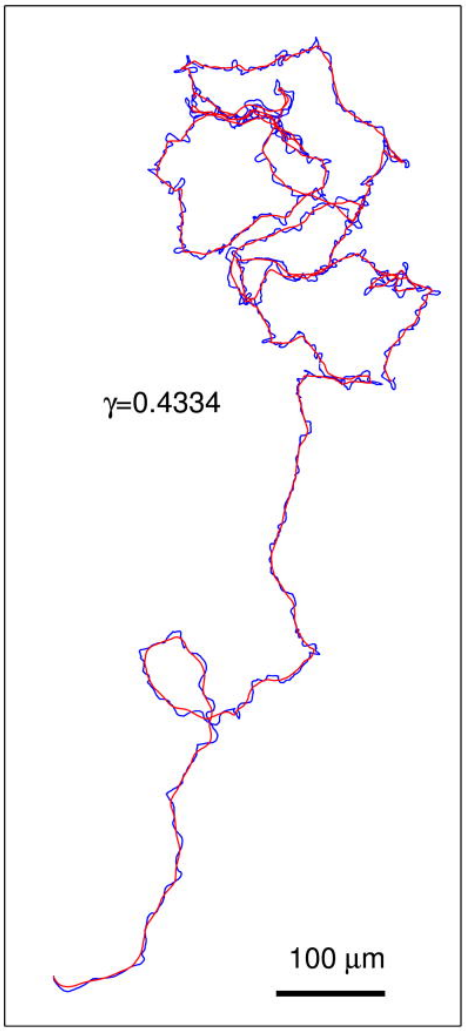


Figure 1.13: An example of trajectory of *Dictyostelium discoideum*, showing persistent motion overlaid with a noisy periodic motion. The red line is an "oscillation-averaged" motion where only the persistent component is present. It is obtained as the time integral of a mean velocity which is a weighted average of past values of velocity, using as weight the exponentially decaying autocorrelation [7].

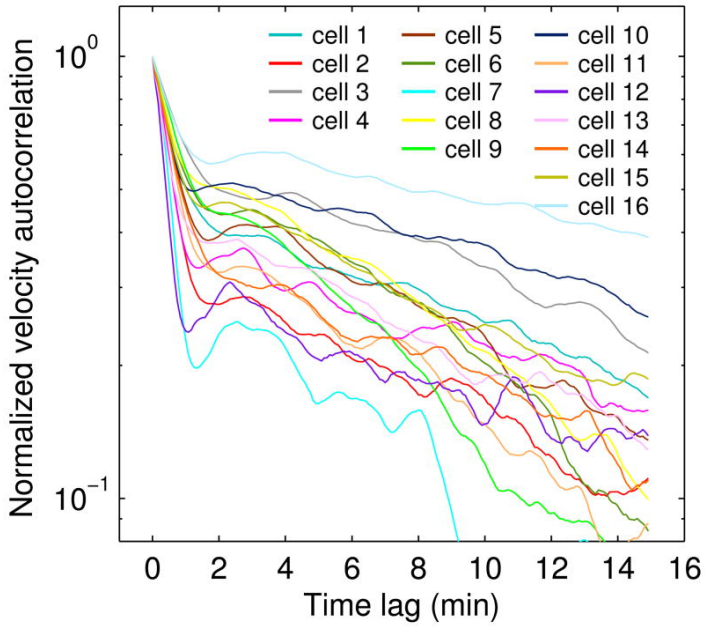


Figure 1.14: Autocorrelation of velocities for trajectiories of *Dictyostelium discoideum*, on a lin-log scale [7]. A double exponential decay is found.

this by taking the Fourier transform of (1.25)

$$\vec{v}(\omega) = \sigma g(\omega)\vec{\xi}(\omega) \tag{1.26}$$

and substituting it into the Fourier transform of (1.24). We find

$$g(\omega) = \frac{1}{K(\omega) + i\omega} \tag{1.27}$$

Now the power-spectrum of the autocorrelation of $\vec{v}$, $S_{\vec{v}}$,(which is the Fourier transform of the autocorrelation, thanks to the Wiener-Khinchin theorem) is expressible in terms of $g(\omega)$ with the following method.
Using (1.25) we can write

$$\begin{aligned} C(\tau) = \langle v(t)v(t+\tau)\rangle = \langle \int_0^t g(t-t')\sigma\xi(t')\mathrm{d}t' \int_0^{t+\tau} g(t+\tau-t'')\sigma\xi(t'')\mathrm{d}t''\rangle = \\ = \sigma^2 \int_0^t \int_0^{t+\tau} g(t-t')g(t+\tau-t'')\langle\xi(t')\xi(t'')\rangle \mathrm{d}t'\mathrm{d}t'' \end{aligned} \tag{1.28}$$

Now using

$$\langle\xi(\omega)\xi(\omega')\rangle = \pi S_\xi(\omega)\delta(\omega-\omega') \tag{1.29}$$

where S_ξ is the power spectrum of ξ, and recalling that if $g(t)$ is real then $g(-\omega) = g^*(\omega)$ we get

$$C(\tau) = \pi \int_{-\infty}^{+\infty} \sigma^2 g(\omega)g^*(\omega)S_\xi(\omega)e^{i\omega\tau}\mathrm{d}\omega \tag{1.30}$$

Then we can identify the power-spectrum of velocity

$$S_v(\omega) = C(\omega) = \sigma^2 g(\omega)g^*(\omega)S_\xi(\omega) \tag{1.31}$$

which using $S_\xi = const$ and (1.27) becomes

$$C(\omega) = \frac{\sigma^2}{(K(\omega)+i\omega)(K(\omega)-i\omega)} \tag{1.32}$$

Thanks to this expression, if we fit the experimental autocorrelation with a certain function and compute its power spectrum, we can deduce from the comparison with (1.31) the exact form of $K(t)$.
In [18] they find that a double exponential decay in $C(t)$ requires $K(t) = -\rho\delta(t) + A\theta(t)e^{-\beta t}$, with A, β constants and $\theta(t)$ the Heaviside function. The equation of the model then becomes

$$\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = -\rho\vec{v} - A\int_0^t e^{-\beta(t-\tau)}\vec{v}(t)\mathrm{d}\tau + \sigma(v)\vec{\xi}(t) \tag{1.33}$$

Note that they have considered a variant with σ dependent on the modulus of velocity v: they choose $\sigma(v) = \sigma_0 + v\sigma_1$, because they find, from data, that $\langle\dot{\vec{v}}\rangle \propto v$. This multiplicative noise allows to find a non-Gaussian $P(\vec{v})$, of the form

$$P(v) = C\frac{e^{-\frac{a}{1+v/v_\sigma}}}{(1+v/v_\sigma)^{2+a}} \tag{1.34}$$

where a and v_σ are parameters of the model.

In conclusion, what can be extrapolated by this approach is that a form of $C(t)$ that is not simply an exponential can be explained with a "memory" that keeps track of past values of velocity, which results in an additional term in the O-U equation, and that a non-Gaussian $P(\vec{v})$ can be obtained if the intensity of the propelling force is dependent on v.
These models succeed in describing the observed data, and do so because they are finely tailored on those data. In fact, both the form of $K(t)$ and $\sigma(v)$ are chosen in such a way that they fit the observed $C(t)$ and $P(v)$. Memory is indeed a possibility to explain the trend of the autocorrelation, and possibly different forms of $K(t)$ could also explain other observed behaviours, as power-laws. On the other hand, we will see in the following and also in chapter 3 that to predict non-Gaussianity of $P(v)$ it is not necessary to suppose that the the intensity of the force depends on v.

The other limit of the O-U process we had identified, and haven't addressed yet, is independence of the motion along the two axes.
In response to some experimental evidences [21] showing that accelerations in the component of velocity parallel to the direction of motion behave differently than those in the perpendicular component, some works (for instance [22]) suggested to choose for the modelling the microorganism's frame of reference, and thus to write an equation for the modulus of the velocity and one for the direction. The form suggested in [22] is

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -\rho(v - v_s) + \sigma_v \xi_v(t) \tag{1.35}$$

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \sigma_\theta \xi_\theta(t) \tag{1.36}$$

in the case of undirected motion, which can be generalized for directed motion by modifying the equation for the angle's dynamic.
The motion described by these equation displays fluctuation of the direction with mean zero, while in the direction of motion the velocity is described by an O-U process with the addiction of a term ρv_s. This represents the propulsion force exerted by the cell to keep its cruise speed constant at v_s. The velocity and the direction of motion are considered independent, an hypothesis that can be realistic for some microorganisms but not for others (as we already noted, in the case of *run and tumble* motion the rate of change of the direction is anticorrelated with velocity).
The probability distribution resulting for v is a Gaussian centered at v_s (note that this is not equivalent to having a Gaussian $\vec{v}$, since that in 2D would result in $P(v) \propto ve^{-v^2}$). This has the problem that it allows for negative values of v, which is of course impossible, and the fit to the data relative to granulocites is not perfect, since data show a slightly asymmetrical distribution around the peak velocity (see fig. 1.15). However, the model is interesting since it introduces the novel idea of considering the microorganism's frame of reference, which was then employed by other models.
For the turn angles a Gaussian distribution is predicted, since they are simply the increments of a Wiener process. This agrees with some of the experimental results presented in fig. 1.8. Knowing the distribution for v and θ and using their independence, one can compute the

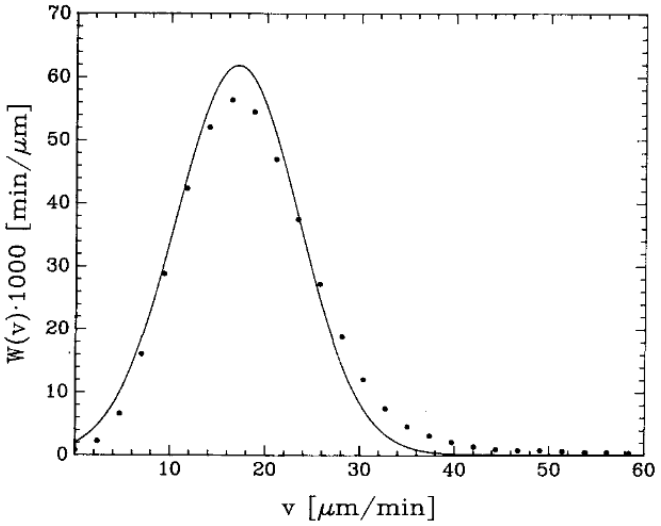


Figure 1.15: Stationary speed distribution for the model (1.35): dots represent experimental results for granulocites, while the line is the model's prediction. [22]

predicted MSD and velocity autocorrelation, with the following results

$$\langle \vec{v}(t)\vec{v}(s)\rangle = \langle v(t)v(t+\tau)\rangle\langle \cos\theta(t)\cos\theta(t+\tau)\rangle = \frac{v_s^2}{2}e^{-\frac{\sigma_\theta^2}{2}\tau} + \frac{\sigma^2}{4\rho}e^{-\Lambda\tau} \tag{1.37}$$

$$\langle x^2(t)\rangle = \langle \Big(\int_0^t v(\tau)\cos\theta(\tau)\mathrm{d}\tau\Big)^2\rangle = \int_0^t\int_0^t \langle v(\tau)v(s)\rangle\langle\cos\theta(\tau)\cos\theta(s)\rangle \mathrm{d}\tau\mathrm{d}s = \tag{1.38}$$

$$= \frac{\sigma^2}{2\Lambda^2\rho}(\Lambda t - 1e^{-\Lambda t}) + \frac{4v_s^2}{\sigma_\theta^4}\Big(\frac{\sigma_\theta^2}{2}t - 1 + e^{-\frac{\sigma_\theta^2}{2}t}\Big) \tag{1.39}$$

where $\Lambda = \rho + \sigma_\theta^2$.
This is a very interesting result, since it presents some feature different from those of the O-U process: both the velocity autocorrelation and the MSD are governed by two different time-scales, Λ^{-1} and $({\sigma_\theta}^2/2)^{-1}$. The velocity autocorrelation is a double exponential, like that observed in some experimental results [18], and the MSD has the form (1.7) which goes from ballistic scaling to diffusive scaling, but in a transient $\Lambda^{-1} - (\sigma_\theta^2/2)^{-1}$ has a behaviour which can look like anomalous scaling, as already explained commenting eq. (1.7). Moreover, the two time-scales can be distinguished only if Λ is sufficiently greater than $\frac{\sigma_\theta^2}{2}$. This could explain why in some cases a single exponential is observed, while in others a double exponential is found.

In [32] a similar model is proposed, where the assumption of independence of v and θ is relaxed. They consider active particle which are polar: they have a head and a tail, and thus a heading direction is defined. This direction, dependent on time, is described by $\vec{e}_h(t) = (\cos\theta(t), \sin\theta(t))$, and let $\vec{e}_\theta(t)$ be the direction perpendicular to $\vec{e}_h(t)$. The modelling assumption they propose is that the stochastic propelling force can be described by two independent Gaussian white noises, one in the direction of motion $\vec{e}_h(t)$ and one in the angular direction $\vec{e}_\theta(t)$, as follows:

$$\vec{\eta}(t) = \sqrt{2D_v}\xi_v(t)\vec{e}_h(t) + \sqrt{2D_\theta}\xi_\theta(t)\vec{e}_\theta(t) \tag{1.40}$$

where D_v and D_θ are the intensities of the two noises.
They write an equation of motion of the form

$$\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = -\gamma(\vec{v})\vec{v} + \vec{\eta}(t) \tag{1.41}$$

where $\vec{\eta}$ is given by (1.40), since $\vec{e}_h(t)$ depends on the current orientation $\theta(t)$ we have obtained a multiplicative noise (giving non-Gaussianity of $P(\vec{v})$) without having to introduce an intensity of noise depending on v (the same happened in [22]).
Let us define $v = \vec{v}\vec{e}_h(t)$, the velocity in the heading direction. This is not the modulus of velocity, given that the particle can move both forward and backward, so both in direction $\vec{e}_h(t)$ and $-\vec{e}_h(t)$. Thus, v may assume also negative values. In the coordinates (v,θ) the equation of the process are

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -\gamma(v)v + \sqrt{2D_v}\xi_v(t) \tag{1.42}$$

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{1}{v}\sqrt{2D_\theta}\xi_\theta(t) \tag{1.43}$$

Thus, we find again a dependence of the turn rate on v.
The form of $\gamma(\vec{v})$ they consider is $\gamma(v) = \rho(1 - v_s/v)$, which gives rise to a cruise speed as that of [22].
One can analytically find the stationary probability distribution of this process, and going back to cartesian coordinates this gives

$$P(v_x, v_y) = \frac{1}{|\vec{v}|}[e^{-A(|\vec{v}|-v_0)^2} + e^{-A(|\vec{v}|+v_0)^2}] \tag{1.44}$$

where $A = \rho/2D_v$. The behaviour of this function for different values of D_v is depicted in figure 1.16. As is clear from its expression (1.44), this function diverges at $\vec{v} = 0$ in a non integrable way. Thus, this model has a pathology, given that a probability distribution must be normalizable.

To conclude this section, let us make a final comment on the models that we have examined. All of them present interesting ingredients, but none is completely satisfactory. In fact, either they do not give correct predictions for the statistical features of the motion, or they make unrealistic assumptions on the motion process, or they have analytical pathologies.
Unrealistic aspects of such models look like the first problem to tackle, since it could be also the cause of the incorrect predictions. The core of any model of active motility is the description of the stochastic propelling force. All the models seen so far employed a Gaussian white noise term, and although some of them abandoned the assumption of independent forces along the two axis, none used a non-white process. Modelling the force by a white process means that it changes value at every time-step, each time assuming a new value completely uncorrelated with the previous ones. But since this is a force exerted by the microorganism, for instance by rotating its flagella which has a certain orientation is space, or by beating its cilia with the body orientated in a certain direction, this would mean that the microorganism keeps changing its orientation while moving, rotating its whole body at every time-step. However, looking at the real motion of microorganisms, it seems that big changes of direction in the motion do not happen at every instant[10], but only after some periods of directed motion. Since direction of motion and direction of the propelling force are clearly connected[11], it seems likely that the force also remains the same for a certain correlation time.
These observations are at the base of the new model proposed in chapter 3 of this thesis.

[10] We are talking of the "average" motion where we neglect the noise due to the particular motility mechanism.
[11] As explained in section 1.1 talking of swimming at low Reynolds number

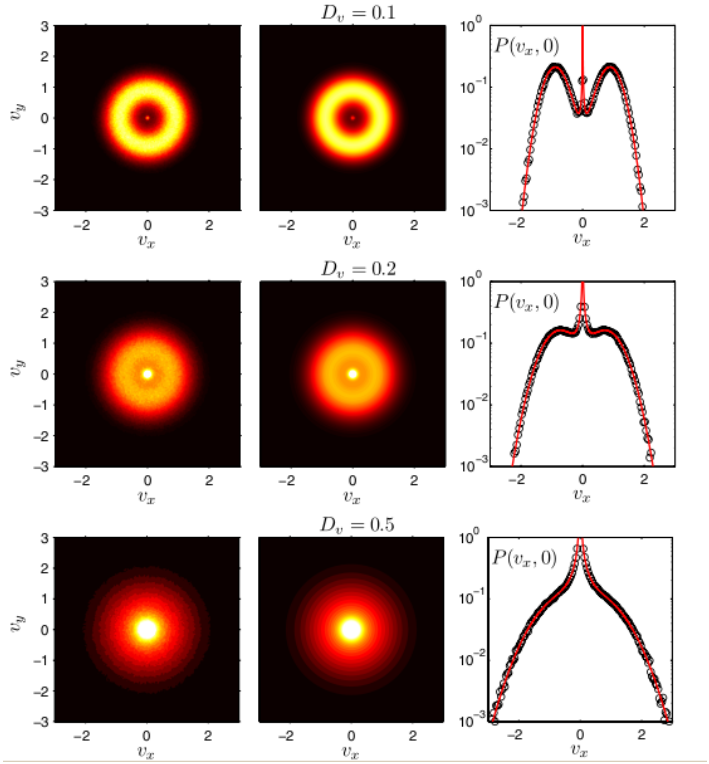


Figure 1.16: $P(v_x, v_y)$ as predicted by the model in [32], for different values of D_v, the intensity of the propelling force. This probability has the pathology of being divergent at $\vec{v} = 0$.

Chapter 2

Analytical solution of a process subject to Markovian dichotomous noise

In this chapter we will show the analytical solution of a one dimensional process subject to Markovian dichotomous noise. This will be useful in chapter 3, where we will make use of a noise which is a 2D generalization of Markovian dichotomous noise. The process subject to this generalized dichotomous noise is, to the best of our knowledge, not analytically solvable, so the 1D case will serve as an analytical point of reference.
As a first thing, in section 2.1 we will define Markovian dichotomous noise. Then in section 2.2 we will find, following [33], the stationary probability distribution of a process v_t subject to this kind of noise, $p_s(v_t)$. Then, in section 2.3 we will show that in a certain limit this noise tends to a Gaussian white noise. Finally, in section 2.4 we will explain the connection of the Markovian dichotomous noise with the 2D process that we will use in chapter 3.

2.1 Markovian dichotomous noise

Markovian dichotomous noise, or *random telegraph signal*, is a continuous time Markov process σ_t with a discrete state space $S = \{\Delta_+,\ \Delta_-\}$. For such a process the infinitesimal generator [1] has the form

$$\mathcal{L}f(\sigma) = w(\sigma'|\sigma)[f(\sigma') - f(\sigma)] \qquad \sigma, \sigma' \in \{\Delta_+, \Delta_-\},\ \sigma' \neq \sigma \tag{2.1}$$

where $w(i|j)$ is the transition rate from state j to state i. Let us call $\alpha := w(\Delta_-|\Delta_+)$ and $\beta := w(\Delta_+|\Delta_-)$.
This means that σ_t jumps from Δ_+ to Δ_- after an exponential time of parameter α, and from Δ_- to Δ_+ after an exponential time of parameter β.
Since there are only two state, the generator can be simply written as the matrix

$$\mathcal{L} = \begin{pmatrix} -\beta & \beta \\ \alpha & -\alpha \end{pmatrix} \tag{2.2}$$

Let $p_{i,j} = p(\sigma_t = i|\sigma_0 = j)$. Then, using the generator above, the resulting equation for the

[1] See appendix B for a brief introduction to infinitesimal generators for stochastic processes.

evolution of the probability is

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} p_{\Delta_-,j}(t) \\ p_{\Delta_+,j}(t) \end{pmatrix} = \begin{pmatrix} -\beta & \alpha \\ \beta & -\alpha \end{pmatrix}\begin{pmatrix} p_{\Delta_-,j}(t) \\ p_{\Delta_+,j}(t) \end{pmatrix} \tag{2.3}$$

Using $p_{\Delta_-,j}(t) + p_{\Delta_+,j}(t) = 1\ \forall t$, we get the time dependent solution

$$(p(t))_{i,j} = \frac{1}{\gamma}\begin{pmatrix} \alpha + \beta e^{-\gamma t} & \alpha(1-e^{-\gamma t}) \\ \beta(1-e^{-\gamma t}) & \beta + \alpha e^{-\gamma t} \end{pmatrix} \tag{2.4}$$

where $\gamma = \alpha + \beta$ and $i,\ j = \Delta_-,\ \Delta_+$.
Letting t go to infinity we find the stationary solution

$$p_s(\Delta_+) = \frac{\beta}{\gamma} \tag{2.5}$$

$$p_s(\Delta_-) = \frac{\alpha}{\gamma} \tag{2.6}$$

We can then compute the stationary mean of the process

$$\langle \sigma_t \rangle_s = \sum_{i=\Delta_+,\Delta_-} i p_s(i) = \frac{\beta}{\gamma}\Delta_+ + \frac{\alpha}{\gamma}\Delta_- \tag{2.7}$$

and the stationary autocorrelation

$$\langle \sigma_t \sigma_0 \rangle_s = \sum_{i,j=\Delta_+,\Delta_-} ij p_{ij}(t) p_s(j) - \langle \sigma_t \rangle_s^2 = \frac{\alpha\beta}{\gamma^2}(\Delta_+ - \Delta_-)^2 e^{-\gamma t} \tag{2.8}$$

This last result means that this process is non-white, meaning that it is not delta-correlated in time but has instead a characteristic time $1/\gamma$ over which it remains correlated with itself. If $\Delta_+ = -\Delta_- = a$ and $\alpha{=}\beta = \gamma/2$ the process is called symmetric. This will be the case in which we are interested.

2.2 Exact solution of a process subject to Markovian dichotomous noise

Let us consider a process v_t defined by the equation

$$\dot{v}_t = h(v_t) + \sigma_t g(v_t) \tag{2.9}$$

where σ_t is a Markovian dichotomous noise. This is one of the few cases of a process subject to coloured noise which is analytically solvable [33]. We are going to show in this section how to find its stationary probability distribution.
Since σ_t is coloured, the process v_t is non-Markovian[2]. But $(v_t,\ \sigma_t)$ is, and its infinitesimal generator is

$$\mathcal{L}f(v,\sigma) = (h(v) + \sigma g(v))\partial_v f(v,\sigma) + w(\sigma'|\sigma)[f(v,\sigma') - f(v,\sigma)] \tag{2.10}$$

[2]In fact, consider an increment of the process $\mathrm{d}v_t$: this depends on the value of the noise at time t, σ_t; since the noise has a certain autocorrelation time, its value at a later time t' will be correlated to that at time t, if $t'-t$ is shorter than the autocorrelation time. Consequently, the increment of the process at time t', $\mathrm{d}v_{t'}$, will be correlated to the increment at time t.

for $\sigma, \sigma' \in \{\Delta_-, \Delta_+\}$, $\sigma \neq \sigma$. So we can write a master equation for the joint probability of the form

$$\partial_t p(v,\sigma,t) = -\partial_v\{[h(v)+\sigma g(v)]p(v,\sigma,t)\} + w(\sigma|\sigma')p(v,\sigma',t) - w(\sigma'|\sigma)p(v,\sigma,t) \tag{2.11}$$

The process σ_t has only two possible states, so we can write this equation as the following system of two equations:

$$\begin{cases} \partial_t p(v,\Delta_-,t) = -\partial_v\{[h(v)+\Delta_- g(v)]p(v,\sigma,t)\} - \beta p(v,\Delta_-,t) + \alpha p(v,\Delta_+,t) \\ \\ \partial_t p(v,\Delta_+,t) = -\partial_v\{[h(v)+\Delta_+ g(v)]p(v,\sigma,t)\} - \alpha p(v,\Delta_+,t) + \beta p(v,\Delta_-,t) \end{cases} \tag{2.12}$$

Now, since we are interested in $p(v,t) = p(v,\Delta_+,t) + p(v,\Delta_-,t)$, we want to turn (2.12) into a closed equation for $p(v,t)$.
Let us introduce

$$q(v,t) = \alpha p(v,\Delta_+,t) - \beta p(v,\Delta_-,t) \tag{2.13}$$

Then $p(v,\Delta_+,t)$ and $p(v,\Delta_-,t)$ can be written as

$$\begin{cases} p(v,\Delta_+,t) = \gamma^{-1}(\beta p(v,t) + q(v,t)) \\ p(v,\Delta_-,t) = \gamma^{-1}(\alpha p(v,t) - q(v,t)) \end{cases} \tag{2.14}$$

Substituting these into (2.12) we get

$$\begin{cases} \partial_t(\alpha p(v,t) - q(v,t)) = -\{\partial_v[h(v)+\Delta_- g(v)](\alpha p(v,t) - q(v,t))\} + \gamma q(v,t) \\ \partial_t(\beta p(v,t) + q(v,t)) = -\{\partial_v[h(v)+\Delta_+ g(v)](\beta p(v,t) + q(v,t))\} - \gamma q(v,t) \end{cases} \tag{2.15}$$

Now substitute to the first equation the sum of the two, and find

$$\begin{cases} \gamma\partial_t p(v,t) = -\partial_v\{\gamma[h(v)+\langle\sigma\rangle g(v)]p(v,t) + (\Delta_- - \Delta_+)g(v)q(v,t)\} \\ \partial_t(\beta p(v,t) + q(v,t)) = -\{\partial_v[h(v)+\Delta_+ g(v)](\beta p(v,t) + q(v,t))\} - \gamma q(v,t) \end{cases} \tag{2.16}$$

where we used $\langle\sigma\rangle = \frac{\alpha}{\gamma}\Delta_+ + \frac{\beta}{\gamma}\Delta_-$.
To find the stationary solution, we set the time derivative equal to zero:

$$\begin{cases} 0 = -\partial_v\{\gamma[h(v)+\langle\sigma\rangle g(v)]p_s(v) + (\Delta_- - \Delta_+)g(v)q_s(v)\} \\ 0 = -\{\partial_v[h(v)+\Delta_+ g(v)](\beta p_s(v) + q_s(v))\} - \gamma q_s(v) \end{cases} \tag{2.17}$$

The first of these equation implies

$$\gamma[h(v)+\langle\sigma\rangle g(v)]p_s(v) + (\Delta_- - \Delta_+)g(v)q_s(v) = C \tag{2.18}$$

where C is a constant. Now, we can argue that C must be equal to zero in the following way. We require that the system is deterministically stable, i.e. that if we fix the noise σ in one of its possible value (let us call this value λ) then the process $v(t)$ does not explode ($|v(t)| \nrightarrow \infty$) when $t \to \infty$. For the equation

$$\dot{v}(t) = h(v) + \lambda g(v) \tag{2.19}$$

this condition is met if $\exists K(\lambda)$ such that

$$h(v) + \lambda g(v) < 0 \qquad \forall v > K(\lambda) \tag{2.20}$$
$$h(v) + \lambda g(v) > 0 \qquad \forall v < -K(\lambda) \tag{2.21}$$

In this case, if $v(0)$ is finite, v remains inside $[-K(\lambda),K(\lambda)]$ at every time.
This means, in turn, that $p_s(v,\lambda)$ has the compact support $[-K(\lambda),K(\lambda)]$. In our case, then

$$p_s(v,\Delta_-) = 0 \qquad \forall v > \max(K(\Delta_+), K(\Delta_-)) \tag{2.22}$$
$$p_s(x,\Delta_+) = 0 \qquad \forall x > \max(K(\Delta_+), K(\Delta_-)) \tag{2.23}$$

and the same for $v < min(-K(\Delta_+), -K(\Delta_-))$. Consequently, both $p_s(v)$ and $q_s(v)$ have a compact support $U(\Delta_+,\Delta_-)$, outside of which they are equal to zero. So, we can always choose v such that both $q_s(v) = 0$ and $p_s(v) = 0$, which by (2.18) implies that $C = 0$.
Then from (2.18) we get the following expression for $q_s(v)$

$$q_s(v) = -\frac{\gamma}{\Delta_+ - \Delta_-}[h(v) + \langle\sigma\rangle g(v)]\frac{p_s(v)}{g(v)} \tag{2.24}$$

which we substitute in the second of equations (2.17) to get for $p_s(v)$ the following differential equation:

$$\begin{aligned} p_s'(v) =& \Big[-\gamma\frac{h(v) + \langle\sigma\rangle g(v)}{[h(v) + \Delta_+ g(v)][h(v) + \Delta_- g(v)]} - \frac{h'(v) + \Delta_+ g'(v)}{h(v) + \Delta_+ g(v)} + \frac{1}{g(v)}\frac{h(v)g'(v) - g(v)h'(v)}{h(v) + \Delta_- g(v)}\Big]p_s(v) \\ =& \Big[-\gamma\frac{h(v) + \langle\sigma\rangle g(v)}{[h(v) + \Delta_+ g(v)][h(v) + \Delta_- g(v)]} - \partial_v \ln\Big(\frac{[h(v) + \Delta_- g(v)][h(v) + \Delta_+ g(v)]}{g(v)}\Big)\Big]p_s(v) \end{aligned} \tag{2.25}$$

where the prime denotes derivation with respect to v.
By integration we finally obtain

$$p_s(v) = N\frac{g(v)}{[h(v) + \Delta_- g(v)][h(v) + \Delta_+ g(v)]}e^{-\gamma\int^v \mathrm{d}z \frac{h(z)+\langle\sigma\rangle g(z)}{[h(z)+\Delta_- g(z)][h(z)+\Delta_+ g(z)]}} \tag{2.26}$$

with N a normalization constant.
A general discussion on the behaviour of this solution at the boundaries of its support $U(\Delta_+,\Delta_-)$ is done in [33]. It is found that different behaviours are possible depending on the value of γ and on the values of the functions h and g at the boundaries. Here we will not undertake this general discussion, but proceed to examine the particular case that will be useful in next chapter.

The process we consider from now on is the one with $h(v) = -\rho v$ and $g(v) = 1$, subject to a symmetric dichotomous noise σ_t. For the symmetric case, the generator has the simple expression

$$\mathcal{L}f(\sigma) = \frac{1}{\tau}[f(-\sigma) - f(\sigma)] \tag{2.27}$$

where σ takes values in $\{a,-a\}$, and $w(a|-a) = w(-a|a) = \frac{1}{\tau} = \frac{\gamma}{2}$.
Thus equation (2.9) becomes

$$\dot{v}_t = -\rho v_t + \sigma_t \tag{2.28}$$

We can verify that in this case conditions (2.22) are satisfied, and the process is confined in the interval $[\,-\frac{a}{\rho},\,\frac{a}{\rho}]$.
Noting that in the symmetric case $\langle\sigma\rangle = 0$, the stationary solution (2.26) becomes

$$p_s(v) = N\frac{1}{\rho^2 v^2 - a^2}e^{\frac{2\rho}{\tau}\int^v \mathrm{d}z \frac{z}{\rho^2 z^2 - a^2}} = N\Big[1 - \Big(\frac{\rho v}{a}\Big)^2\Big]^{\frac{1}{\rho\tau}-1} \tag{2.29}$$

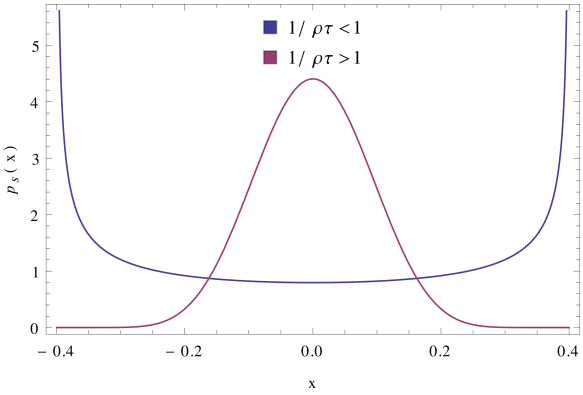


Figure 2.1: The solution (2.29) in the two regimes.

The support of $p_s(v)$ is $[\ -\frac{a}{\rho},\ \frac{a}{\rho}]$, in fact outside of this interval it would assume negative values. The behaviour of $p_s(v)$ depends on the relative value of parameters ρ and γ as follows:

- if $\frac{1}{\rho\tau} > 1$ the extremants are $-\frac{a}{\rho}$, 0, $-\frac{a}{\rho}$. The first and the last are minima, while 0 is a maximum (see figure 2.1);
- if $\frac{1}{\rho\tau} = 1$ the solution is constant on its support;
- if $\frac{1}{\rho\tau} < 1$ in $\pm\frac{a}{\rho}$ the solution diverges, while the extremant in 0 is a minimum (see figure 2.1).

We see then that the behaviour of the solution depends not on the intensity of noise a, but on τ which by eq. (2.8) is the autocorrelation time of the noise.
It is interesting to note that in the limit $\tau \to 0$, $a \to \infty$ such that $\frac{\tau a^2}{2} = \text{const} =: D$, the solution tends to a Gaussian:

$$\lim_{\substack{\tau\to 0\\ a\to\infty}} p_s(v) = \lim_{\substack{\tau\to 0\\ a\to\infty}} N\Big(1 - \frac{\rho v^2}{\tau a^2}\rho\tau\Big)^{\frac{1}{\rho\tau}-1} = Ne^{-\frac{\rho v^2}{2D}} \tag{2.30}$$

This is a consequence of the fact that the symmetric Markovian dichotomous noise in the above mentioned limit tends to a Gaussian white noise, as will be shown in next section.

2.3 Gaussian limit of Markovian dichotomous process

In order to prove that the symmetric Markovian dichotomous noise tends to a Gaussian white noise in the limit

$$\begin{cases} \tau \to 0 \\ a \to \infty \end{cases} \quad \text{such that} \quad D := \frac{a^2\tau}{2} = \text{const} \tag{2.31}$$

we are going to show, following [34], that its integral process x_t converges in distribution to the integral process of the Gaussian white noise, i.e. the Wiener process. In particular, we will see that in the above limit the master equation for the process x_t becomes the diffusion equation

$$\frac{\partial}{\partial t}p(x,t) = D\ \frac{\partial^2}{\partial x^2}p(x,t) \tag{2.32}$$

which is exactly the Fokker-Planck equation of the Wiener process [30].
The integral process x_t is defined by the equation

$$\dot{x}_t = \sigma_t \tag{2.33}$$

where $\sigma_t \in \{a, -a\}$.
This is a particular case of equation (2.9) where we have chosen $h(x) = 0$, $g(x) = 1$. Thus, in analogy with what was done in the previous section, we may write for $p(x, a)$ and $p(x, -a)$ the following system of equations:

$$\begin{cases} \partial_t p(x,a,t) = -a\partial_x p(x,a,t) + \frac{1}{\tau}[p(x,-a,t) - p(x,a,t)] \\ \partial_t p(x,-a,t) = a\partial_x p(x,-a,t) + \frac{1}{\tau}[p(x,a,t) - p(x,-a,t)] \end{cases} \tag{2.34}$$

Now we substitute to the second of eqs. (2.34) the sum of the two, and in the first we use $p(x, -a, t) = p(x, t) - p(x, a, t)$.

$$\begin{cases} \partial_t p(x,a,t) = -[a\partial_x - \frac{2}{\tau}]p(x,a,t) + \frac{1}{\tau}p(x,t) \\ \partial_t p(x,t) = a\partial_x p(x,t) - 2a\partial_x p(x,a,t) \end{cases} \tag{2.35}$$

From the first, taking as initial conditions $p(x, a, 0) = 0$, we obtain

$$p(x,a,t) = \int_0^t \mathrm{d}s\ e^{-[a\partial_x - \frac{2}{\tau}](t-s)} \frac{1}{\tau} p(x,s) \tag{2.36}$$

which substituted in the second gives us the following differential equation for $p(x, t)$:

$$\partial_t p(x,t) = a\frac{\partial}{\partial x}p(x,a,t) - 2a\frac{\partial}{\partial_x}\int_0^t \mathrm{d}s\ e^{-[a\partial_x - \frac{2}{\tau}](t-s)} \frac{1}{\tau} p(x,s) \tag{2.37}$$

It is interesting to note that this equation shows explicitly that x_t is not Markovian: in fact its evolution depends, through the integral, on the value of the process at all preceding times. In the limit (2.31), the argument of the exponential tends to $-\infty$, so the dominant term in the integral will be that for $s \approx t$. Then, assuming that $p(x, t)$ is regular enough, i.e. that its derivatives do not diverge too fast, we may make the approximation of calculating it in t, and thus we have:

$$\partial_t p(x,t) = a\frac{\partial}{\partial x}p(x,a,t) - 2a\frac{\partial}{\partial_x}\int_0^t \mathrm{d}s\ e^{-[a\partial_x - \frac{2}{\tau}](t-s)} \frac{1}{\tau} p(x,t) \tag{2.38}$$

Computing the integral

$$\partial_t p(x,t) = a\frac{\partial}{\partial x}p(x,a,t) - 2a\frac{\partial}{\partial_x}\frac{1}{a\frac{\partial}{\partial x} + \frac{2}{\tau}} \frac{1}{\tau} p(x,t) \tag{2.39}$$

where we have used the fact that $e^{-[a\partial_x - \frac{2}{\tau}]t}$ tends to zero in the limit. Now, recalling that $\tau a \to 0$, we get

$$\begin{aligned} \partial_t p(x,t) &= a\frac{\partial}{\partial x}p(x,a,t) - a\frac{\partial}{\partial x}\frac{1}{\frac{\tau a}{2}\frac{\partial}{\partial x} + 1}p(x,t) \simeq a\frac{\partial}{\partial x}p(x,a,t) - a\frac{\partial}{\partial x}(1 - \frac{\tau a}{2}\frac{\partial}{\partial x})p(x,t) \\ &= D\frac{\partial^2}{\partial x^2}p(x,t) \end{aligned} \tag{2.40}$$

which is eq. (2.32).

2.4 Generalizations of the Markovian dichotomous noise

A possible generalization of the Markovian dichotomous noise is what is sometimes called *compound* Markovian dichotomous noise [34]. Loosely speaking, the idea at the base of the generalization is the following: the process has two possible states, + and −, but the value of the process in a certain state is not a fixed value but a random variable with a certain distribution. Thus, when the process is in state +, its value is a variable Δ_+ with distribution $\rho_+(\Delta_+)$, and similarly in the state − it is a variable Δ_- with distribution $\rho_-(\Delta_-)$. These distributions could be, for instance, Gaussians centred respectively in a and $-a$. In this case the noise, instead of jumping between two values Δ_+ and Δ_- as in the analytically solvable case above, jumps from a value gaussianly distributed around a to a value gaussianly distributed around $-a$. In the limit in which each of the two distributions tends to a δ-function on one value, this compound process reduces to the simple process.

In chapter 3 we will make use of a noise which is a "double" generalization of the one discussed in this chapter.
The symmetric Markovian dichotomous process has only two states, so we can view it as jumping between the point $+a$ and $-a$ of the x axis. Now consider a process which states are the points of the circle of radius a in the plane, thus having a vectorial value $\vec{\sigma}$. As in the former case, the rate of transition from any state to any other is $1/\tau$. Then, this process has infinite equiprobable state.
Now let us take a process $\vec{x}_t$ subject to this two-dimensional noise, with an equation of the same form of (2.28). Although to our knowledge this equation is not analytically solvable, we may deduce the behaviour of $p_s(\vec{v})$ with the following reasoning. For the one dimensional process, each time the noise changes value, the process v_t relaxes to its limit value σ_t/ρ with a characteristic time of relaxation $1/\rho$. Depending on whether this time of relaxation is shorter or longer than τ, the process is able or not to reach the limit value between two jumps of the noise, consequently we get the two regimes. One where the probability of v is concentrated around the limit values, and one where it is concentrated at small values. In the generalised process, $p_s(v)$ is built in a similar way: what changes is that the process is pulled not to two possible limit values, but to infinite ones, all equiprobable. What we conclude from this reasoning is that there will be again two regimes, one where the probability will be concentrated near the circle of radius a/ρ, and one where it will be concentrated at the origin. These assertions are supported by numerical simulations of the process, as presented in next chapter.
Then we take an additional generalization: when the process is in one state, which corresponds to choosing an angle θ for the vector $\vec{\sigma}_t$, its modulus is not a but $\sigma \sim \mathrm{N}(a, \xi)$, as in the compound process discussed above. In the limit in which the variance of the Gaussian tends to zero, this process reduces to the non compound one. The solution of a process subject to this noise will be investigated in next chapter with the help of numerical simulations.

Chapter 3

Our model

In this chapter we introduce a new stochastic model [4] to describe the motion of self-propelled particles, such as bacteria or protists, in a viscous fluid in the absence of external stimuli. The proposed model is presented in section 3.2.
To verify the efficacy of this model, in section 3.3 we compare its predictions with data relative to a protist species, *Colpidium* sp. Information on this microorganism and on the data is gathered in section 3.1.

As explained in chapter 1, microorganisms move through various mechanisms, mainly employing cilia and flagella, but it is not our intent to include such specificity in the modelling, which should remain as general as possible. The hope is that the model will be apt to describe qualitatively the motion of a large class of organisms, independently from the particular means of propulsion they might use.
The reason why we expect a qualitative description is intrinsic to our approach: since the various means of propulsion influence in a certain way the motion (as clear from chapter 1), to finely reproduce the trajectories one should keep in consideration every detail. In this way, however, one renounces to generality and arrives to a model valid for a certain particular microorganism, that fails to enlighten about the common basis of microscopic motion. On the other hand, if there are some common statistical features characterizing the motion of microorganisms, to understand their emergence a general description is needed. Such a general description should provide a theoretical basis on which to build, if necessary, a further more specific model and possibly also a theory of directed motion.
One of these common statistical features we are looking for is non-Gaussian probability distributions for velocities, which has been confirmed by many experiments, as explained in section 1.2. Another characteristic of the motion which is well established by observation is persistence, i.e. the fact that the direction tends to remain the same for a certain persistence time. These are then two elements that a model seeking a general description of microorganisms motility should reproduce.
The model we propose in this chapter accounts for these two features by proposing a new form of the propelling force term in the equation of the process, different from the Gaussian white noises employed by previous models (see section 1.3). This proposal is motivated not only by the need of describing a non-Gaussian probability distribution for velocity, but also by the request of a certain realism in the modelling of the propelling force. As explained at the conclusion of section 1.3, a Gaussian white noise does not seem a lifelike description of this force, due to the fact that it changes its value at every time-step without any self-correlation: this is not compatible with the observed persistence of the motion. Therefore the need of a different term for the propelling force.

We find there is a good agreement between the data and the predictions of the proposed model, and in particular that the model fits accurately the experimental $P(v)$.

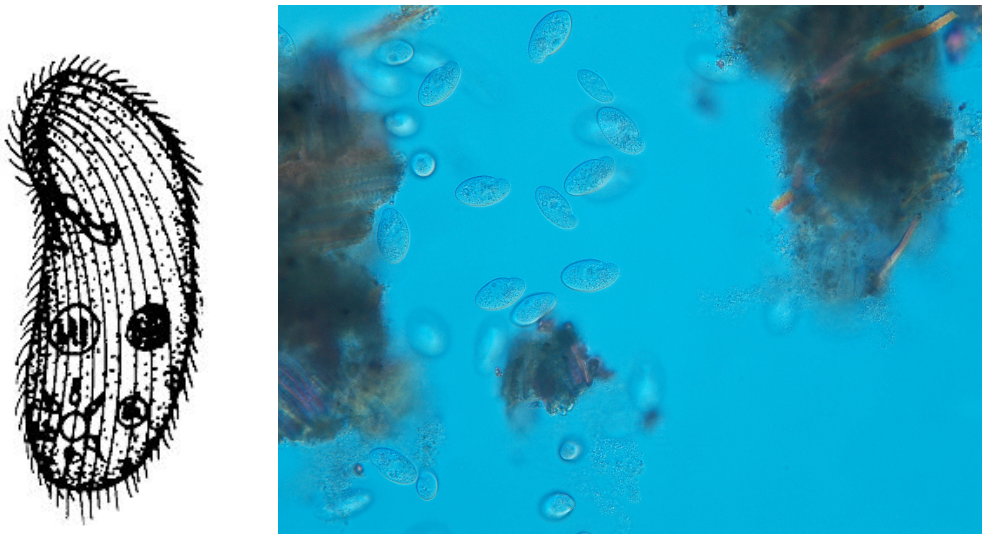
3.1 Data

The dataset consists of trajectories of microorganisms of the protist species *Colpidium* sp. (see fig. 3.1), swimming in a viscous solution.
The study organism was acquired at Carolina Biological Supply (USA). The culture medium was made of local spring water and Protozoan Pellets (Carolina Biological Supply) at a density of 0.45 g/L. Three bacteria species (*Serratia fonticola*, *Breviacillus brevis*, *Bacillus subtilis*) were added to the culture and acted as a food source for *Colpidium* sp. The culture of *Colpidium* sp. was initialized two weeks before the experiment and kept under fluorescent light at a constant temperature of 22 °C. The experiment was performed by introducing 1 mL of *Colpidium* sp. in a Sedgewick Rafter Counting Cell S52 (Pyser - SGI, UK) made of glass. The counting cell was placed under the objective of a stereomicroscope and a video was recorded at a framerate of 0.15 s^{-1}; the visible area in the video was $11,0x8,3$ mm. The experiment was repeated three times by extracting 1 mL of culture for each measurement. Videos were analysed with the software *Mathematica*, version 9.0, which was used to extract the position of each individual at all times. The plugin *MOSAIC* [35] for the software *ImageJ* was used to reconstruct trajectories. A total of 1035 trajectories were recorded. Organisms are tracked until they remain in the observation window, typically between 30 and 60s, consequently each trajectory consists of 200 to 400 points. Microbial motion in the experiment can be considered two dimensional, being the vertical component of the motion negligible, as results from a visual inspection.

Colpidium is a ciliate, which means that it moves employing cilia, which are organized in rows covering all its body. The size of the organisms may vary depending on the living conditions, but the usual measures are 60 to 90 μm long and 20 to 30μm wide. The specimens employed in this experiment were measured to have a mean radius of 14 μm.
Cilia are 10 μm long, arranged in around 30 rows. Each row beats in a synchronized way and there is a small time lag between the beating of successive rows. The beating, called *metachronal wave*, starts from the anterior part of the cell and sweeps to the posterior end. Since successive rows have slightly different axes, the resulting thrust causes the microorganism to advance spiralling [36].

Examples of the trajectories obtained are in figure 3.2. These trajectories display a directional persistence of 10-20 s, alternating runs in a given direction at an about constant speed to turns, with noise due to the spiralling pattern. This motion is similar to the *run and tumble* of which we wrote in section 1.2, although in the case of *Colpidium* tumbles are instantaneous, differently from what is observed in other species.
We computed the mean velocity between two measures of the position as $\vec{v}(t) = \frac{\vec{x}(t+\Delta t)-\vec{x}(t)}{\Delta t}$ with Δt=0.45 s. This is approximately the duration of one spiral, and was chosen in order not to follow too closely the spiraling pattern. In fact, since our model aims at describing the motion of a large class of organisms, it is deemed reasonable that we overlook these specific details.
Note that the velocity distribution depends on the size of Δt: in the limit of a very large Δt[1] $P(v)$ will be peaked at the origin. In fact the measured velocity will be an average of

[1]Where very large means much greater than the typical duration of a straight segment of the trajectories.



(a) **(b)**

Figure 3.1: (a) Portrayal of a *Colpidium*; (b) A photo of *Colpidia* in the culture medium.

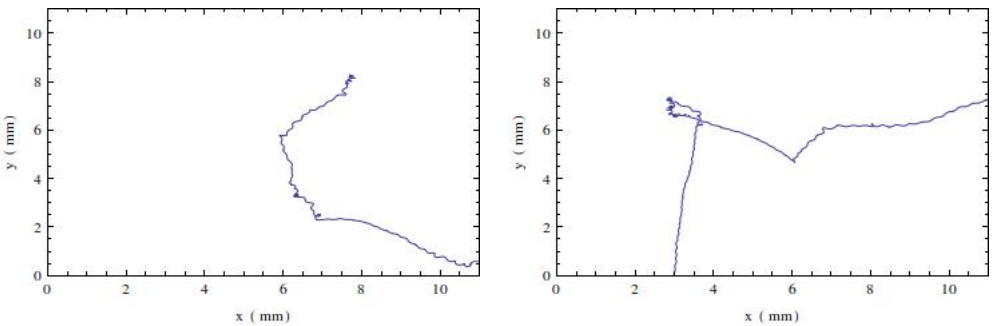


Figure 3.2: Two examples of reconstructed trajectories of *Colpidium* individuals.

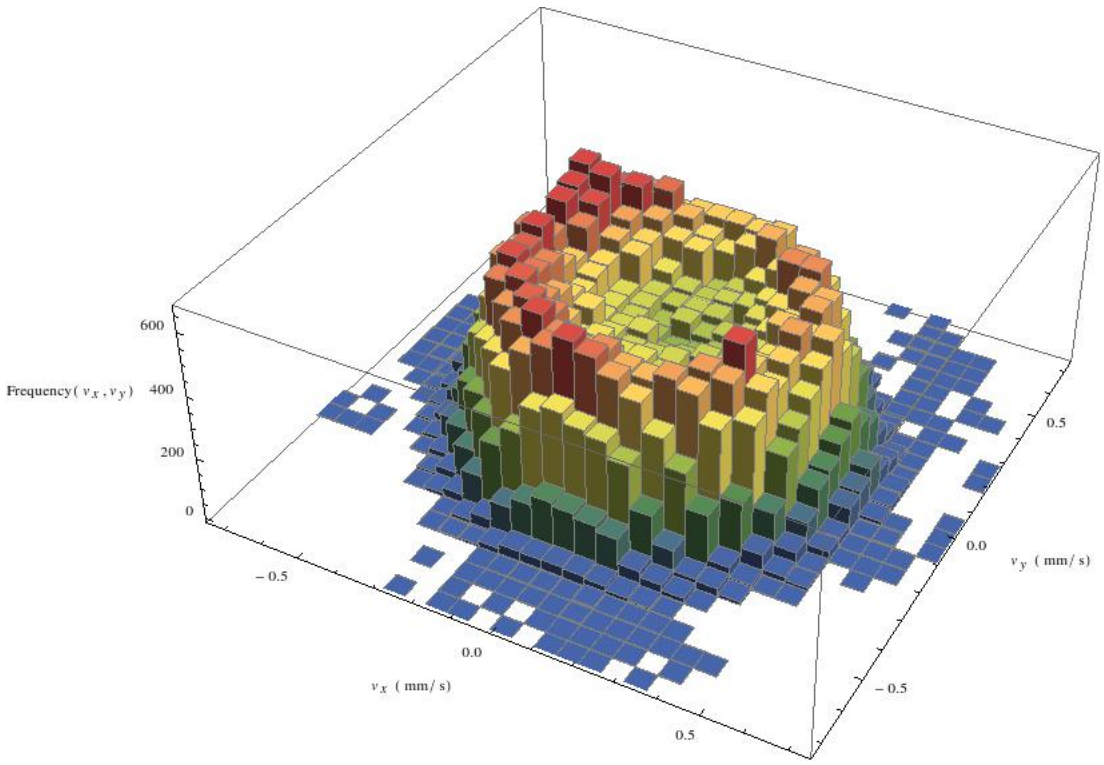


Figure 3.3: Velocity distribution from data of Colpidium trajectories.

velocities of segments of trajectories with different orientation. Since these orientation are isotropic, the average velocity will be close to zero.
The distribution $P(v)$ exhibit a volcano-like shape (see figure 3.3) peaked at speeds of approximately 0.4mm/s. This shape cannot be predicted by a model which employs Gaussian white noise as a propelling force, as that would predict a Gaussian distribution, nor by any of the models reviewed in section 1.3. The distribution appears rotationally invariant, as expected from the lack of any directional stimulus in the environment. Thus, we restrict our investigation to the distribution of moduli $P(v)$ (fig. 3.4 (a)) instead of the full two dimensional distribution.

From the distribution of moduli we obtained a slice of the rotational-invariant distribution $P(v_x, v_y)$, which allows visual inspection of the volcano-like shape and facilitates comparison between the model and the data (figure 3.4(b)). In fact, using isotropy, $P(v) = 2\pi P(v, \theta) = 2\pi v P(v_x, v_y)$ being v the determinant of the Jacobian of the transformation of variables. Then $P(v_x, 0) = \frac{P(|v_x|)}{2\pi|v_x|}$.
The velocity autocorrelation

$$C(t) = \frac{\langle \vec{v}(t) \cdot \vec{v}(0) \rangle}{\langle \vec{v}(0)^2 \rangle} \tag{3.1}$$

was also computed from the experimental data and is shown in figure 3.4 (c).
Furthermore, we also observed the Mean Square Displacement (MSD)

$$MSD(t) = \langle (x(\vec{t}) - x(\vec{0}))^2 \rangle \tag{3.2}$$

shown in figure 3.4(d). The curve saturates for t$\approx$ 25 s at a value MSD $\approx$ 35 mm^2. This is due to the finite dimensions of the observation window (11mm x 8mm), which makes it impossible to follow the whole trajectory of the microorganisms. The microorganisms which move more in terms of square displacement will exit the window first, so for longer times

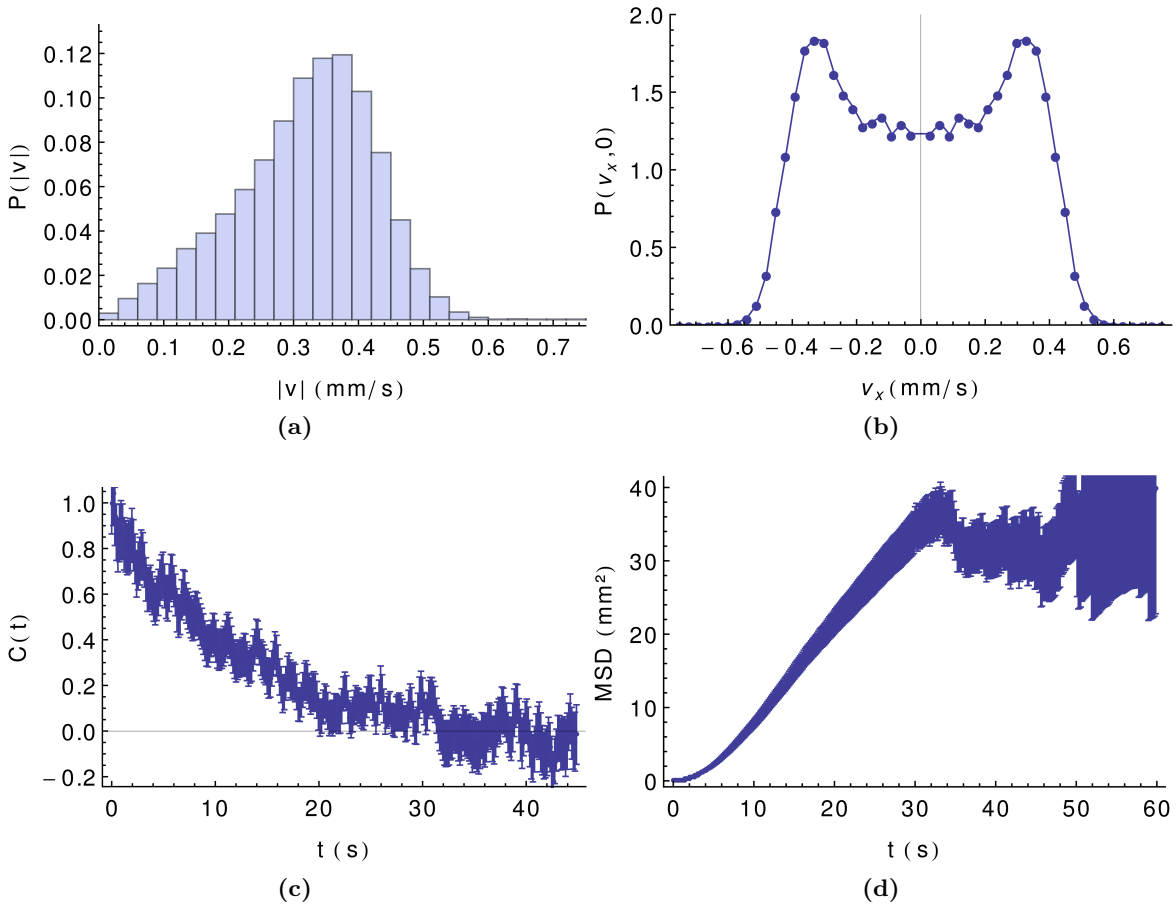


Figure 3.4: Experimental data for: (a) velocity's moduli distribution; (b) velocity's distribution $P(v_x, 0)$; (c) velocity autocorrelation; (d) MSD.

the statistic will comprise only the microorganisms remained closer to their starting points, which lower the MSD and cause its saturation. That this happens at about 25s is plausible, being this approximately the time lapse required for an organism moving on a straight line at the most frequent speed 0.4 mm/s (see fig. 3.4(a)) to cross the observation window: this means that after that time the organisms giving the higher contributions to the MSD will begin to exit. The consequence is that the MSD does not give us meaningful information about the motion, because we cannot follow the singular trajectory for enough time to tell how it behaves in the long term. The definitive trend of the MSD is in fact reached after some persistence times have passed, but our tracking lasts only two or three. The same goes for the other moments of the position, like kurtosis, which show the same saturation. None of these, then, are useful to characterize the motion. Thus, we were only be able to judge the model on its prediction regarding the velocity process (comparing $P(v)$ and $C(t)$) and not on its description of the spatial feature on the motion.

3.2 The model

The particles our model seeks to describe move in a two dimensional landscape and are subject to the force they exert to propel themselves. Since as explained in section 1.3 the best frame-

work to describe a process that shows persistence, as that of the velocity of microorganisms, is the O-U process, we write a generalization of the O-U equation as follows:

$$\dot{\vec{v}} = -\rho\vec{v} + \vec{\sigma} \tag{3.3}$$

where ρ is a parameter of the model and $\vec{\sigma}$ is a suitable stochastic process modeling the self-propulsion force. The proposed form of this process is based on the observation of microbial trajectories: in this sense the model is "phenomenological". If we ignore the spiralling pattern which is a feature connected to the particular motility mechanism of *Colpidium*, the microorganisms' trajectories alternate segments of directed movement to neat changes of direction. Recalling that at low Reynolds numbers a change of the propelling force results in an immediate change of direction due to the negligibility of inertia, we may suppose that during the straight segments the force remains the same. Consequently, our hypothesis for the force is a Markovian process $\vec{\sigma_t} = \sigma_t(\cos\theta_t, \sin\theta_t)$ with infinitesimal generator (see appendix for a derivation)

$$\mathcal{L}f(\theta,\sigma) = \frac{1}{2\pi\tau}\int_0^{2\pi} d\theta' \int_{-\infty}^{+\infty} d\sigma' \frac{1}{\sqrt{2\pi\xi^2}} e^{-\frac{(\sigma'-a)^2}{2\xi^2}} [f(\theta',\sigma') - f(\theta,\sigma)] \tag{3.4}$$

that is, jumps are exponentially distributed with rate $1/\tau$ and each jump is associated to the choice of both a new direction (an angle θ uniformly distributed in $[0, 2\pi]$) and an updated modulus for the propelling force, σ, where $\sigma \sim$N(a,ξ). The average time of permanence of the process in a certain state is τ.
Note that this generator allows σ to take negative values, which are not acceptable if it is the modulus of a vector. Consequently, in principle we should take instead a σ distributed as a truncated Gaussian, assuming only positive values. However, as we will see in next section, the values of a and ξ that will be of interest to us are such that the probability of negative value is negligible, so this issue will not constitute a concrete problem.
To summarize, the stochastic force exhibits no directional preference (because we are modeling movement in the absence of external stimuli) and has a modulus that fluctuates around a non-zero mean.
This process is a generalization of the Markovian dichotomous noise studied in chapter 2, in the sense explained in section 2.4.
We can characterize the process $\vec{\sigma}_t$ with its mean and variance, using

$$P(\vec{\sigma}_{t+t'} \in [\vec{\sigma}', \vec{\sigma}' + \mathrm{d}\vec{\sigma}'] | \vec{\sigma}_t \in [\vec{\sigma}, \vec{\sigma} + \mathrm{d}\vec{\sigma}]) = \frac{1}{2\pi}\frac{1}{\sqrt{2\pi\xi^2}} e^{-\frac{(\sigma'-a)^2}{2\xi^2}} (1 - e^{-t'/\tau}) + \delta(\vec{\sigma}' - \vec{\sigma}) e^{-t'/\tau} \mathrm{d}\sigma' \tag{3.5}$$

and

$$P_s(\vec{\sigma}) = \frac{1}{2\pi}\frac{1}{\sqrt{2\pi\xi^2}} e^{-\frac{(\sigma-a)^2}{2\xi^2}} \tag{3.6}$$

which simply comes from 3.5 letting t' go to infinity.
The mean of the stochastic force at stationarity is then

$$\langle\vec{\sigma}\rangle = \int \mathrm{d}\vec{\sigma}\ \vec{\sigma}\ P_s(\vec{\sigma}) = 0 \tag{3.7}$$

and its autocorrelation

$$C(s) = \frac{\langle\vec{\sigma}(t+s)\cdot\vec{\sigma(t)}\rangle}{\langle\sigma(t)^2\rangle} = \frac{1}{\xi^2}\int \mathrm{d}\vec{\sigma} \int \mathrm{d}\vec{\sigma}'\vec{\sigma}\cdot\vec{\sigma}' P(\vec{\sigma}', t+s|\vec{\sigma}, t) P_s(\vec{\sigma}) = e^{-s/\tau} \tag{3.8}$$

Having established the properties of the stochastic force, eq. (3.3) defines the velocity process, $\vec{v}_t$. We may picture this process in the following way: if at time t the stochastic force has a value $\vec{\sigma}_t$, the velocity will approach the limit velocity $\vec{v}_l = \vec{\sigma}_t/\rho$ with a characteristic time $\tau_c = 1/\rho$. With probability $e^{-t'/\tau}dt'$ a time in the interval $(t',\ t' + dt']$ is chosen and a new value of the noise, $\vec{\sigma}_{t+t'}$ is drawn and the velocity will be pulled towards a new limit velocity. The process ($\vec{v}_t$,θ_t,σ_t) is Markovian, and the equation for the stationary joint probability P($\vec{v}$,θ,σ) can be written as

$$\mathcal{L}^* P_s(\vec{v}, \theta, \sigma) = 0 \tag{3.9}$$

where $\mathcal{L}^*$ is the adjoint of the infinitesimal generator of the process (v_t,θ_t,σ_t). Inserting the expression for $\mathcal{L}$, the derivation of which can be found in appendix, we find

$$-\nabla_{\vec{v}}[(-\rho\vec{v} + \vec{\sigma})p(\vec{v}, \theta, \sigma)] + \frac{1}{2\pi\tau}\int_0^{2\pi} d\theta' \int_{-\infty}^{+\infty} d\sigma' \frac{1}{\sqrt{2\pi}\xi} e^{-\frac{(\sigma'-a)^2}{2\xi^2}}[p(\vec{v}, \theta', \sigma') - p(\vec{v}, \theta, \sigma)] = 0 \tag{3.10}$$

Equation (3.10) is, to the best of our knowledge, not solvable analytically. Instead, we can study the stationary velocity distribution for the process $\vec{v}_t$, $P_s(\vec{v})$, with the help of numerical simulations based on the Gillespie algorithm [37].
The equation of motion for the two components of the velocity are

$$\dot{v}_x = -\rho v_x + \sigma\cos\theta \tag{3.11}$$

$$\dot{v}_y = -\rho v_y + \sigma\sin\theta \tag{3.12}$$

Between two successive updates of the process, v_x and v_y are updated using the analytical solution of eqs. (3.11) and (3.12) with the current value of the force:

$$v_x(t) = e^{-v_x t}(v_x(0) - \sigma\cos\theta) + \sigma\cos\theta \tag{3.13}$$

$$v_y(t) = e^{-v_y t}(v_y(0) - \sigma\sin\theta) + \sigma\sin\theta \tag{3.14}$$

The timing of the changes of the force are extracted from an exponential distribution of rate $1/\tau$. The values of $\vec{v}$ are then sampled every Δt. Ergodic properties of the system ensure that the histogram of such velocities, after waiting an initial transient in order to reach stationarity, is close to the theoretical stationary distribution $P(\vec{v})$ [30].
Simulations show that $P(\vec{v})$ is strongly affected by the relative values of ρ and τ and exhibits, in analogy with what was shown in chapter 2 for the one dimensional process, a crossover between two regimes (see figure 3.5):

1. if $\rho\tau > 1$, the limit velocity $\vec{v}_l = \frac{\vec{\sigma}}{\rho}$ is reached before the stochastic force changes, so $P(\vec{v})$ is concentrated near $\vec{v}_l$ s and displays the volcano-like shape observed experimentally;

2. if $\rho\tau < 1$, the characteristic time of direction update is too small to allow complete relaxation, so $P(\vec{v})$ is peaked in $\vec{v} = 0$.

Depending on the study species, regime 1) or 2) might be of interest. In the case of *Colpidium* the distribution of velocity has a shape similar to that of regime 1) and as shown in next section it is not only qualitatively similar: the model provides a quantitatively satisfactory fit of the experimental distribution.

With respect to the previous models explored in section 1.3, this new model conserves certain positive features achieved by those previous models and moreover brings the significant novelty of a force which is non-Gaussian and non-white.
The forces along the two axes are not independent, as they are the projections of a unique

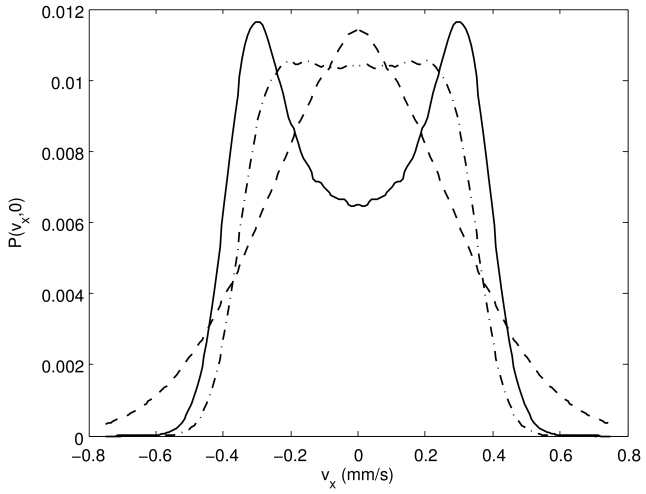


Figure 3.5: $P(\vec{v})$ in the two different regimes of the model and at the crossover between the two regimes: $\rho\tau > 1$ (*solid line*), $\rho\tau = 1$ (*dash-dot line*), $\rho\tau < 1$ (*dashed line*).

vectorial force $\vec{\sigma}_t$. In previous models, this was obtained by choosing the microorganisms' frame of reference [22, 32], and thus describing the motion in the coordinates (v, ϕ), where v is the modulus of velocity and θ is the direction of motion. These approaches led however to other problems: in one case the dynamic of v and θ were independent (while experimental result suggest otherwise), in another the resulting $P(\vec{v})$ was non-integrable at the origin.
The model we propose, on the other hand, contains both the non-independence of the force on the two axes and the non-independence of the dynamic of speed and direction. Moreover it introduces a form of the stochastic propelling force which, as argued at the beginning of this chapter, is more realistic in describing the process than a white noise.

3.3 Comparison between data and model

As mentioned above, our model predicts a volcano-shaped velocity distribution P(v) in certain parameter ranges. This allows to perform a quantitative comparison between model and data and to find the set of parameters that best describes the experimentally observed motion. The comparison of the velocity distribution between model and data, however, only allows fitting three of the four parameters (ρ, τ, a, and ξ) of the model. In fact, if we divide eq. (3.3) by a constant c and redefine $t' = ct$, we get the same equation with $a \to a/c$, $\xi \to \xi/c$, $\rho \to \rho/c$, $\tau \to c\tau$. Since the stationary distribution is not affected by the redefinition of time, it will be left unchanged. Consequently, choosing for example $c = \rho$, we get that $P_s(\vec{v})$ depends on a/ρ, ξ/ρ, $\tau\rho$.
To fit all four parameters, we first fit a/ρ, ξ/ρ, $\tau\rho$ to the experimental velocity distribution and subsequently fit ρ to the experimental velocity autocorrelation. The fitting of the velocity distribution was performed via the MCMC algorithm DREAM [38], which allows computing the posterior distribution of the parameters given the data. Fig. 3.6 shows such distribution (obtained marginalizing the stationary distribution of the MCMC), while table 3.1 shows the parameter values that give the maximum likelihood fit. Errors were estimated as the width of the posterior distribution at half height.

The resulting best fit for the velocity distribution is shown in figure 3.8(a). The best fit for the autocorrelation was found, instead, by minimization of the χ^2 with a bootstrap

method (in figure 3.7 the posterior distribution of the fitted parameter ρ). Figure 3.8(b) shows the best fit of the autocorrelation curve and table 3.2 reports the estimated values of the parameters. The plot with logarithmic scale shows that the model's autocorrelation follows the experimental one in an exponential decay.

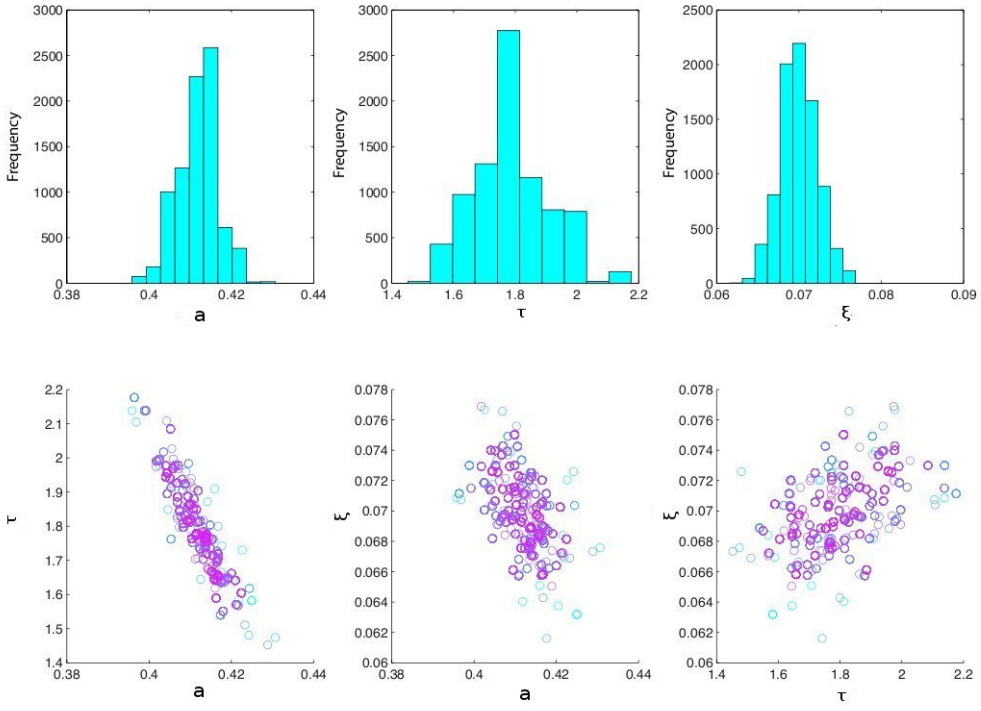


Figure 3.6: Top: Posterior marginal probability distributions for the parameters, ensuing from the MCMC algorithm, with ρ set to $\rho = 1$. Bottom: Scatter plots of the parameters explored after reaching stationarity. The color indicates the log-likelihood for each choice of the parameters (light blue = low, purple = high).

a/ρ	0.41 ± 0.01 m/s
ξ/ρ	0.070 ± 0.005 m/s
$\tau\rho$	1.7 ± 0.1

Table 3.1: Best fit parameters

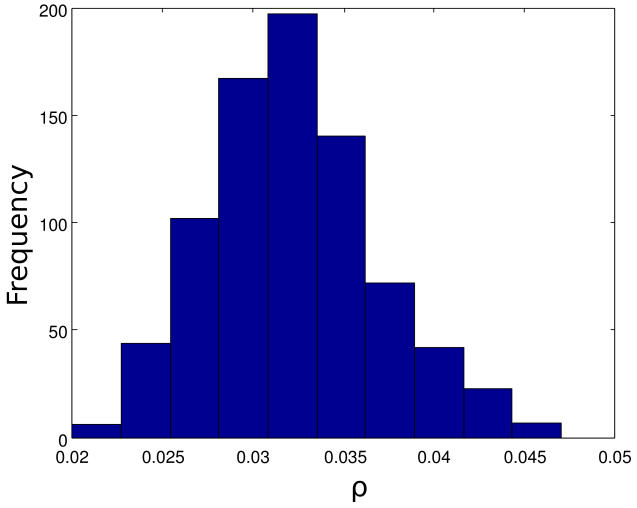
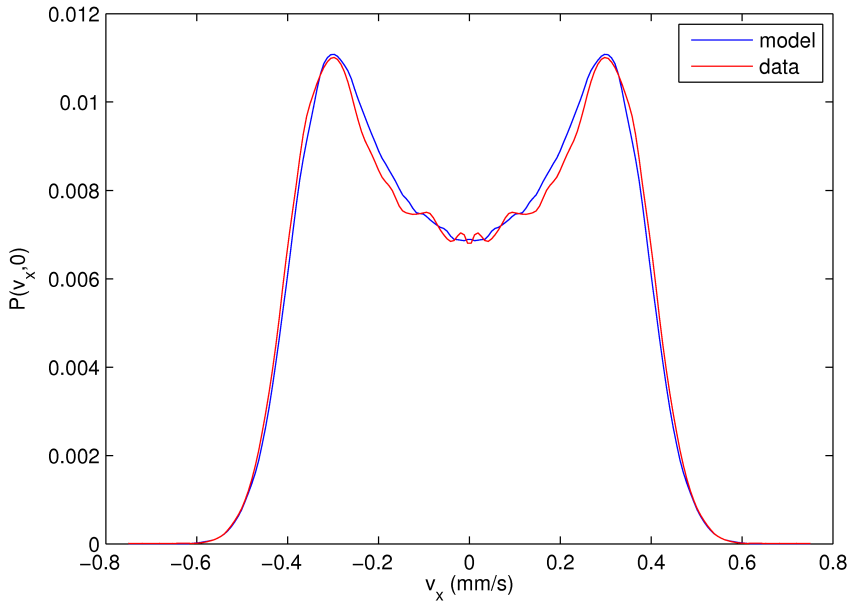


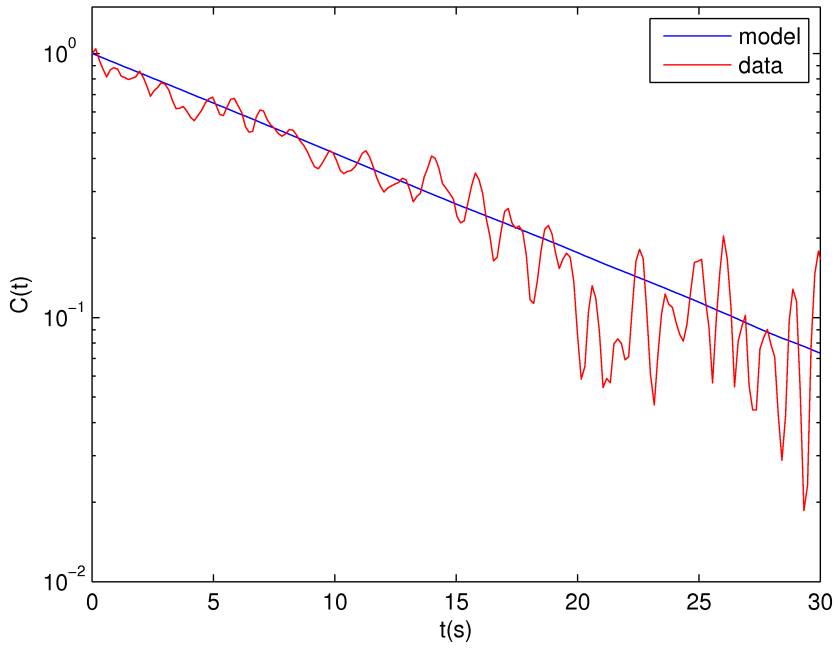
Figure 3.7: Posterior distribution of ρ.

ρ	0.033 ± 0.005 s^{-1}
a	0.014 ± 0.002 mm/s^2
ξ	0.0023 ± 0.0004 mm/s^2
τ	54 ± 8 s

Table 3.2: Best fit parameters.



(a)



(b)

Figure 3.8: (a) Best fit of the stationary velocity distribution $P_s(v_x, 0)$; (b) Best fit for the velocity autocorrelation. The oscillations in the experimental curve (red) are due to the spiraling pattern of the motion.

Conclusions

In this thesis we have examined microbial motility from the statistical point of view, trying to deduce if there are characteristics of the motion common to different microorganisms. Reviewing previous experiments in section 1.2, we have seen that some interesting observations stem from the data. First, the motion of active, living creatures in liquid media is not erratic as that of a Brownian particle, which passively diffuses with a randomly fluctuating velocity as a result of collisions with the molecules of the liquid. Instead, microorganisms tend to keep their velocity, and thus their direction, for a certain persistence time [3]. This means that the velocity has a non-zero autocorrelation time. This persistence of velocity is a well established property of motion, and allows us to make an interesting inference about the motility process. In the case of active living matter, the cause of motion and of changes of direction are not collisions but the propelling action of the particle itself. Since these organisms swim at low Reynolds number, a condition in which inertia is negligible, we must deduce that this persistence of velocity is the result of a persistence in the propelling force they exert. In fact, were they to stop swimming, they would immediately come to a halt, without coasting. Thus we can conclude that the observed persistence of velocity implies persistence of the propelling force.
While this qualitative observation comes already from a visual inspection of trajectories, where constant velocity results in approximately straight segments, and changes of velocity result in turns, additional observations can be made on the distribution of velocities. Velocities are computed from the tracking of individual trajectories, and the result of a large number of experiments is that the distribution $P(v)$ of velocities is non-Gaussian. However, not all experimental findings agree on the shape of this distribution: some experiments discover an enhanced[2] probability of small velocities [18, 21], others report bigger tails [16], others still observe a suppressed probability of small velocities, resulting in a volcano-shaped distribution [17]. Thus, what is well established is that $P(v)$ is non-Gaussian, but its shape seems to be non unique. The conclusion we draw from this is that a model of the motility must account for non-Gaussianity of v, and hopefully be able to explain different shapes of $P(v)$.
Experiments have also looked at other statistical features of motion: the autocorrelation of the velocity, for example, has been often observed to display an exponential decay, although in some cases a double-exponential [18, 7] or a power-law decay [16] was seen.
The interpretation of spatial features of the motion is even more difficult: many experiments found, computing the Mean Square Displacement of trajectories, what seems anomalous diffusion [16, 19]. However, it is not easy to track microorganisms for very long, and authors of [7] advanced the hypothesis that this seemingly anomalous behaviour could only be a transient before a diffusive behaviour, unseen in experiments due to shortness of the tracking.

To sum up, experiments allow to establish some common features (persistence, non-Gaussianity),

[2] With respect to a Gaussian distribution.

but on other aspects of the motion results are not univocal (exponential decay of autocorrelation, anomalous scaling). In the absence of clarifying experiments [3], the modelling should start from the characteristics of the motion which have a solid base. In section 1.3 we saw that to model persistence the Ornstein-Uhlenbeck process is a good start, and that there have been many previous attempt to extend it in two dimensions. What is common to all these attempts is that the stochastic propelling force is modelled through a Gaussian white noise, so its value is uncorrelated in time. The claim of this thesis is that a more realistic choice for the force is a coloured process, i.e. a process with a non-zero autocorrelation time. In fact, as explained above, the observation of the motion hints that the microorganism keeps exerting a force in the same direction for a certain lapse of time. In particular, we examined the motion of the ciliates *Colpidium* sp., protists which trajectories alternate straight segments and neat changes of direction (a motion similar to the *run and tumble* motion seen in *E. coli*). Our guess of the propelling force is that exposed in chapter 3: a Markovian stochastic process $\vec{\sigma}_t$ which direction θ is chosen uniformly in $[0, 2\pi]$ and which modulus varies around a non-zero mean. A model of the form $\dot{\vec{v}}_t = -\rho\vec{v}_t + \vec{\sigma}_t$ allowed us to recover both the experimental $P(v)$, which displays a volcano-like shape, and the experimental autocorrelation function $C(t)$, which decays exponentially. That we succeeded in simultaneously recovering both these feature is promising, and so is the fact that the model, for a different regime of the parameters, predicts for $P(v)$ a shape peaked at the origin but non-Gaussian [4], which could be apt to describe the motion of different organisms, for which such a shape is found.
We were not able to study the spatial features of the motion due to short tracking time of trajectories, but in the near future we plan to do experiments with a wider observation window, that will allow a longer tracking time. This will hopefully clarify whether the scaling of the experimental Mean Square Displacement is diffusive or anomalous, and whether our model succeeds in describing it correctly.
We also plan to verify the fitness of our model comparing it with data relative to other microbial species, to discover if the introduction of a non-Gaussian and coloured term in the model brings good results not only for the species we examined so far, *Colpidium* sp., but also for other species.

[3] A longer tracking time in particular seems essential to capture the MSD behaviour.
[4] Except in a certain limit.

Appendix A

The Ornstein-Uhlenbeck process

The O-U process in one dimension is described by the equations

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v \tag{A.1}$$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -\rho v + \sigma\xi \tag{A.2}$$

where ξ is a Gaussian white noise process.
Although we write the equations in this form because the process was initially meant to describe a particle moving with that equation of motion, i.e. subject to a stochastic force with the above properties, they do not have a rigorous mathematical meaning, since ξ is not a well-defined mathematical entity, due to the fact that it has a diverging variance (see property 2 above).
However, we can write a mathematically sound equation for the process v(t) if we notice ([30]) that $\int_0^t \xi(t')\mathrm{d}t' = W(t)$, where $W(t)$ is the Wiener process, a perfectly well-defined mathematical entity.
Thus by integration we can write for the process v(t) the following equation:

$$v(t) = v_0 - \rho\int_0^t v(t)\mathrm{d}t + \sigma\int_0^t \mathrm{d}W(t) \tag{A.3}$$

where the integral in the last member is a stochastic integral, which in this case may be defined simply as[1]

$$\int_0^t \mathrm{d}W(t) := \underset{n\to\infty}{\text{ms-lim}} \sum_{i=1}^{n} [W(t_i) - W(t_{i-1})] \tag{A.4}$$

being $0 = t_1, t_2, ..., t_n = t$ a partition of the interval [0,t] in n subintervals. Since we are dealing with stochastic variables, when taking the limit of the partial sums we take the *mean-square limit* [2].
Equation (A.3) can be written in the so called *Stochastic Differential Equation* (SDE) form

$$\mathrm{d}v(t) = -\rho v(t)\mathrm{d}t + \sigma\mathrm{d}W(t) \tag{A.5}$$

[1]Since the Gaussian white noise ξ in eq. (A.2) is additive, i.e. it is not multiplied by a function of the process v, there is no need to introduce the distinction between Ito and Stratonovich stochastic integral. In fact, in this case the two formulations are equal. For the definition of stochastic integral in the general case see [30].

[2]The process X_t is said to converge in mean square to X if $\lim_{t\to\infty}\langle(X_t - X)^2\rangle = 0$ [30]

which form reminds us of the our original equation, although its meaning is exactly the same as (A.3).

Now that we have written the equation for the process in a clearly interpretable form, we may proceed to solve it. But first, let us establish a useful fact about the integral with measure $\mathrm{d}W(t)$.
It can be proved [30] that

$$\mathrm{d}W(t)^2 = \mathrm{d}t \tag{A.6}$$

in the sense that

$$\int_0^t \mathrm{d}W(t')^2 = \int_0^t \mathrm{d}t' \tag{A.7}$$

This can be interpreted saying that $\mathrm{d}W(t)$ is an infinitesim of order $\frac{1}{2}$ in t.
A consequence of this is that if we want to change variables in eq. (A.5) to $f[v(t)]$, to be consistent at the order $\mathrm{d}t$ we have to consider $\mathrm{d}W(t)^2$ too. In particular:

$$\begin{aligned} \mathrm{d}f[v(t)] :&= f[v(t+\mathrm{d}t)] - f[v(t)] = \partial_v f[v(t)]\mathrm{d}v + \partial_v^2 f[v(t)]\mathrm{d}v^2 + ... \\ &= [-\rho v \partial_v f + \sigma^2 \partial_v^2 f]\mathrm{d}t + \sigma \partial_v f \mathrm{d}W \end{aligned} \tag{A.8}$$

Let us employ a change of variables to put the equation (A.5) into a form easier to solve. If we take $y = ve^{-\rho t}$, we get the new equation

$$\mathrm{d}y(t) = \partial_t y \mathrm{d}t + \partial_v y \mathrm{d}v + \partial_v^2 y \mathrm{d}v^2 = \sigma e^{\rho t}\mathrm{d}W(t) \tag{A.9}$$

which can be straightaway integrated to obtain

$$y(t) = \sigma \int_0^t e^{\rho t'}\mathrm{d}W(t') + y(0) \tag{A.10}$$

Changing back to $v(t)$ we get

$$v(t) = \sigma \int_0^t e^{-\rho(t-t')}\mathrm{d}W(t') + v(0)e^{-\rho t} \tag{A.11}$$

From this result we notice that $v(t)$ is a Gaussian variable, since the increments of the Wiener process are Gaussian and the deterministic term can be seen as Gaussian with null variance. This means that its distribution is completely determined by its first two moments, $\langle v(t)\rangle$ and $\langle v(t)^2\rangle$, which we compute in the following way:

$$\langle v(t)\rangle = \langle \sigma \int_0^t e^{-\rho(t-t')}\mathrm{d}W(t')\rangle + \langle v(0)e^{-\rho t}\rangle = v(0)e^{-\rho t} \tag{A.12}$$

having used $\langle \mathrm{d}W(t)\rangle = 0$.

$$\begin{aligned} \langle v(t)^2\rangle &= v(0)^2 e^{-2\rho t} + \sigma^2 \langle \int_0^t e^{-\rho(t-t')}\mathrm{d}W(t') \int_0^t e^{-\rho(t-t'')}\mathrm{d}W(t'')\rangle \\ &= v(0)^2 e^{-2\rho t} + \sigma^2 \int_0^t \int_0^t e^{-\rho(2t-t'-t'')}\langle \mathrm{d}W(t')\mathrm{d}W(t'')\rangle \\ &= v(0)^2 e^{-2\rho t} + \sigma^2 \int_0^t e^{-2\rho(t-t')}dt' \\ &= v(0)^2 e^{-2\rho t} + \frac{\sigma^2}{2\rho}(1 - e^{-2\rho t}) \end{aligned} \tag{A.13}$$

having used that the increments of the Wiener process at different times are independent [30].

The variance is then

$$\sigma_v^2 = \langle v(t)^2 \rangle - \langle v(t) \rangle^2 = \frac{\sigma^2}{2\rho}(1 - e^{-2\rho t}) \tag{A.14}$$

and we can write the probability distribution of v as

$$P(v,t) = \frac{e^{\frac{(v-\langle v(t) \rangle)^2}{2\sigma_v^2}}}{\sqrt{2\pi\sigma_v^2}} \tag{A.15}$$

We may also compute the autocorrelation as

$$\begin{aligned}
\langle v(t)v(s) \rangle &= v(0)^2 e^{-\rho(t+s)} + \sigma^2 \langle \int_0^t e^{-\rho(t-t')} \mathrm{d}W(t') \int_0^s e^{-\rho(s-s')} \mathrm{d}W(s') \rangle \\
&= v(0)^2 e^{-\rho(t+s)} + \sigma^2 \int_0^{\min(t,s)} e^{-\rho(t+s-2t')} dt' \\
&= [v(0)^2 - \frac{\sigma^2}{2\rho}] e^{-\rho(t+s)} + \frac{\sigma^2}{2\rho} e^{-\rho(t+s-\min(t,s))}
\end{aligned} \tag{A.16}$$

meaning that, if we take $s = t + \tau$, the decay has a simple exponential form in τ:

$$\langle v(t)v(t+\tau) \rangle \sim \frac{\sigma^2}{2\rho} e^{-\rho\tau} \tag{A.17}$$

Finally, let us have a look at the integral process $x(t)$ and compute its second moment, which is the Mean Square Displacement. From the integration of eq. (A.1) we get

$$x(t) = \int_0^t v(t') \mathrm{d}t' \tag{A.18}$$

having chosen $x(0) = 0$ for simplicity.
We notice that since v is Gaussian, x is also Gaussian, so it is also completely determined by mean and variance, which we proceed to compute.

$$\langle x(t) \rangle = \int_0^t \langle v(t') \rangle \mathrm{d}t' = \frac{v(0)}{\rho}(1 - e^{-\rho t}) \tag{A.19}$$

$$\begin{aligned}
\langle x(t)^2 \rangle &= \int_0^t \int_0^t \langle v(t')v(t'') \rangle \mathrm{d}t' \mathrm{d}t'' \\
&= \int_0^t \int_0^t \Big(\Big[v(0)^2 - \frac{\sigma^2}{2\rho} \Big] e^{-\rho(t'+t'')} + \frac{\sigma^2}{2\rho} e^{-\rho(t'+t''-\min(t',t''))} \Big) \mathrm{d}t' \mathrm{d}t'' \\
&= \frac{v(0)^2}{\rho^2}(1 - e^{-\rho t})^2 + \frac{\sigma^2}{\rho^2} \Big[t - \frac{2}{\rho}(1 - e^{-\rho t}) + \frac{1}{2\rho}(1 - e^{-2\rho t}) \Big]
\end{aligned} \tag{A.20}$$

This result is particularly interesting because we can distinguish two regimes in which the MSD behaves in two different ways:

- when $t \ll 1/\rho$ expanding the exponentials we get $\langle x(t)^2 \rangle \approx v(0)^2 t^2$;
- when $t \gg 1/\rho$ all the exponentials are negligible, thus we get $\langle x(t)^2 \rangle \approx \frac{\sigma^2}{\rho} t$.

The first regime is called *ballistic*. It lasts only in the first instants of the motion, when the velocity is still strongly correlated with the initial velocity $v(0)$.
After many persistence times ρ^{-1} have passed the direction of motion will have changed several times, and as long as the Mean Square Displacement is concerned the motion will be indistinguishable from diffusion. This regime is called *diffusive*.

A characteristic of the O-U process is that it admits a stationary distribution $P_s(v)$, the mean and variance of which we find making t tend to ∞. We get

$$\langle v(t)\rangle_s = \lim_{t\to\infty} v(0)e^{-\rho t} = 0 \tag{A.21}$$

$$\sigma^2_{v,s} = \langle v(t)^2\rangle_s = \lim_{t\to\infty}\Big[v(0)^2 e^{-2\rho t} + \frac{\sigma^2}{2\rho}(1-e^{-2\rho t})\Big] = \frac{\sigma^2}{2\rho} \tag{A.22}$$

consequently the stationary distribution is

$$P_s(v) = \frac{e^{-\frac{\rho v^2}{\sigma^2}}}{\sqrt{\pi\frac{\sigma^2}{\rho}}} \tag{A.23}$$

Appendix B

Infinitesimal generator for Markov Processes

In this appendix we briefly give the definition of infinitesimal generator for a Markovian stochastic process and recover the expressions given in section 3.2 for the generators of the stochastic force and of the velocity process. For an exhaustive description of Markov processes with continuous time and generic state space see [39].
Let $(X_t)_{t\geq 0}$ be a Markov process with values in S. Then, if f $\in$ C(S), its infinitesimal generator is defined as

$$\mathcal{L}f(x) := \lim_{t\downarrow 0} \frac{S_t f(x) - f(x)}{t} \tag{B.1}$$

where S_t is the transition semigroup of the process,i.e.

$$S_t f(x) := E(f(x_t)|x_0 = x) \tag{B.2}$$

In the case of the stochastic force of our model the process is (θ_t,σ_t) with state space $S = [0, 2\pi] \times [-\infty, +\infty]$. The semigroup is then

$$\begin{aligned} S_t f(\theta,\sigma) :&= E(f(\theta_t,\sigma_t)|\theta_0 = \theta, \sigma_0 = \sigma) \\ &= \int_0^{2\pi} d\theta' \int_{-\infty}^{+\infty} d\sigma' f(\theta',\sigma') P(\theta_t = \theta', \sigma_t = \sigma'|0) \end{aligned} \tag{B.3}$$

where we denoted, for brevity,

$$P(\theta_t = \theta', \sigma_t = \sigma'|0) := P(\theta_t = \theta', \sigma_t = \sigma'|\theta_0 = \theta, \sigma_0 = \sigma) \tag{B.4}$$

Now, let N_t be the Poisson process of parameter $1/\tau$ counting the number of jumps of our process. Then the probability on the right of eq. (B.3) is

$$\begin{aligned} P(\theta_t \in [\theta', \theta' + \mathrm{d}\theta'], \sigma_t \in [\sigma', \sigma' + \mathrm{d}\sigma']|0) &= \sum_{n=0}^{\infty} P(N_t = n) P(\theta_n \in [\theta', \theta' + \mathrm{d}\theta'], \sigma_n \in [\sigma', \sigma' + \mathrm{d}\sigma']|0) \\ &= \Big[e^{-\frac{t}{\tau}} \delta(\theta - \theta')\delta(\sigma - \sigma') + \sum_{n=1}^{\infty} e^{-\frac{t}{\tau}} \frac{\left(\frac{t}{\tau}\right)^n}{n!} \frac{1}{2\pi} \frac{1}{\sqrt{2\pi}\xi} e^{-\frac{(\sigma'-a)^2}{2\xi^2}} \Big] \mathrm{d}\theta' \mathrm{d}\sigma' \\ &= \Big[e^{-\frac{t}{\tau}} \delta(\theta - \theta')\delta(\sigma - \sigma') + \frac{1}{2\pi} \frac{1}{\sqrt{2\pi}\xi} e^{-\frac{(\sigma'-a)^2}{2\xi^2}} (1 - e^{-\frac{t}{\tau}}) \Big] \mathrm{d}\theta' \mathrm{d}\sigma' \end{aligned} \tag{B.5}$$

where we have used the fact that the probability of jumping to a certain state is independent of the starting point, so $P(\theta_n \in [\theta', \theta'+\mathrm{d}\theta'], \sigma_n \in [\sigma', \sigma'+\mathrm{d}\sigma']|\theta_0 \in [\theta, \theta+\mathrm{d}\theta], \sigma_0 \in [\sigma, \sigma+\mathrm{d}\sigma]) =$

$P(\theta_n \in [\theta', \theta' + \mathrm{d}\theta'], \sigma_n \in [\sigma', \sigma' + \mathrm{d}\sigma'])$.
From the definition of infinitesimal generator (B.1) we then get

$$\begin{aligned}
\mathcal{L}f(\theta,\sigma) :&= \lim_{t\downarrow 0} \frac{S_t f(\theta,\sigma) - f(\theta,\sigma)}{t} \\
&= \lim_{t\downarrow 0} \frac{1}{t}\Big[\int_0^{2\pi} d\theta' \int_{-\infty}^{+\infty} d\sigma' f(\theta',\sigma') e^{-\frac{t}{\tau}} \delta(\theta-\theta')\delta(\sigma-\sigma') + \\
&+ \frac{1}{2\pi}\frac{1}{\sqrt{2\pi}\xi} e^{-\frac{(\sigma'-a)^2}{2\xi^2}}(1-e^{-\frac{t}{\tau}}) - f(\theta,\sigma)\Big] \\
&= \lim_{t\downarrow 0}\Big[f(\theta,\sigma)\frac{e^{-\frac{t}{\tau}}-1}{t} + \int_0^{2\pi} d\theta' \int_{-\infty}^{+\infty} d\sigma' f(\theta',\sigma')\frac{1}{2\pi}\frac{1}{\sqrt{2\pi}\xi} e^{-\frac{(\sigma'-a)^2}{2\xi^2}}\frac{1-e^{-\frac{t}{\tau}}}{t}\Big] \\
&= \frac{1}{\tau}\Big[-f(\theta,\sigma) + \int_0^{2\pi} d\theta' \int_{-\infty}^{+\infty} d\sigma' f(\theta',\sigma')\frac{1}{2\pi}\frac{1}{\sqrt{2\pi}\xi} e^{-\frac{(\sigma'-a)^2}{2\xi^2}}\Big] \\
&= \frac{1}{2\pi\tau}\int_0^{2\pi} d\theta' \int_{-\infty}^{+\infty} d\sigma' \frac{1}{\sqrt{2\pi}\xi} e^{-\frac{(\sigma'-a)^2}{2\xi^2}}[f(\theta',\sigma') - f(\theta,\sigma)]
\end{aligned} \tag{B.6}$$

We recovered eq. (3.4).

Let us now obtain the expression of the generator of the velocity process ($\vec{v}_t$,θ_t,σ_t). From the definition given above

$$\mathcal{L}f(\vec{v},\theta,\sigma) = \lim_{t\downarrow 0}\frac{1}{t}\Big[\int d\vec{v}' \int_0^{2\pi} d\theta' \int_{-\infty}^{+\infty} d\sigma' P(\vec{v}_t = \vec{v}', \theta_t = \theta', \sigma_t = \sigma'|0) f(\vec{v}',\theta',\sigma') - f(\vec{v},\theta,\sigma)\Big] \tag{B.7}$$

where again we denoted

$$P(\vec{v}_t = \vec{v}', \theta_t = \theta', \sigma_t = \sigma'|0) := P(\vec{v}_t = \vec{v}', \theta_t = \theta', \sigma_t = \sigma'|\vec{v}_0 = \vec{v}, \theta_0 = \theta, \sigma_0 = \sigma) \tag{B.8}$$

We can calculate the probability in the expression above neglecting terms $o(t)$, since in the limit they will go to zero. This means that the probability of having more than one jump of the stochastic force in the time interval t is negligible: it would be $e^{-\frac{t}{\tau}}\frac{(t/\tau)^n}{n!}$, which for $n < 1$ is $o(t)$.
We have then two cases to consider:

- if the force does not jump, which happens with probability $1 - \frac{t}{\tau} + o(t)$, the velocity at time t will be $\vec{v}_t = \vec{v}_0 + t(-\rho\vec{v}_0 + \vec{\sigma}_0) + o(t)$
- if the force jumps once, with probability $\frac{t}{\tau}$+o(t), at time νt with $\nu \in$[0,1], the velocity at time t will be $\vec{v}_t = \vec{v}_0 + \nu t(-\rho\vec{v}_0 + \vec{\sigma}_0) + (1-\nu)t(-\rho\vec{v}_0 + \vec{\sigma}_t) + o(t)$

Consequently

$$\begin{aligned}
&P(\vec{v} \in [\vec{v}' + \mathrm{d}\vec{v}'], \theta \in [\theta' + \mathrm{d}\theta'], \sigma \in [\sigma' + \mathrm{d}\sigma']|0) = \\
&= \delta(\vec{v}' - [\vec{v} + t(-\rho\vec{v} + \vec{\sigma})])\delta(\theta-\theta')\delta(\sigma-\sigma')(1-\frac{t}{\tau})\mathrm{d}\vec{v}'\mathrm{d}\theta'\mathrm{d}\sigma' + \\
&+ \delta(\vec{v}' - [\vec{v} + \nu t(-\rho\vec{v} + \vec{\sigma}) + (1-\nu)t(-\rho\vec{v} - \vec{\sigma}')])\frac{t}{\tau}\frac{1}{2\pi}\frac{1}{\sqrt{2\pi}\xi} e^{-\frac{(\sigma'-a)^2}{2\xi^2}}\mathrm{d}\vec{v}'\mathrm{d}\theta'\mathrm{d}\sigma' + o(t)
\end{aligned} \tag{B.9}$$

Substituting this expression into (B.7) we get

$$\begin{aligned}\mathcal{L}f =& \lim_{t\downarrow 0}\frac{1}{t}\Big[\int d\vec{v}'\int_0^{2\pi}d\theta'\int_{-\infty}^{+\infty}d\sigma'\delta(\vec{v}'-[\vec{v}+\nu t(-\rho\vec{v}+\vec{\sigma})+(1-\nu)t(-\rho\vec{v}-\vec{\sigma}')])\cdot\\ &\cdot\frac{t}{\tau}\frac{1}{2\pi}\frac{1}{\sqrt{2\pi}\xi}e^{-\frac{(\sigma'-a)^2}{2\xi^2}}(1-\frac{t}{\tau})f(\vec{v}',\theta',\sigma')+f(\vec{v}+t(-\rho\vec{v}+\vec{\sigma}),\theta,\sigma)\big(1-\frac{t}{\tau}\big)-f(\vec{v},\theta,\sigma)\Big]\\ =&\frac{1}{2\pi\tau}\int_0^{2\pi}d\theta'\int_{-\infty}^{+\infty}d\sigma' f(\vec{v},\theta',\sigma')\frac{1}{\sqrt{2\pi}\xi}e^{-\frac{(\sigma'-a)^2}{2\xi^2}}+\\ &+\lim_{t\downarrow 0}\frac{1}{t}\Big[f(\vec{v},\theta,\sigma)\big(1-\frac{t}{\tau}\big)+t(-\rho\vec{v}+\vec{\sigma})\nabla_{\vec{v}}f(\vec{v},\theta,\sigma)+o(t)-f(\vec{v},\theta,\sigma)\Big]\\ =&(-\rho\vec{v}+\vec{\sigma})\nabla_{\vec{v}}f(\vec{v},\theta,\sigma)+\frac{1}{2\pi\tau}\int_0^{2\pi}d\theta'\int_{-\infty}^{+\infty}d\sigma'\frac{1}{\sqrt{2\pi}\xi}e^{-\frac{(\sigma'-a)^2}{2\xi^2}}[f(\vec{v},\theta',\sigma')-f(\vec{v},\theta,\sigma)]\end{aligned} \tag{B.10}$$

The knowledge of the infinitesimal generator for the process allows us to write the equation for its stationary probability. In fact, a probability measure μ on S is stationary for the process with infinitesimal generator $\mathcal{L}$ if

$$\int \mathcal{L}f d\mu = 0 \qquad \forall f \tag{B.11}$$

which if the measure is absolutely continuous with respect to the Lebesgue measure becomes

$$0=\int \mathcal{L}f\mu \mathrm{d}x = \langle \mathcal{L}f,\mu\rangle = \langle f,\mathcal{L}^*\mu\rangle \quad \forall f \quad \Leftrightarrow \quad \mathcal{L}^*\mu = 0 \tag{B.12}$$

This last equation, substituting the adjoint of the generator we have just found ((B.10)), gives equation (3.10) of section 3.2 for the stationary distribution.

Bibliography

[1] Bray D.: *Cell Movements: from Molecules to Motility* New York : Garland Publishing, 1992.

[2] Berg H.C.: *E. Coli in Motion* Springer 2004

[3] Mendez V., Campos D., Bartumeus F.: *Stochastic Foundations in Movement Ecology* Springer 2014

[4] Zaoli S., Giometto A., Formentin M., Azaele S., Rinaldo A., Maritan A.: *Phenomenological modeling for the motility of self-propelled microorganisms.* preprint: http://arxiv.org/abs/1407.1762

[5] Czirok, A., Schlett, K., Madarasz, E., Vicsek, T.: *Exponential distribution of locomotion activity in cell cultures.* Phys. Rev. Lett. **81**(14), 3038–3041 (1998). doi:10.1103/PhysRevLett. 81.3038

[6] Berg H.C.: *Random Walks in Biology.* Princeton University Press, 1993

[7] Li L., Cox E.C., Flyvbjerg H.: *"Dicty Dynamics": Dictyostelium motility as persistent random motion.* Phys Biol. 2011 August ; 8(4): 046006. doi:10.1088/1478-3975/8/4/046006.

[8] Purcell E.: *Life at low Reynolds-number.* Am. J. Phys. **45**(1), 3–11 (1977). doi:10.1119/1.10903

[9] Berg H.C., Brown D.A.: *Chemotaxis in Escherichia coli analysed by three-dimensional tracking.* Nature 239:500–504(1972)

[10] Van Haasert, P.J.M.: *Amoeboid cells use protrusions for walking, gliding and swimming* PLOS Comput. Biol. **6**(8)(2010)

[11] Bosgraaf L.,Van Haasert, P.J.M.: *The ordered extension of pseudopodia by amoeboid cells in the absence of external cues.* PLoS ONE. 2009;4(4):e5253.

[12] Dusenbery D.: *Minimum size limit for useful locomotion by free-swimming microbes.* Proc.Natl.Acad.Sci.USA **94**(20),10949–10954 (1997).

[13] Dusenbery D.: *Fitness landscapes for effects of shape on chemotaxis and other behaviours of bacteria.* J. Bacteriol. **180**(22) 5978–5983 (1998)

[14] Sokolov A., Aranson I.S., Kessler J.O., Goldstein R.E.: *Concentration dependence of the collective dynamics of swimming bacteria.* Phys. Rev. Lett. **98**(15) (2007). doi:10.1103/PhysRevLett.98.158102

[15] Vicsek T., Zafeiris A.: *Collective Motion.* Physics Reports **517**(3–4), 71–140 (2012)

[16] Upadhyaya, A., Rieu, J.P., Glazier, J.A., Sawada, Y.: *Anomalous diffusion and non-gaussian velocity distribution of Hydra cells in cellular aggregates.* Physica A 293(3–4), 549–558 (2001). http://dx.doi.org/10.1016/S0378-4371(01)00009-7

[17] Liang L., Nørrelykke S. F., Cox E.C.: *Persistent Cell Motion in the Absence of External Signals: A Search Strategy for Eukaryotic Cells.* PLoS ONE 3(5), e2093 (2008)

[18] Selmeczi D., Mosler S., Hagedorn P.H., Larsen N.B., Flyvbjerg H.: *Cell Motility as Persistent Random Motion: Theories from Experiments.* Biophysical Journal, Volume 89 (2005), 912–931

[19] Dieterich P., Klages, R., Preuss R., Schwab A.: *Anomalous dynamics of cell migration.* PNAS USA **105**(2), 459–463 (2008). doi: 10.1073/pnas.0707603105

[20] Selmeczi, D., Li, L., Pedersen, L.I.I., Nrrelykke, 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. Spec. Top. **157**(1), 1–15 (2008). doi:10.1140/epjst/e2008-00626-x. http://dx.doi.org/10.1140/epjst/e2008-00626-x

[21] Takagi H., Sato M.J., Yanagida T., Ueda M.: *Functional Analysis of Spontaneous Cell Movement under Different Physiological Conditions.* PLOS ONE **3**(7) (2008). doi:10.1371/journal.pone.0002648

[22] Schienbein M., Gruler H.; *Langevin equation, Fokker-Planck equation and cell migration.* Bull. Math. Biol. **55**(3), 585-608 (1993)

[23] Levandowsky M., White B.S., Schuster F.L.: *Random movements of soil amebas.* Acta Protozool. **36**(4), 237–248 (1997). http://www.nencki.gov.pl/abstr/abs36/abs36-26.htm

[24] Potdar A.A., Jeon J., Weaver A.M., Quaranta V., Cummings P.T.: *Human mammary epithelial cells exhibit a bimodal correlated random walk pattern.* PLOS ONE **5**(3) (2010). doi:10.1371/journal.pone.0009636

[25] Potdar A.A., Lu J., Jeon J., Weaver A.M., Cummings P.T.: *Bimodal analysis of mammary epithelial cell migration in two dimensions.* Ann. Biomed. Eng. **37**(1), 230–245 (2009). doi:10.1007/s10439-008-9592-y

[26] Van Haastert P.J.M.: *A model for a correlated random walk based on the ordered extension of pseudopodia.* PLOS Comput. Biol. **6**(8) (2010). doi:10.1371/journal.pcbi.1000874

[27] Gruver J.S., Potdar A.A., Jeon J., Sai J., Anderson B., Webb D., Richmond A., Quaranta V., Cummings P.T., Chung C.Y.: *Bimodal analysis reveals a general scaling law governing nondirected and chemotactic cell motility.* Biophys. J. **99**(2), 367–376 (2010). doi:10.1016/j.bpj.2010.03.073

[28] Shenderov A., Sheetz M.: *Inversely correlated cycles in speed and turning in an Ameba: an oscillatory model of cell locomotion.* Biophys. J. **72**(5), 2382–2389 (1997)

[29] Stokes C., Lauffenburger D., Williams S.: *Migration of individual microvessels endothelial-cells–stochastic model and parameter measurement.* J.Cell Sci **99**(Part 2), 419–430 (1991)

[30] Gardiner C. *Handbook of stochastic methods for physics, chemistry and the natural sciences.* Springer-Verlag (2004)

[31] Metzler R., Klafter J.: *The random walk's guide to anomalous diffusion: a fractional dynamics approach.* Phys Rep 339:1–77. (2000)

[32] Romanczuk P., Schimansky-Geier L.: *Brownian motion with active fluctuations.* Phys. Rev. Lett. **106**(23) (2011)

[33] Kitahara K., Horsthemke W., Lefever R.: *Coloured-noise-induced transitions: exact results for external dichotomous Markovian noise* Phys. Lett. A70, 377 (1979)

[34] Van Den Broeck C.: *On the Relation between White Shot Noise, Gaussian White Noise, and the Dichotomic Markov Process* Journal of Statistical Physics, Vol. 31, No. 3, 1983

[35] Sbalzarini I.F., Koumoutsakos P. *Feature point tracking and trajectory analysis for video imaging in cell biology.* J. Struct. Biol. **151**(2):182–195 (2005).

[36] Wilson G.B., Jahn T.L., Fonseca J.R.,: *Helical nature of the ciliary beat of Colpidium Striatum.* Acta Protozoologica, Volumen XII, Fasc. 34, Warszawa (1975)

[37] Gillespie D.T.: *Exact Stochastic Simulation of Coupled Chemical Reactions.* The Journal of Physical Chemistry **81**(25): 2340–2361 (1977) doi:10.1021/j100540a008

[38] Vrugt J.A., ter Braak C.J.F., Diks C.G.H., Robinson B.A., Hyman J.M., Higdon D. *Accelerating Markov Chain Monte Carlo Simulation by Differential Evolution with Self-Adaptive Randomized Subspace Sampling.* International journal of nonlinear sciences and numerical simulation **10**(3):273 –290.(2009)

[39] Liggett T.M.: *Continuous time Markov process: an Introduction.* American Mathematical Society (2010)