Statistical Analysis and Stochastic Modelling of Foraging Bumblebees.

Lenz, Friedrich

Statistical Analysis and Stochastic Modelling of Foraging Bumblebees

Friedrich Lenz

School of Mathematical Sciences

Queen Mary University of London

A thesis submitted in partial fulfilment

of the requirements for the Degree of

Doctor of Philosophy

18/01/2013

Abstract

In the analysis of movement patterns of animals, stochastic processes play an important role, providing us with a variety of tools to examine, model and simulate their behaviour. In this thesis we focus on the foraging of specific animals - bumblebees - and analyse experimental data to understand the influence of changes in the bumblebees' environment on their search flights. Starting with a discussion of main classes of stochastic models useful for the description of foraging animals, we then look at a multitude of environmental factors influencing the dynamics of animals in their search for food. With this background we examine flight data of foraging bumblebees obtained from a laboratory experiment by stochastic analyses. The main point of interest of this analysis is the description, modelling and understanding of the data with respect to the influence of predatory threats on the bumblebee's foraging search flights. After this detail-oriented view on interactions of bumblebees with food sources and predators in the experimental data, we develop a generalized reorientation model. By extracting the necessary information from the data, we arrive at a generalized correlated random walk foraging model for bumblebee flights, which we discuss and compare to the experimental data via simulations. We finish with a discussion of anomalous fluctuation relations and some results on spectral densities of autocorrelation functions. While this part is not directly related to the analysis of foraging, it concerns a closely related class of stochastic processes described by Langevin equations with non-trivial autocorrelation functions.

Acknowledgements

Without support from others this thesis would not have been possible. I would therefore like to thank my colleagues, the departmental staff, and especially:

Person(s) \ Reason	supervision and inspiration	interesting discussions	productive collaborations	providing experimental data	thorough thesis examination	friendly atmosphere
Thomas Ings		x	x	x		x
Lars Chittka		x	x	x		x
Aleksei Chechkin		x	x			x
Jonathan Pitchford		x			x	x
Mark Broom		x			x	x
The Curve Crowd		x				x
Rainer Klages	x	x	x			x

Contents

List of Figures

Overview

The desire to understand the behaviour of animals gave rise to a broad field of research. A specific but still large part of this field is concerned with the analysis of the movement patterns of foragers. While the topic of food search of animals has been analysed for a long time, many interesting questions remain under investigation due to the inherent intricacy of the field: a large variety of environmental factors, competing evolutionary pressures, and the complexity of the analysed forager itself make the analysis of experimental foraging data challenging.

Since a complete modelling of all the biological factors relevant to describe the food search of an animal is typically not feasible, and even such a model would likely still be non-deterministic, stochastic models have been introduced into foraging research. Consequently, stochastic processes play an important role by providing us with a multitude of tools to examine, model and simulate animal movement patterns.

In the first chapter we start with a discussion of environmental factors influencing the dynamics of animals in their search for food (see section 1.2). We then present the main classes of stochastic models used to describe the foraging of animals in section 1.3. At the boundary between optimal foraging theory and stochastic processes the idea of the optimality of specific random walks to find randomly distributed targets developed into the Lévy flight hypothesis, which we discuss in section 1.4 in the context of the biological factors and of its relation to the alternative foraging models.

In the following two chapters we focus on a specific foraging animal: the bumblebee. These two chapters are based on a laboratory experiment by Thomas C. Ings and Lars Chittka [1], who collaborated with me on the topic of predation threats together with Aleksei V. Chechkin and Rainer Klages (published in

[2]). In chapter 2 we examine flight data of foraging bumblebees in order to understand the influence of predatory threats on the bumblebee's foraging search flights. While the threat of predation is only one of the biological factors affecting the foraging behaviour, the set-up of the experiment as described in section 2.1 has the advantage of keeping all other environmental influences constant. After the main section 2.2 of the chapter, which consists of the analysis and interpretation of the effects of predators and a partial model thereof, we also connect our findings to the discussion on the applicability of the Lévy flight hypothesis (see section 1.4.2) in section 2.3.

From the detail-oriented view on interactions of bumblebees with food sources and predators in chapter 2, we turn our attention to the search flights between flower visits in chapter 3. In this chapter the goal is to arrive at a generalized reorientation model for bumblebee flights, which we develop by extracting the necessary information from the experimental data in section 3.2. After comparison of the resulting generalized correlated random walk foraging model with a simple reorientation model, the model is validated by simulation and comparison to the data of the bumblebee experiment in section 3.3. A paper written in collaboration with Aleksei V. Chechkin and Rainer Klages with the main results of this chapter has been published in [3].

A common theme recurring through the previous chapters — apart from foraging — are Langevin equations and their generalizations. In chapter 4 we finish the thesis with a discussion of anomalous fluctuation relations and some results on spectral densities of autocorrelation functions. While the class of stochastic processes we investigate here are not directly related to the analysis of foraging, they are also described by a (differently) generalized Langevin equation with nontrivial autocorrelation functions. The content of this chapter is closely related to a publication together with Aleksei V. Chechkin and Rainer Klages [4], who are its main authors.

Chapter 1

Foraging and the Lévy Flight Hypothesis

1.1 Embedding into Foraging Research

Understanding the behaviour of foraging animals is an endeavour which is challenging due to the complex environment in which the search for food happens. Correspondingly broad are the topics in the area of foraging research. In the following two chapters we will analyse experimental data to answer more specific questions about foraging bumblebees, i.e. can we understand the interaction between bumblebees and their predators and how can we model the foraging behaviour. However, in this chapter we first want to introduce the relevant biological factors and the essential stochastic foraging models, as the background to discuss optimal foraging. Specifically, *"What is the best statistical strategy in order to search efficiently for randomly located objects?"* has been used as a guiding question to research the movement patterns of foraging animals. A search model was proposed which predicts that *Lévy walks*[1] are optimal to search for sparse and revisitable food sources [5]. The basic idea is that instead of a random walk with a constant or normally distributed step length l, a random walk whose flight lengths are distributed as a power-law is used to model the movement of a forager, that is:

[1] Misleadingly called *Lévy flights* in [5]. See also section 1.3.2.

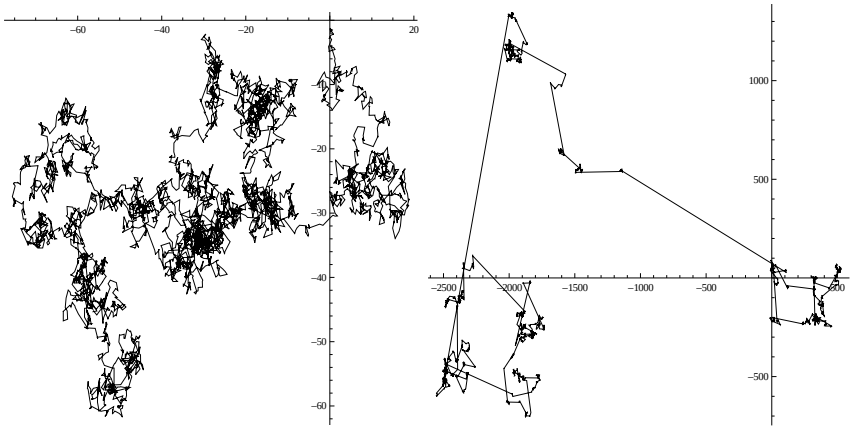

Figure 1.1: Example trajectories (both with $n = 5000$ steps) of a Brownian motion (left) and a superdiffusive Lévy walk with typical scale-free step lengths (right).

$\rho(l) \sim l^{-\beta}$ for $1 < \beta \leq 3$, where ρ is the probability density function of the step length. As this random walk exhibits the property that very large step length are much more common than in the case of Brownian motion, leading to superdiffusive movement dynamics as shown in Fig. 1.1. In section 1.3.2 we will look at Lévy flights, Lévy walks and why they are interesting for modelling diffusion and especially foraging.

The optimality claim of the Lévy flight hypothesis is interesting from the biological point of view as one would expect evolutionary pressure on the forager from the energy and time spent for searching food, which could lead to a corresponding adaptation of the forager towards an optimal foraging strategy. It is however unclear if the Lévy walk model is applicable to real animals and even if it is, whether the resulting advantage for survival would be important enough in comparison to other evolutionary pressures to push the animals towards an adaptation of this foraging scheme. A second reason why Lévy walks were investigated as a viable strategy were claims that they arise naturally from the interaction of the forager with the food source distribution [6, 7]. Both reasons lead to the search for Lévy walks in experimental data. Foraging data of many animals has been analysed to find evidence for Lévy walks as a search strategy: albatross [6, 8]; deer, bumblebees [5, 8]; Drosophila [9]; tuna, cod, turtles, shark and penguins

[10] – with mixed results, as discussed in section 1.4.2. In quite a few cases the analysis had to be revised due to errors in the data, insufficient data or methodological problems [8]. The evidence for the existence of Lévy walks in foraging data therefore still remains under discussion.

In this chapter we will begin in section 1.2 with an introduction to the main biological factors which affect the movement of animals while searching for food. After this exposition of the complexity of the environment of foragers we will then introduce common classes of stochastic processes which are typically used to model animal movement in section 1.3. In section 1.4 we will then concern ourselves with optimality in foraging, which has been developed under the name of *optimal foraging theory* (see section 1.4.1). There we will focus on the necessary assumptions and experimental evidence for the *Lévy flight hypothesis*, which arose at the interface of the theory of stochastic processes and optimal foraging theory (see section 1.4.2). This chapter also functions as a reference with respect to stochastic models and the biology of foraging for the following chapters 2 and 3. We will meet the Lévy flight hypothesis again in section 2.3 in the context of our analysis of experimental data of foraging bumblebees.

1.2 Biological Factors in Foraging

The ability of an animal to forage and the resulting movement patterns depend on a large number of biological factors. While some of them are related to the environment the animal is living in, others are given by the internal constraints acting on it, e.g., its energy needs and energy storage capacity.

In this section we can only give a short overview of the main aspects which play a role in the discussion of search strategies of foragers. A nice introduction to the field of foraging animals presenting the variety of ecological factors and matching theories can be found in [11].

1.2.1 Habitat and Home Range

Many animals have a limited space in which they can search for food. There are different reasons for these spacial limitations of the habitat. On the one hand

the area may be bounded, e.g., by physical barriers or by boundaries to adjacent territories of rivals. A different kind of spatial restriction is generated if the animal has to return periodically to a specific location. Examples are sleeping places and locations where its social group gathers.

When discussing foraging, the effects of bounded motion have to be recognized, and may themselves give rise to interesting technical and also biological questions, e.g., how boundaries between territories are maintained [12]. For the discussion of search strategies, in many cases an unbounded foraging space is assumed and the need to return to the origin is neglected, since the focus on optimal search strategies is already complicated enough without these complications. However for the analysis of experimental data these concerns have to be addressed, e.g., by removing movements from and to a sleeping place [8] – effectively assuming that the dynamics during long search periods is independent from transient movement phases to access the area of search.

1.2.2 Heterogeneous Environments

Nearly all animals live in a highly complex and heterogeneous environment. One common cause for a heterogeneous environment is a non-uniform food source distribution, which therefore has been examined analytically [13, 14, 15, 16] and experimentally [17, 18, 19] and is of concern when optimal foraging strategies are investigated (see section 1.4). Even seemingly monotonous environments such as the ocean surface have spatially heterogeneous food sources for a foraging sea bird, in this case structured plankton which is aggregated by water eddies [18].

External spatially varying parameters, e.g., food availability, temperature or water depth, can affect parameters of the movement of the foraging animal.

When the internal parameters of the animal can be adequately described by a low number of “internal states”, the movement can be modelled by a composite random walk (see section 1.3.5). However, if the number of states is very large or even infinite, the idea of switching between different modes might not make much sense in those models – they are usually not considered when speaking about composite random walks in the context of foraging animals. The parameter dependencies in these models might either be phenomenologically treatable, e.g.,

by superstatistical methods[20], or they have to be modelled explicitly. Understanding the dependence on heterogeneities is extremely important since neglecting them can lead to a false classification of the movement process.

Destructive Foraging

Depending on the nature of the food, food sources can be described as either replenishing (a flower) or only once visitable (a single fish) by a forager. If the rates of replenishing the sources are high in comparison to the time between returns of a forager, the former can also be called *revisitable*. Due to the changes to their own environment, which can induce heterogeneities, *destructive* foragers on the other hand are also of interest: especially in cases of collective behaviour. We will see in section 1.4.2 that the ability to revisit sources changes optimal search behaviour drastically.

1.2.3 Risks while Foraging

The search for food is a risky endeavour for many foragers. Especially the risk from predators has to be considered and weighed against the risk associated with not foraging, which at some point means starvation. The dynamics of the interaction between predators and their prey has been studied with various approaches, e.g., by Lotka-Volterra equations. Various analyses in optimal foraging theory (e.g. [21, 22]) have tried to quantify the risks and benefits of foraging in order to find foraging strategies with ideal trade-offs.

In chapter 2 we will have a closer look at a specific example of the influence of predation threat on the movement of a forager.

1.2.4 Heterogeneous Populations

Among many animals cooperative behaviour exists between the individuals. While we will restrict ourselves in the following to the analysis of the movement of a single individual without interaction with its peers, we will look at a few effects of animal cooperation which can influence the movement of foragers. The analysis of the movement of animals in a collective, e.g., in swarms, though interesting

on its own [23], is not our focus here since it is strongly governed by in-swarm interactions and only loosely connected to food search.

The tendency of animals to return to their group can induce restrictions on its movement similar to those of a habitat or home range (see section 1.2.1). Several experimental studies have analysed animal dispersal, i.e., the spreading of a group of animals from a single source site, finding a decay of the population density which has fat tails [24]. While at first this has been seen as evidence for super-diffusive movement processes (see section 1.3.2), more detailed analyses of experimental data revealed the heterogeneity in the animal populations as the source of the seeming anomalous diffusion [24, 25, 26, 27, 28]. The diffusion appeared to be anomalous because, while the movement of each individual was well-described by a normal diffusion, the diffusion constant varied between the individuals. Notice that, while this is not exactly the same as a composite random walk (see section 1.3.5), the effect of finding a seemingly anomalous movement process from averaging over data with an unaccounted parameter is the same.

1.2.5 Perception of the Forager

For the search behaviour of foraging animals their perception plays an important role. Only through the limits on their senses does it become necessary to move around in order to search for food. While there are some analyses investigating the role of perception on search behaviour (see e.g. [23, 15, 29, 30]), in most studies of search behaviour the modelling of the perception is simplified by assuming that the animal has a fixed range of perception r: all targets closer than r are automatically recognized, while no other targets are perceived. This assumption can be interpreted as a simple model for undirected local search whose movement patterns are too small to be resolved in the larger model. If however this local search is important enough that it has to be modelled explicitly as a separate and maybe different stochastic process, intermittent models become a quite natural choice for the movement analysis (see section 1.3.5).

The range of perception of the forager is an important parameter to take into consideration when the optimality of stochastic foraging strategies is analysed (see section 1.4.2). If the range would be large enough that the animal always perceives

nearby food sources, the problem of foraging shifts away from stochastic search (see section 1.2.6).

1.2.6 Deterministic Foraging and Memory

When modelling the foraging behaviour of an animal, the question of how much knowledge the forager possesses about its food sources arises. With complete knowledge the problem of efficient foraging reduces to finding a solution, or an approximation thereof, to a *travelling salesman problem* [31], where the physical distances between the food sources might be modified by environmental conditions and risks when specifying the corresponding problem.

Another way determinism can enter into the discussion of foraging is if the forager always has sufficient information to know the nearest food source, e.g., due to a large range of perception (see section 1.2.5), and always chooses this source as the next target. This kind of deterministic foraging in a random environment has been analysed and compared to stochastic foraging models, e.g., when analysing the effects of the shape of the home range of the forager ([32, 33], see also section 1.2.1).

In addition to a small perceptive range, a typical assumption of stochastic search models is that the forager has no memory of the already explored part of its environment. However, while this is reflected in the basic stochastic foraging models (see section 1.3), most animals do have the capability to gather information. The resulting effects on foraging behaviour have been recognized as important for many foragers, e.g., the spatial memory of bumblebees was analysed [34] and the effects of learning on movement patterns investigated [2, 31, 35]. For example the development of trap lines, i.e. fixed foraging routes between revisitable targets (see section 1.2.2) has been studied [36, 31, 35, 37].

1.3 Stochastic Movement Models

In order to understand the dynamics of animal movement, a large number of stochastic models have been developed over time. In this section we can only give a brief overview of the most essential classes of models which have been studied

in relation to animal foraging. All the models presented here have in common that an organism is modelled as an ideal particle without any internal structure or learning capability, usually moving in an unstructured two-dimensional space. For real animals these assumptions will not always hold. Nonetheless the presented model classes have been shown to be useful first approximations for the description of animal movement. It should be kept in mind that the models usually have to be modified to incorporate the main environmental factors (see section 1.2) for a comparison to experimental data.

1.3.1 Random Walk

The Wiener Process

The *Wiener process* $W(t)$ for $t \in \mathbb{R}^+$ is a time-continuous stochastic process starting with $W(0) = 0$, whose increments $W(t+\tau) - W(t)$ are independent and normally distributed with a variance[38]

$$\left\langle (W(t+\tau) - W(t))^2 \right\rangle = \tau \tag{1.1}$$

for all $\tau \geq 0$ and $W(t)$ is *almost surely* continuous: the probability of a sample path to be continuous is one. The usefulness of the Wiener Process as a model for normal diffusion and random searches, e.g., in foraging, is in large parts a result of the central limit theorem. As experimental data is by construction discrete in time, a discretised Wiener Process, i.e. a random walk with Gaussian step lengths and fixed time step τ, is often used for comparison to experiments.

1.3.2 Lévy Flights and Lévy Walks

Stable Distributions

Given a family of independent random variables $\{X_i\}$, $i \in \mathbb{N}$, which are all drawn from the same distribution with finite mean μ and finite variance σ^2, the position of a *random walker*[2] after n steps is given by $S_n = \sum_{i=1}^{n} X_i$. The *central limit*

[2] While the random walk is presented here in one dimension, it can be generalized to more dimensions, e.g. by choosing a direction uniformly. For a random walker with a normal step

theorem states that the distribution of S_n converges to a normal distribution after scaling:

$$\sqrt{n}\left(\frac{S_n}{n} - \mu\right) \to \mathcal{N}(0, \sigma^2). \tag{1.2}$$

In this limit the random walk converges to a Wiener process which is therefore used to model random motion. For the normal distributions the central limit theorem applies as well — as they converge to themselves they are an example of stable distributions.

A real, non-degenerate distribution $\mathcal{D}$ is called *stable* iff for all independent random variables X, X_1, X_2 with distribution $\mathcal{D}$ and all $a, b \in \mathbb{R}$, $aX_1 + bX_2$ is distributed like $cX + d$ for some $c, d \in \mathbb{R}$. The central limit theorem ensures that in the family of distributions with finite mean and variance only the normal distributions are stable.

If one eliminates this restriction, and considers the partial sums $S_n = \sum_{i=1}^{n} Y_i$ of arbitrary independent identically-distributed (i.i.d.) random variables Y_i, $i \in \mathbb{N}$, the family of limit distributions is larger. The only distributions Z which are possible as limits for the recentred and rescaled partial sums, that is $\frac{S_n - a_n}{b_n} \to Z$ for suitable coefficients a_n, b_n are the *Lévy alpha-stable distributions*, also called the *stable laws* [39, 40]. These are defined by their characteristic functions [41]

$$\phi_Z(\omega) := \left\langle e^{i\omega Z} \right\rangle = \exp\left(i\delta\omega - |\gamma\omega|^{\alpha}\left(1 + i\beta \mathrm{sgn}(\omega) K(\alpha, \omega)\right)\right) \tag{1.3}$$

where

$$K(\alpha, \omega) = \begin{cases} -\tan(\pi\alpha/2) & : \quad \alpha \neq 1 \\ 2\log|\omega|/\pi & : \quad \alpha = 1. \end{cases} \tag{1.4}$$

The restricted parameters are the index $\alpha \in (0, 2]$, the skewness $\beta \in [-1, 1]$, the scale $\gamma > 0$, and the location δ. Here we are only interested in random variables Y_i with an even probability density function, which result in symmetric $(\beta = 0)$ and centred $(\delta = 0)$ stable distributions with:

$$\phi_Z(\omega) = e^{-|\gamma\omega|^{\alpha}}. \tag{1.5}$$

length distribution $\mathcal{N}(0, \sigma^2)$ this would be the same as using independent normal step length distributions in each dimension separately.

The central limit theorem is generalized in the following way: the stable distributions attract other distributions when summing their random variables depending on the asymptotic behaviour of the tail probabilities of Y_i. For a finite variance $\langle Y_i^2 \rangle$ (or $y^2 P[|Y_i| > y] \to 0$) the resulting index is $\alpha = 2$, giving a normal distribution as this is the case for the usual central limit theorem. However, if $P[|Y_i > y| \sim cy^{-\alpha}$ for some $c > 0$ and $\alpha \in (0, 2)$ as $y \to \infty$, this α is also the index of the stable distribution Z [39, 40].

One reason why stable distributions for step lengths are of special interest in movement models is that coarse-graining experimental data by always treating, e.g., two consecutive movement steps as a single step, does not change the step length distribution (up to a scale). This is a nice property for the analysis especially since it might be difficult to define, and hard to determine experimentally, when a step ends [42, 43, 44]. However this does not mean that, when analysing animal movement, models based on stable step lengths distributions are the only available choices (see e.g. sections 1.3.4, 1.3.5).

Lévy Flights

In the context of foraging it was questioned whether a normal diffusion is a good model for the random search behaviour of animals. As an alternative which models a super-diffusive behaviour, random walks in one and two dimensions with scale-free step lengths l have been used. Let us assume that the step length distribution $\rho(l)$ has a power-law tail, that is $\rho(l) \sim l^{-\beta}$ for large l. For $\beta \leq 1$ the distribution $\rho(l)$ cannot be normalised as $\int_0^\infty \rho(l) dl$ diverges. For $\beta > 3$ the first and second moments exist. This means that in this case the central limit theorem applies and the position distribution S_n converges to a Gaussian for large n. This leaves the range of $1 < \beta \leq 3$ where the variance diverges. By the generalized central limit theorem (see above) the process converges to a Lévy stable distribution which conserves the power-law tail. In these random walks, which are called *Lévy flights*, the time used for each step is assumed to be a constant. This means that the total time is just the number of steps, and the velocity is proportional to the step length. Since this means that the velocity is unbounded, Lévy flights are not very useful as a foraging model as animal velocities are always bounded.

Lévy Walks

A *Lévy walk* distinguishes itself from the Lévy flights model by using a constant speed v_0 for the random walker. This means that instead of jumping from one position to the next in constant time, the walker moves with a speed v_0 from position S_n to S_{n+1} in a time span proportional to the step length. The Lévy walk is a more realistic foraging model than the Lévy flights even if the speed of animals is rarely constant. The model can be seen as an approximation where v_0 corresponds to the mean speed of an animal. Therefore when scale-free processes are considered as animal movement models, Lévy walks are nearly always preferred to Lévy flights. The question of whether Lévy walks are a good description of real animal movement will be discussed in section 1.4.2.

In a similar way, classic random walks have also been generalized to another class of stochastic processes: *continuous time random walks* [45, 46, 47]. For these models, not only is the step size drawn from a distribution, but the time between one step and the next is also drawn from another, different distribution, where both random variables are typically drawn i.i.d. The interpretation of the random update times is typically that they are induced by a random environment, which causes the object to stick and wait after each step. Due to the additional waiting times continuous time random walks can exhibit subdiffusion, which makes them interesting in the context of crowded environments [48]. Continuous time random walks can also be superdiffusive as a result of heavy-tailed step size distributions, e.g., a Lévy walk can be seen as a special case of a continuous time random walk. However, apart from Lévy walks continuous time random walks are only rarely [49] used for modelling foraging animals for the same reason as the Lévy flights: the typically unbounded velocities do not match well to the physics of animal movement.

1.3.3 Correlated Random Walk (Reorientation Models)

Typical candidates for modelling diffusion-like processes are e.g. (generalized) Langevin equations or continuous time random walks [50]. In the case of foraging models, it has to be taken into account, that animals often have a “front”-direction in which they move and have to turn their body to change their movement direc-

tion. This is commonly modelled by *reorientation models* (also called *correlated random walks* (CRW)) and has been analysed [42, 14, 43] and used to describe the movement of a variety of animals [44, 51].

In (two-dimensional) *reorientation models*, the movement of an animal, with position $(x, y) \in \mathbb{R}^2$ heading into the direction given by the angle α in a static frame of reference, is described by:

$$\alpha(t + \tau) = \alpha(t) + \beta(t) \tag{1.6}$$

$$x(t + \tau) = x(t) + l(t) \cos(\alpha(t)) \tag{1.7}$$

$$y(t + \tau) = y(t) - l(t) \sin(\alpha(t)) \tag{1.8}$$

where l is the *step length*, τ is the discrete time-step and β is the *turning angle*, i.e. the change in direction in a single time step. Many variations to this description are used, for example the time-continuous version in [42]. Proportional to the step length is the speed $v(t) := l(t)/\tau$.

The turning angle β and the step length $l \geq 0$ are drawn independently from probability densities $p(\beta)$ and $q(l)$. These densities are usually estimated from animal trajectories. In some models (e.g. [51]) the analysis is simplified by assuming a constant step length l_0 of the animal (and therefore also a constant speed), which means that $q(l) = \delta(l - l_0)$. Most reorientation models ignore autocorrelations of β and l: each random variable is drawn i.i.d. If the autocorrelations decay fast enough, i.e. exponentially, the model is diffusive. The diffusive properties, e.g., mean squared displacement and diffusion constant have been derived analytically for various subclasses of CRWs [44, 43, 52, 25]. However autocorrelations have been rarely[53] used to analyse experimental movement data of foragers [54, 55]. Processes with anomalous diffusion are often excluded from the class of correlated random walks and treated separately.

Reorientation models are not only used when directional correlations occur because of an asymmetry of the animal and the necessity to turn its body. In many applications the CRW is used to model the intended direction of movement of the animal instead of the orientation of the body. In these cases, the CRW describes the dynamics of the intended direction, which can give rise to directional

persistence of animals over long time scales even though the animal changes the orientation of its body on much shorter time scales.

If the autocorrelation time scale is large it can become difficult to distinguish a correlated random walk from a Lévy walk. For this determination given a finite amount of experimental data da Luz et.al. [56] gave a necessary criterion relating the time scale of the exponential autocorrelation of a Markovian correlated random walk to the distribution of its turning angles.

1.3.4 Generalized Langevin Equation (Active Brownian Particles)

While many models of animal movement use a time-discrete description with clearly discernible movement steps, most time-continuous models are in essence Langevin equations or generalizations thereof.

The Langevin Equation

A *Langevin equation* is a stochastic differential equation with a deterministic part, called $\mathbf{f}$, and an added noise term $\mathbf{\Gamma}$ which is multiplied by the matrix $\mathbf{k}$ of coefficient functions:

$$\frac{d}{dt}\mathbf{X}(t) = \mathbf{f}(\mathbf{X}(t), t) + \mathbf{k}(\mathbf{X}(t), t)\mathbf{\Gamma}(t) \tag{1.9}$$

where $\mathbf{\Gamma}$ is called a stochastic force or *Langevin force*: it is a vector of *white noise*, meaning that $\langle \Gamma_i(t) \rangle = 0$ and $\langle \Gamma_i(t)\Gamma_j(t') \rangle = \delta_{i,j}\delta(t - t')$ for all dimensions, where $\delta_{i,j}$ is Kronecker's delta and $\delta(t - t')$ is the Dirac delta function.[3] An equivalent restatement of Eq. (1.9) is:

$$d\mathbf{X}(t) = \mathbf{f}(\mathbf{X}(t), t)dt + \mathbf{k}(\mathbf{X}(t), t)d\mathbf{W}(t). \tag{1.10}$$

From the Langevin equation alone it is not clear which system we describe, as we have not defined yet how to integrate it. As the Wiener process is nowhere

[3] There is also the possibility to define the Langevin equation for stochastic forces Γ which are not δ-correlated. Such *coloured noise* will be used in sections 3.2.6 and 3.2.7.

differentiable it is not integrable in the Riemann sense. There are two different ways to define a stochastic integral called the *Itō integral* and the *Stratonovich integral*[4] [38, 46] . While the Riemann integral is independent of the supporting points of the discretisation, the stochastic integrals differ for varying approximation approaches.

Both integration methods are defined by

$$\int_0^t u(x_s, s)dW(s) := \lim_{n\to 0} \sum_{i=0}^{n-1} u(x_{\tau_i}, \tau_i)(W(t_{i+1}) - W(t_i)) \tag{1.11}$$

with $0 = t_0 < t_1 < \ldots < t_n = t$ for any function $u(x_s, s)$ of the Wiener process. The two definitions of the stochastic integrals differ only in the choice of τ_i as a function of t_i and t_{i+1}:

- The Itō integral uses $\tau_i = t_i$. It is non-anticipating which means that for numeric integration f only has to be evaluated at the previous time step as described in section A.3.
- The Stratonovich integral uses $\tau_i = \frac{t_i + t_{i+1}}{2}$ and is symmetric in time.

The Stratonovich integral has the advantage that it corresponds to the calculus of the Riemann integral whereas the Itō integral needs a special one: the *Itō calculus*.

Given that we specify the integration method by saying that we use the Itō or Stratonovich interpretation of the Langevin equation we completely describe a Markov process (see section 3.2.1). It depends on the process we want to model which interpretation is appropriate.

Given the Langevin coefficients in one of the interpretations it is possible to convert them to the other interpretation with the equations [38]:

$$f_i(\mathbf{X}, t) = \widetilde{f}_i(\mathbf{X}, t) + \frac{1}{2} \sum_{j,l} k_{j,l}(\mathbf{X}, t) \frac{\partial k_{i,l}}{\partial X_j}(\mathbf{X}, t) \tag{1.12}$$

[4]To be precise, there are not only two ways to define a stochastic integral, but an infinite number, as you are free to choose the supporting points of the approximation. Itō and Stratonovich have analytical advantages the other definitions do not have. This means that there is no reason not to use one of the two.

where $\mathbf{f}$ is the deterministic part of the Langevin equation in Itō interpretation and $\widetilde{\mathbf{f}}$ in the Stratonovich interpretation. $\mathbf{k}$ is identical for both interpretations.

From Eq. (1.12) it follows that for given $\mathbf{f}$ and $\mathbf{k}$ the Itō and the Stratonovich interpretation describe the same process if $\mathbf{k}(\mathbf{X}, t)$ is constant over phase space. In this case the non-deterministic term is called *additive noise* in contrast to the general *multiplicative noise*.

The integration of a Langevin equation makes it possible to generate sample paths of a Markov process if the Langevin coefficients are known, e.g., with the Euler-Maruyama approximation (see section A.3). As a special case deterministic systems are modelled by Langevin equations with $\mathbf{k} \equiv 0$, however the convention is to restrict the term only to systems with non-trivial $\mathbf{f}$ and $\mathbf{k}$.

Langevin Movement Models

One example of how the Langevin equation is used for modelling animal movement are *active Brownian particles*. The basic model describes the position $\mathbf{r}(t)$ of the animal by the dynamics of its velocity $\mathbf{v}(t) = \dot{\mathbf{r}}(t)$ via

$$m\dot{\mathbf{v}} = -\gamma(\mathbf{v})\mathbf{v} + \sqrt{2D}\mathbf{\Gamma}(t), \tag{1.13}$$

where m is the particle mass, $\gamma(\mathbf{v})$ is a velocity-dependent "friction" and the diffusion constant D scales the Langevin force (see e.g. [57, 58]). For active particles, the "friction" $\gamma(\mathbf{v})$ is allowed to be negative, resulting in an active acceleration which is usually powered by the metabolism of the animal. A nice introduction to active Brownian particles including many-particle interactions can be found in [23].

A variety of different generalizations of the Langevin equation (Eq. (1.9)) related to active Brownian particles will be used in this thesis. A Langevin equation with an additional potential will be used in section 2.2.5. In chapter 3 a non-Markovian version of a generalized Langevin equation in polar coordinates will be extracted from experimental data to model foraging bumblebees using the connection (see section 3.2.3) of the Langevin equation to the Fokker-Planck equation (see section 3.2.2). In the final chapter a generalized Langevin equation with memory kernel will be studied.

1.3.5 Composite Random Walks and Intermittent Search

One assumption which is common to all animal movement models shown above is that only one process is responsible for generating the path of an animal. While this focus on a single explanatory mechanism might be aesthetically pleasing, it has to be questioned when dealing with the movement of highly complex organisms in complex environments. It is natural to start from a simple description by a diffusive random walk (section 1.3.1) and, when observing that the model is not consistent with the experimental data, continue by developing more general models. However insisting that the resulting models stay simple may lead us astray in understanding animal movement. For example, if one looks at a typical recorded trajectory of a foraging animal and finds that there are many small step lengths but also a non-negligible amount of much larger step lengths, one might be tempted to say: "Since there are too many large steps for a Brownian random walk, we need a process with a step length distribution with a heavy tail. And since the steps should be made of (not observed) sub-steps, only a stable distribution is plausible. (see section 1.3.2)" This explanation simplifies by assuming that a process has only one relevant scale. But for many animals, movement serves different purposes which can have different relevant spacial scales and time scales. Therefore it is plausible, that animals switch between different internal states governing different movement phases.

Composite random walks explicitly model these states $s_1, \ldots, s_n$ and switches between them. The switching between states s_i and s_j in one time step Δt is then specified by a (time-independent) switching probability $p(s_i \to s_j)$ for each pair i, j, with $\sum_{j=1}^{n} p(s_i \to s_j) = 1$ for each i, and the switching process is usually assumed to be uncorrelated. These probabilities can sometimes be reconstructed from time series, e.g., via *hidden Markov models* (HMM) [59]. Associated to each state is a stochastic process, which generates the trajectory of the animal while the state is active. In principle any process could be used for a state, but using Lévy strategies is only done occasionally [60] as scale-free strategies, while possible, are a bit of a mismatch when one explicitly wants to explain the scales of the involved processes.

Although in theory one could use models with many states, often just two

states are used, with a Bernoulli switching process. In many cases of foraging animals one movement phase corresponds to a local search for food while another corresponds to movement with larger step lengths or stronger directional persistence. Composite random walks are therefore sometimes called *intermittent search processes*, even though the underlying process of the phenomenon is not directly related to intermittency of dynamical systems. [5] These two-state models can be understood as a compromise between *exploration* of food abundance and *exploitation* of local food sources. However, there are many other reasons why composite random walks are used for modelling since switching between different movement phases is a good description for a variety of biological factors. Examples are spatially inhomogeneous environments leading to a switching between different kinds of behaviour (see section 1.2.2), switching between directed and undirected modes of movement [62], and behaviour induced by external changes in the environment, e.g., day and night cycles. In our analysis of experimental data of foraging bumblebees in chapter 2 we will encounter an example of intermittency induced by spatial inhomogeneities (see section 2.2.2).

Due to the flexibility in describing different biological aspects for animal movement, composite random walks have been used to model a variety of experiments [63, 64, 59, 62, 65, 66] and a large number of analyses of their properties have been done [66, 67]. A review of intermittent search processes can be found in [48].

The step length distribution ρ_l of a composite random walk is a mixture of the step length distributions of each of the contributing processes, with weights which depend on the transition probabilities between the states. Even with very simple processes for each state, e.g., scaled Wiener processes, the resulting distribution ρ_l can be hard to distinguish from other those of other models given experimental data due to the large variety of possible ρ_l [67]. This has been especially important in the search for Lévy walks in animal movement data. Typically a preference of a power-law tail of a step length distribution over an exponential[6] tail has been interpreted as evidence supporting the biological Lévy hypothesis (see section 1.4.2),

[5] Notice that in some cases the term "intermittency" has been used for a model with only one process. In [61] a Levy Walk model was used and all steps below a threshold were retroactively assigned to a local search phase and all other steps to a relocation phase.

[6] or even thinner

e.g., in [10, 68]. However, since there are many biological factors which can in effect lead to a composite random walk, at least a few simple cases of composite random walks should be excluded before one can attribute experimental data to a Lévy walk. Otherwise, e.g., if the only alternative model is a Wiener process, the step-length distribution of a composite random walk can be easily mistakenly identified as the power law of a Lévy walk [69].

1.4 Optimal Foraging

1.4.1 Classical Optimal Foraging Theory

In the long and exciting process of biology developing from a natural science with a stronger descriptive focus to a more quantitative science, the question of how to explain the complex behaviour of animals proved to be a resistant one. While early research gave to questionable descriptions of their behaviour, e.g., the "bad wolf" or the "greedy cow", the tables turned with the advent of the theory of evolution through natural selection [70]. *Optimal foraging theory* arose as the attempt to examine foraging through a set of core principles [11]:

- a goal function which will be maximized, e.g., energy,
- options from which the forager can choose and
- environmental constraints acting on the forager, including internal constraints.

With the assumption that the goal function is positively correlated with the chance of survival of the species of the forager, natural selection provides the selection pressure, such that the animal is pushed towards choosing those available options, which under the environmental constraints maximize the goal function. While the field of optimal foraging theory (see e.g. [71, 21, 22, 72]) diversified until today [11] it also lost its name due to cosmetic reasons [11]. Part of the diversification came from considering more complex goal functions, which model survival chances more realistically. This means that also trade-offs, e.g., between gathered food and predation risk have been considered [11].

The Lévy hypothesis (see below in section 1.4.2) integrates nicely into this framework, by considering different stochastic food search processes as options of the forager. The constraints are given by the lack of memory of the searcher, limited perception and a random environment. In this context it is important to realize that the choice of the goal function plays a decisive role: it should consider both gain and costs, typically the food gained per distance travelled for a *cruise forager*, i.e. a forager which continuously scans the environment while moving. This might however not be the correct choice of a goal function. Examples are so-called *saltatory predator*, that switch between predation attempts and "blind" movement phases, which often has the largest energetic costs associated to feeding, e.g.' his predation attempts, and not to the travelling between attempts ([73], compare sections 1.2.5, 1.3.5). Another example is an animal which has a very limited capability to store energy. This animal might want to optimize towards a more regular/predictable uptake of food at the cost of the total amount of food in order to avoid starvation [73].

In the context of optimal foraging theory another aspect of the biological Lévy hypothesis might be of interest to investigate. If the hypothesis would be correct in case of a specific application, the resulting stochastic movement process would be scale-free [5, 6, 74]. In its strong interpretation (section 1.4.2) the resulting dynamics would be quite inflexible: a model with more possible parameters corresponding to different temporal or spatial scales might be good for the adaptability of the animal. This is another reason why we might not expect the strong interpretation to hold. For example a composite random walk may not have the optimal step length distribution for a particular search problem, but might be easy to produce, compose and be flexible. The differences to some optimal distribution might not be large enough to give rise to evolutionary pressure [75].

1.4.2 Lévy Hypothesis

In the context of early experiments on foraging animals [6, 76, 9] and some theoretical work on Lévy processes [77, 5] and their applicability to movement data of animals, the *Lévy hypothesis* was born. However, *the* Lévy hypothesis is actually two (main) hypotheses, which should be considered separately. Though they are

usually collectively known as the "Lévy flight hypothesis", where the "flight"-part is actually a historical misnomer since most biologically interesting models are variations of Lévy walks, we give them distinct names here.

Mathematical Lévy Hypothesis

Quantifying foraging behaviour of organisms by statistical analysis has raised the question of whether biologically relevant search strategies can be identified by mathematical modelling [71, 78, 79, 75, 80, 48, 74].

In short, the "Lévy flight hypothesis" predicts that a random search with jump lengths following a power law minimizes the search time for sparsely, randomly distributed, replenishing food sources [77, 5, 74]. In the following we will call this the *mathematical Lévy hypothesis*. While this can be examined as a theoretical question about stochastic processes, it also makes predictions in the context of optimal foraging theory (section 1.4.1). Here we will first clarify the class of processes under consideration and the assumptions needed for the hypothesis to hold. Whether there are actually any organisms which perform Lévy walks (or try to approximate them) is a different question, which we will look at in the next subsection.

For all analyses of optimal foraging discussed here, it is assumed that the foraging animal has no prior knowledge about the position of the randomly distributed food sources, searches stochastically and has no memory (see section 1.2.6): the step-lengths are drawn i.i.d. from a power-law distribution. The optimized goal function, i.e. the search efficiency, used here is the visited food sources per distance travelled (see section 1.4.1).

A major reason why Lévy walks were considered as a model class of interest is that they fill the gap between ballistic motion and a normal diffusion depending on the power β of the power-law decay of the step length distribution (see section 1.3.2). For β approximating 1 from above, the behaviour of Lévy walks is dominated by a few largest steps, making it effectively ballistic for most purposes. This limit is ideal in case of destructive foraging (section 1.2.2) since it decreases the probability to revisit food sources which are not available any more [5]. The non-trivial case is therefore *non-destructive* foraging or cases which can

be approximated by it, e.g. a non-uniform distribution of food sources: if the food sources are distributed *patchily* [81], i.e. in clusters, they can collectively act as re-visitable food patches for long time scales, even though the individual food source is destroyed on visits [5].

It has been shown that the search efficiency can only depend on the chosen search process if the non-destructive forager stops *during* its movement steps when a food source is in the range of perception r_v ([73], see section 1.2.5). These foraging strategies are also called *target-truncated* [82]. Since the interest lies in target-truncated Lévy walks it is important to notice, that the mean free path λ to the targets induces an extra decay of the actual step length distribution [5, 82]. The optimal exponent for a target-truncated Lévy walk is

$$\beta_{opt} = 2 - \left(\ln \frac{\lambda}{r_v}\right)^{-2} \tag{1.14}$$

which means that for sparse food sources ($\lambda \gg r_v$) a Cauchy distribution ($\beta_{opt} = 2$) is the optimal step length distribution, i.e. a target-truncated Lévy walk is better than Brownian motion and ballistic motion [5]. For this result, after each visit to a food source the forager has to be placed near the food source at a distance corresponding to the perception range r_v. If it is placed further and further away, the relative efficiency of the Lévy walk versus ballistic motion decreases — as does β_{opt} [82]. Together with the quite strong assumptions needed, this raises the question of how robust the hypothesis is.

The result on the optimality is dependent on the restriction to Lévy walks. If one allows also, e.g., composite random walks (see section 1.3.5), the situation gets more complicated. In particular, models have been analysed [66, 7] which distinguish a fast relocation phase (ballistic or Lévy walk) in which no food is collected, and a phase of slow local food searches (typically a Brownian walk or a correlated random walk [83]). The results depended on a variety of model details, e.g. the time spent in each phase [66, 61]. Overviews of this zoo of different models can be found, e.g., in [74, 82]. In summary, the required conditions for the optimality of Lévy walks are very strict, which suggests that the mathematical Lévy hypothesis should not be seen as a general paradigm for search strategies, but rather as a remarkable exceptional case.

Biological Lévy Hypothesis

A different hypothesis related to the mathematical Lévy hypothesis is the question whether any animals actually perform Lévy walks when foraging, which we call the *biological Lévy hypothesis*. While this hypothesis is motivated by the optimality of Lévy walks under quite specific conditions (see above), many studies have tried to find Lévy walks in movement data of animals under a variety of environmental conditions (e.g. [6, 76, 84, 10, 68, 85]).

The interest in Lévy walks was motivated by optimal foraging theory (see section 1.4.1), that is, by an argument via evolutionary pressure: if Lévy walks offer animals a more efficient way to forage in a random environment than other stochastic foraging strategies, then it is likely that animals have evolved which at least approximate this behaviour. Notice that the evolutionary argument does not guarantee that the optimum is reached — suboptimal behaviour might be good enough. This raises the question of whether the biological Lévy hypothesis should be understood in the sense that the underlying search process actually *is* a Lévy walk, i.e. that it is directly generated via some bio-chemical or bio-physical process. This *strong* interpretation of the biological Lévy hypothesis is usually not assumed to be valid since no such process has been found and since in classical optimal foraging theory (see section 1.4.1) the optima are not assumed to be realized by the organisms [11]. Instead, usually a *weaker* biological Lévy hypothesis is investigated: the assumption is that the animals movement is driven by another stochastic process, which is well approximated by a Lévy walk. The immediate question which arises is: how is "very well" measured?

The distinction between the strong and weak Lévy hypothesis is sometimes discussed as the difference between *adapted* and *emergent* behaviour [32]. the strong interpretation corresponds to an internal mechanism which the animal developed to adapt to evolutionary pressure, while in the weak interpretation the Lévy movement pattern emerges from the interaction with the environment ([6, 7], see section 1.2.2).

The problem of finding evidence for or against the Lévy hypothesis is further complicated by the fact that the animals' step lengths have to be estimated from imperfect discretely sampled data giving only the pattern but not the process of the

movement [42, 43, 44]. This is done either by definition and analysis of turning points in the recorded trajectory [19, 9], or by only recording the animals' position at the turning points when they are well defined, e.g., as landing points of a foraging sea bird [6]. In the first corresponding studies of experimental data [6, 76] after Lévy dynamics were introduced into foraging theory [77], the tail of the step length histogram was compared to straight lines in log-log plots to find power-laws in the distribution. This has been shown to be unreliable irrespective of the binning method used for the histogram [86, 87], instead to reliably distinguish a power-law tail from, e.g., an exponential tail, maximum likelihood estimation[7] has been shown to be necessary [8, 87].

Experimental evidence [6, 84, 10, 68, 76] supporting the weak biological Lévy hypothesis were challenged by refined statistical data analyses [8, 87, 88, 25].

While in most analyses which claimed to have found Lévy walks the null-hypothesis was a Wiener process with normal diffusion, this comparison is questionable: a variety of mechanisms (see section 1.2) may naturally lead to different foraging dynamics on different length and time scales, e.g., individuality of animals [25, 24, 26], an intermittent switching between quasi-ballistic persistent dynamics and localized search modes [88, 48], or the averaging over non-negligible quantities like the time of day [68]. In section 1.3 we have seen that models like the reorientation models (section 1.3.3) and especially composite random walks (section 1.3.5) arise quite naturally from many of these environmental factors. As ignoring these mechanisms can lead to spurious power laws [8, 87], it is important to look for the reasons of the occurrence of non-trivial distributions, e.g., animals switching between different search modes. Only with this additionally gained knowledge is it then possible to effectively try to answer the biological Lévy hypothesis by excluding that factors other than search efficiency are the reason for the observed movement patterns.

For more elaborate movement models the velocity autocorrelations play a large role. Lévy flights and Lévy walks generate trivial (induced) functional forms for the velocity correlations [89, 90]. Accordingly, experiments testing the biological Lévy hypothesis have focused on probability distributions, not on correlation

[7] We will use a similar technique in section 2.2.2 to reliably distinguish between models describing experimental velocity distributions.

decay [6, 84, 10, 68]. In section 2.2.4 we will find an example of a change in autocorrelations induced by changes in the environment, giving another hint that the Lévy walks, which are inflexible with respect to the autocorrelations, are difficult to reconcile with data from experiments on the movement of foragers.

Although the evidence for Lévy walks as foraging strategies seems to be getting weaker and weaker [69, 91, 8, 87, 88, 25], the lure of the (weak) biological Lévy hypothesis as a way to explain experimental data is still present [68, 85]. In some cases the reason for the interpretation of movement data as Lévy walks is that they were preferred over a limited variety of alternative stochastic models (e.g. by comparing only to Brownian motion), which match even worse. This preference is seen as evidence for the Lévy flight hypothesis despite the fact that some other models would give a much better explanation of the data. For example in the case of [68] the seeming similarity to Lévy walks is very likely to be explained by a bistable day-night cycle for the off-shelf shark movement. Therefore a bistable model or an approximation of the switching by a composite random model (see section 1.3.5) would be more appropriate than either a Lévy walks or a Wiener process.

In summary, while the fundamental question 'What is the mathematically most efficient search strategy of foraging organisms?' has been studied in detail (see above), the mathematical Lévy flight hypothesis describes only one case of a variety of foraging situations. Since its necessary conditions are quite restricting it does not capture the full complexity of a biological foraging problem [74], which incorporates both the dependence of foraging on 'internal' conditions of a forager as well as 'external' environmental constraints (see section 1.2). While the biological Lévy flight hypothesis has been useful by renewing the interest in cooperation between biologists and the stochastic processes community, its use for modelling real animals does not seem to hold up to initial expectations.

A crucial problem is how dispositions of a forager like memory [34], sensory perception [30] or individuality [25, 24, 26] as well as properties of the environment [19, 65, 83, 10, 18, 68], can be tested in a statistical foraging analysis [71, 78, 79, 80, 74]. Especially for data obtained from foraging experiments in the wild, it is typically not clear to what extent extracted search patterns are determined by forager dispositions, or reflect an adjustment of the dynamics of organ-

isms to the distribution of food sources and the presence of predators [10, 68, 80]. This problem can be addressed by statistically quantifying search behaviour in laboratory experiments where foraging conditions are varied in a fully controlled manner [19, 68]. One such experiment has been performed by Ings and Chittka [1, 92], who studied the foraging behaviour of bumblebees with and without different types of artificial spiders mimicking predators. In the following chapter we will examine the resulting experimental data in order to gain insight into the effect of the environment on the movement patterns of foraging bumblebees.

Chapter 2

Bumblebee Flights under Predation Threat

In nature the interplay of a variety of factors, ranging from food source distributions and other spatial inhomogeneities in the environment to sensory capabilities and memory of the forager, as described in section 1.2, make it very hard to analyse foraging data. An important part before one can attempt to build concrete foraging models is to figure out which of those environmental factors have a large influence over foraging behaviour.

In the following two chapters we analyse experimental foraging data of bumblebees under two different aspects. The experiment will give us the opportunity to examine the search behaviour of bumblebees in a well-defined environment (see section 2.1). The goal of this chapter is to analyse the effect that predators have on the bumblebee flights. Therefore artificial predators have been introduced into a foraging arena as a controlled environmental variation, such that the reaction of the bumblebees to the change can be analysed. The main questions are therefore, whether there are qualitative or quantitative changes in their flight behaviour depending on the presence or visibility of predators, in which statistical properties these changes manifest themselves, and what we can say about learning and memory of bumblebees. We will also look at the experimental data in the context of the Lévy Hypothesis, although the experimental data is not suitable to directly test it – mainly because of the boundedness of the experiment due to the confinement

of the bumblebees in the arena (compare section 1.4.2). Nevertheless, the analysis of the data will give us some indication regarding the applicability of the Lévy Hypothesis.

In the next chapter we will then step away from the description of the interaction with flowers and predators and construct a bumblebee flight model from the experimental data focussing on the search flights between flower visits.

We start this chapter with an introduction to the experiment in section 2.1, and a first overview of the data by examining the position probability density function (PDF) in section 2.2.1. The main part of the examination of the bumblebee data then consists of the analysis of the velocity distributions in section 2.2.2 and the velocity autocorrelations in section 2.2.4 with respect to their variation under predation threat. The former also includes a discussion of the individuality of bumblebees. We will then distinguish different spatially localised effects of the presence of artificial predators on the foraging behaviour of the bumblebees in section 2.2.3. In section 2.2.5 we aggregate the gained knowledge about the bumblebee flights in a model, which gives a qualitative explanation for the observed velocity autocorrelations. In section 2.3 we connect the results of our analysis with the biological Lévy flight hypothesis (see section 1.4.2) and finish by summing up the chapter in section 2.4.

2.1 Set-up of the Bumblebee Experiment

In the analysed experiment [1] 30 bumblebees (*Bombus terrestris*) were trained to forage in a flight arena with side lengths of $l_x = 1\,\mathrm{m}$, $l_y = 0.72\,\mathrm{m}$ and $l_z = 0.73\,\mathrm{m}$. The flight arena included a 4×4 grid of artificial flowers on one of the walls. Each of the 16 flowers (see Fig. 2.2) consisted of a landing platform, a yellow square floral marker and an artificial feeder: a replenishing food source offering sucrose syrup at a rate of $1.85\,\mu\mathrm{l}/\mathrm{min}$ [1]. Figure 2.1 shows a diagram of the arena together with data from a typical flight path of a bumblebee. Given the small size of the foraging arena compared to the space available to free flying bumblebees, the flights should be interpreted as the behaviour of bumblebees when foraging in a patch of flowers and not as free flights in an unconstrained environment. The influence of the boundedness of the flight arena on the bumblebee behaviour is

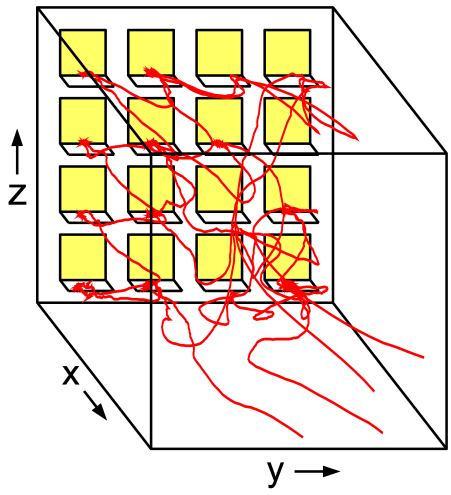

Figure 2.1: **Diagram of the foraging arena.** Included is a part of the flight trajectory of a single bumblebee. The bumblebees forage on a grid of artificial flowers at one wall of the box. While being on the landing platforms, the bumblebees have access to a food supply. All flowers can be equipped with spider models and trapping mechanisms simulating predation attempts.

discussed in section 3.3.2. However, the main confinement of the bumblebees does come from the tendency to return to the food sources, while the walls of the flight arena are not as important (compare section 2.2.4).

The 3D flight trajectories of the bumblebees were tracked by two cameras with a temporal resolution of $\Delta t = 0.02\,\mathrm{s}$. The individual bumblebee behaviour was recorded by letting them fly, one at a time, in the flight arena. Each bumblebee was approximated as a point mass with a spatial resolution of $0.1\,\mathrm{cm}$: internal degrees of freedom were not recorded. The positions of the bumblebees were estimated by the centre of mass of all image pixels corresponding to the bumblebee via background subtraction.

The bumblebees vary individually, e.g., by mass, age and size, measured for instance by the thorax widths of the bumblebees: they have a mean width of $5.6\,\mathrm{mm}$ with a standard deviation of $0.4\,\mathrm{mm}$. Therefore the data analyses of this chapter have been done for each bumblebee separately unless specified otherwise below; the individuality is addressed explicitly in section 2.2.2.

In 7 experimental stages the bumblebees are trained to feed, and their reaction

Figure 2.2: Image of a single artificial flower (left) consisting of a landing platform, a yellow square floral marker, a replenishing food supply and a trapping mechanism. The trapping mechanism is present on all flowers, but it is only activated on flowers which are additionally marked with a 3D spider model. To its right: camouflaged crab spiders in situ, waiting on flowers to attack foraging bumblebees. Photos by Thomas C. Ings.

to — and memory of — the presence of artificial spiders is recorded. The artificial spiders are mechanical traps, which squeeze and release the bumblebees together with life-sized ($l = 12\,\mathrm{mm}$) models of crab spiders (*Misumena vatia*), which simulate predation attempts. The experimental stages are:

(1) *Pretraining* Feeding without predator threat.

(2) Training Artificial spiders are introduced.

(3) Neutral Feeding with no spider models.

(4) *Mid-term Memory Test* Spider models visible, but trapping mechanism is blocked.

(5) Reinforcement Training Spiders with active traps.

(6) Remotivation a day later Feeding with no spiders.

(7) *Long-term Memory Test* Same as Mid-term Memory Test.

The stages used for our analysis are (1) *Pretraining*, (4) *Mid-term Memory Test* and (7) *Long-term Memory Test*, to see whether the bumblebees learn from predation attempts and adapt their movements under the threat of predation. The

Training stages are not useful for an analysis of predation *threats* because one would only measure the agitated flights of the bumblebee after being trapped.

The bumblebees were trained on two kinds of artificial spiders: half of them on easily visible spiders and half of them on cryptic spiders [1]. In vivo, the spider type which was emulated here is able to camouflage itself by adapting its colour to the surroundings (see Fig. 2.2). However for our analysis this difference between the two types turned out to be irrelevant. We did not see any major differences between the two groups of bumblebees in any observed variable. Therefore we will not differentiate between them in the following.

More details on the experimental set-up can be found in [1]. The possibility to change parts of the environment while keeping all other conditions constant is the main advantage of this and related [92, 93] experiments. This is in strong contrast to in situ experiments which have the advantage to capture the behaviour of animals in their natural surroundings. However, they rarely offer the opportunity of completely controlled modifications to the environment, since the number of influential factors is usually large (see section 1.2).

2.2 Analysis of Bumblebee Flights

In our analysis of the experimental data we examined the velocity distributions and autocorrelations in the different spatial directions for all stages of the experiment. As we are only interested in the flight behaviour, we excluded all data corresponding to crawling behaviour of the bumblebees on the artificial flowers by removing all data within 1 cm of each landing platform, leaving from 2000 to 15000 data points (average: 6000) per bumblebee for each stage.

For our analysis, the experimental flight data was classified distinguishing data near flowers and data away from flowers: for that purpose roughly cubical *flower zones* around the artificial flowers have been defined – see section A.2.2 for details. While we will mainly be using the whole data set in this chapter, in chapter 3 the data inside flower zones has been removed in order to analyse the foraging search behaviour, excluding the interaction of the bumblebees with the food sources.

Given that the experimental data contained measurement errors, gaps and other artefacts, e.g., position data of bumblebees when crawling on flowers, the

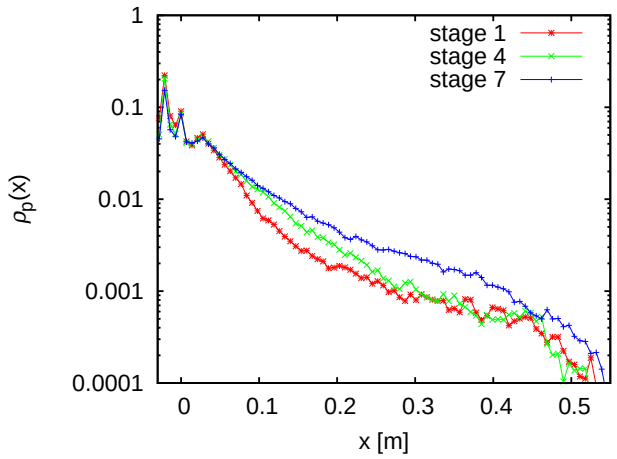

Figure 2.3: Semi-logarithmic plot of estimated x-position distributions for all experimental stages. The flower platforms extend from the flower wall at $x = -0.03\,\mathrm{m}$ up to $x = 0.04\,\mathrm{m}$ into the foraging arena.

data had to be cleaned as described in section A.2. Trajectories were split at larger gaps or when visiting the flowers (see section A.2.3), to exclude correlations induced by flower visits. For individual bumblebees an average of 51 search trajectories between flower visits have been sampled and analysed.

In total ≈ 170000 data points were available for each experimental stage after cleaning the data – in cases where complete gap-less trajectories from flower to flower are needed this reduces to ≈ 135000 data points.

2.2.1 Position Distributions

To get an overview of the data, let us start with an examination of the distribution of bumblebee positions. In all directions the positions concentrate near the flowers and the position distributions decay from there with increasing distance (see Figs. 2.3 and 2.4). Figure 2.3 shows the dependence of the position-PDF ρ_p on the distance from the flower wall. While the exact functional shape of $\rho_p(x, y, z)$ is not easy to pin down, first differences between the experimental stages can be observed. The clearest difference is between the threat-less stage (1) and the memory test one day later (stage (7)): in all directions, the bumblebees' posi-

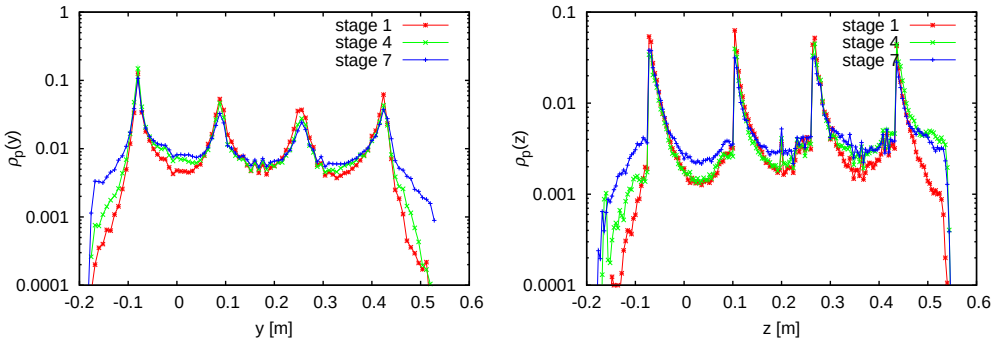

Figure 2.4: Semi-log plot of estimated y- and z-position distributions for all experimental stages.

tion density increases away from the flower positions in stage (7). A plausible explanation for this effect is that the bumblebees now have been trained on two occasions (stages (2) and (5)) with dangerous spiders and might therefore be more motivated to leave the flower patch and search for food elsewhere when they are again exposed to spider models. This is consistent with the increase of ρ_p near the other walls in figures 2.3 and 2.4. The same effect, although less pronounced, can be observed for the first memory test with predation threat in stage (4). It might seem that in stage (4) the bumblebees have not yet sufficiently trained on artificial predators to change their behaviour significantly, but we will see in sections 2.2.3 and 2.2.4 that they have already learned from the first training phase and adapted their flight patterns.

Fig. 2.4 also nicely demonstrates the asymmetry in the vertical z-direction: approaches to and from flowers and inspections of them happen dominantly from above which can also be seen in Fig. 2.5. The effects of the correlation of the flight direction in x- and z-direction due to starting and landing bumblebee flights will be discussed in section 2.2.4.

2.2.2 Velocity Distributions

While the distribution of bumblebee positions already shows some effect of predation threats on the bumblebees, we expect that the reactions to the artificial

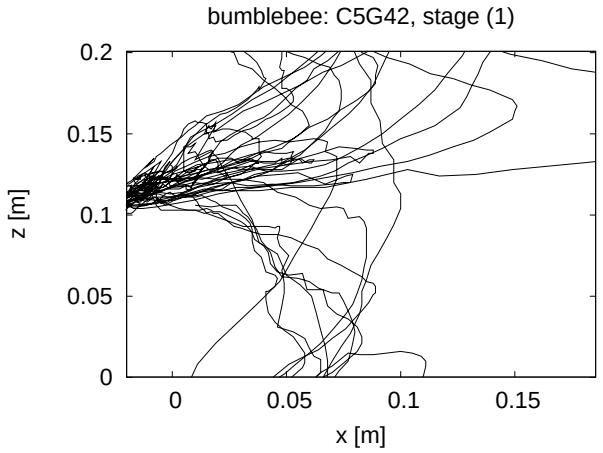

Figure 2.5: Flight trajectories at a single flower (on the left border) projected on x and z: bumblebees starting from and approaching flowers are fairly consistent in their direction.

spiders might show a clearer signal of changed behaviour through changes in the velocities. We therefore analyse the velocity distributions in the 3 qualitatively different directions for all the experimental stages. Since we can expect to find differences between individual bumblebees [1] (see p. 49f), we have to be careful and look at each bumblebee individually before combining the results. With the amount of data given, we have to make sure that our results are consistent and have a convincing interpretation.

The velocity distributions vary for the different spatial directions due to asymmetries induced by physical and biological constraints as well as the spatial arrangement of the flowers. Figure 2.6 shows a typical normalized histogram of the horizontal velocities parallel to the flower wall (cf. y-direction in Fig. 2.1) for a single bumblebee showing a characteristic peak at low velocities. Direct fitting of distributions on the histogram and a visual comparison with some assumed distribution was shown to be unreliable [87], as is illustrated by Fig. 2.6: only the power law and the Gaussian distribution can be ruled out by visual inspection. However, the Gaussian mixture and an exponential function appear to be equally likely. In the following, we therefore use maximum likelihood estimation for a number of candidate distributions to obtain the optimal parameters for each can-

didate and then compare the different distribution types by their weights using the Akaike and Bayesian information criteria [94, 95].

Maximum Likelihood Estimation

In order to fit candidate distributions to the experimental velocity data we estimated their parameters by maximising the likelihood of each distribution [8, 87]. Our candidate distributions are:

a) Exponential: $\rho_\lambda(v) = ce^{-\lambda|v|}$,

b) Power law: $\rho_\mu(v) = c\,|v|^{-\mu}$,

c) Normal distribution with zero mean: $\rho_\sigma(v) = N_\sigma(v)$,

d) Mixture of two normal distributions: $\rho_{a,\sigma_1,\sigma_2}(v) = aN_{\sigma_1}(v)+(1-a)N_{\sigma_2}(v)$,

where $N_{\sigma_i}(v) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{v^2}{2\sigma_i^2}}$, $i = 1, 2$, and $0 \leq a \leq 1$.

Given a set of measured velocities $D = \{v_1, v_2, \ldots, v_n\}$ and a probability density function $\rho_\lambda(v)$, where λ is a vector of k parameters, the *log-likelihood* of the probability density function for a finite resolution of the data ($\Delta v = 5\,{}^{\mathrm{cm}}/_{\mathrm{s}}$) simplifies to

$$\ln L(\lambda|D) = \sum_{v_j \in D} \ln P_\lambda(v_j) = \sum_{b \in bins} h[b] \ln \int_{min(b)}^{max(b)} \rho_\lambda(v) dv \qquad (2.1)$$

where $h(b)$ is the observed frequency in bin b.

For each candidate distribution $\rho^i_{\lambda_i}$, $i \in \{1, \ldots, 4\}$, we locally maximised the log-likelihood $\ln L_i$ with relation to λ_i with the *Nelder-Mead* downhill simplex algorithm [96].[1] We then used a Monte Carlo method to search for global maxima. Figure 2.6 shows a typical result of fitted distributions to data of a single bumblebee.

[1]This algorithm was chosen as it is quite fast, so that we could sample for many starting parameters with Monte Carlo.

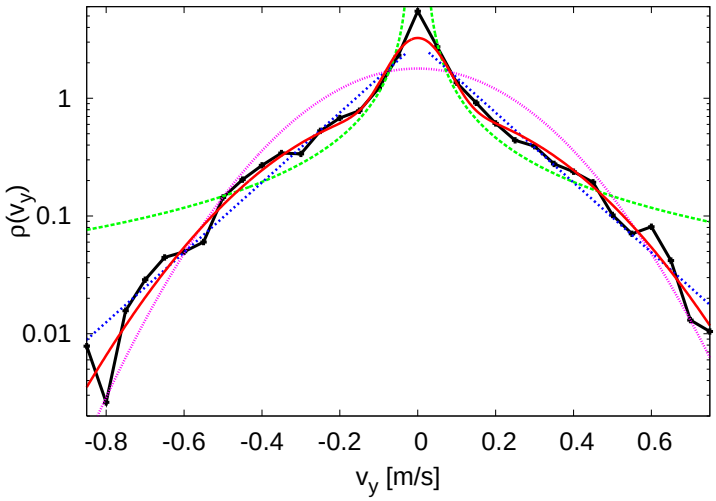

Figure 2.6: Estimated velocity distributions: Semi-logarithmic plot of the normalised histogram of velocities v_y parallel to the y-axis in Fig. 2.1 (black crosses) for a single bumblebee in the spider-free stage (1) together with a Gaussian mixture (red line), exponential (blue dotted), power law (green dashed), and Gaussian distribution (violet dotted), fitted via maximum likelihood estimation.

Information Criteria

For checking which of the distributions fits best we used the Akaike information criterion [8]. We made sure that the results do not depend on the chosen criterion by also checking the Bayesian information criterion.

To find the preference between the different model distributions whose likelihoods L_i are maximised at λ_i^{max} the information criteria are

$$IC_i = -2\ln(L_i(\lambda_i^{max}|D)) + s(n)k_i \tag{2.2}$$

with $s(n) = 2$ for the Akaike information criterion and $s(n) = \ln(n)$ for the Bayesian information criterion as a penalty on the number of parameters k_i. The best model, denoted by $*$, is the one which minimises the information criterion $IC_* = \min_i(IC_i)$. The Akaike/Bayesian weights then give the preference of each

Table 2.1: Model weights and estimated parameters. Akaike and Bayesian weights both give preference to the mixture of two Gaussian distributions for v_y for most of the bumblebees. The weights are estimated individually for each bumblebee and their mean and standard deviation (in brackets) over all bumblebees are shown. Below the mean and standard deviation over all bumblebees of the individually estimated distribution parameters are given.

Model:	a) Expon.	b) Power law	c) Normal	d) Normal Mixture		
Akaike weight	0.00 (0.00)	0.00 (0.00)	0.04 (0.19)	0.96 (0.19)		
Bayesian weight	0.04 (0.18)	0.00 (0.00)	0.08 (0.26)	0.88 (0.30)		
Parameter X	λ	μ	σ	a	σ_1	σ_2
mean(X)	5.61	1.11	0.25	0.67	0.06	0.29
stdev(X)	1.07	0.16	0.03	0.13	0.04	0.03

model over the others as a probability

$$w_i = \alpha \mathrm{e}^{-(IC_i - IC_*)/2} \,, \tag{2.3}$$

where α normalises the weights to $\sum_i w_i = 1$. In our case, the choice of the information criterion makes no strong difference for the model selection in this experiment.

Of our list of candidate distributions the Gaussian mixture turned out to be best for all stages of the experiment independent of environmental parameters (see Table 2.1). With the Akaike information criterion the Gaussian mixture is chosen with a weight of over 95% for all bumblebees and all experimental stages. The Bayesian information criterion agrees with the Akaike information criterion on 90% of all data sets. For the other 10% it prefers a single Gaussian or an exponential distribution — however, these data sets turned out to be those with the least amount of data available.

The mixture of two Gaussian distributions can be biologically interpreted as two modes of flight: one for flying near flowers and one for 'free' search flights, which bears some resemblance to intermittent dynamics [65, 48, 74]. This has been verified by splitting the data into flights far from the flower wall vs. flights in the flower zones (see section A.2.2) and examining both data sets separately. With

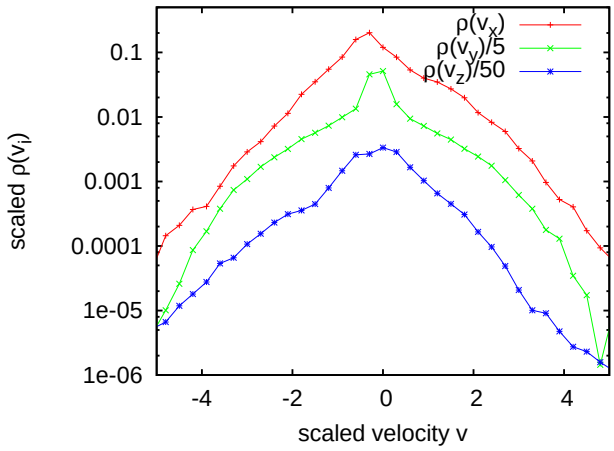

Figure 2.7: Semi-log plot of normalized velocity distributions in each direction in experimental stage (1). The velocity distributions have been scaled by their variance for each bumblebee before accumulating the data of all bees. ($\rho(v_y)$ and $\rho(v_z)$ are shifted down for better visibility.)

a growing risk of overfitting we could continue to increase the "zoo" of candidate distributions, e.g. by considering exponential mixtures. However, the Gaussian mixture is not only satisfying because we can explain why it is mixed, but also because the Gaussians are stable, which is consistent with an interpretation of the discretely measured flight steps as sums of substeps (see section 1.3.2).

The velocity distributions in the other directions, i.e. v_x and v_z can be seen in figure 2.7. For this comparison data from all bumblebees in stage (1) was used. To allow the use of all data, the velocities had to be scaled by the variance for each bumblebee (see below). Fig. 2.7 shows that while the y-velocity distribution is nicely symmetric, $\rho(v_x)$ and $\rho(v_z)$ have asymmetries induced by gravity for v_z respectively by the difference between flying towards the flower wall and flying away for v_x. For an analysis of the causes of the functional shape of these distributions a more comprehensive bio-mechanical model for starting and landing on flowers would be needed. It is interesting that $\rho(v_x)$ is consistently exponential for negative velocities v_x over all individual bumblebees, that is, for flights towards the flower wall – however, the reason is unclear.

Table 2.2: Weights and estimated parameters of the Gaussian mixture for the different experimental stages. Weights and parameters are estimated for each bumblebee. Shown are the mean over all individuals and the standard deviation (in brackets). The mixture of two Gaussians is the best fit in all stages. In the parameters of the distribution we observe no significant effect of the threat of predators on the bumblebees.

Stages	Akaike w.	Bayesian w.	a	σ_1	σ_2
(1) No Risk	0.97 (0.15)	0.93 (0.23)	0.64 (0.11)	0.06 (0.02)	0.29 (0.03)
(4) Pred. risk	0.99 (0.04)	0.90 (0.27)	0.68 (0.13)	0.06 (0.02)	0.29 (0.02)
(7) Risk+1day	0.89 (0.29)	0.80 (0.38)	0.72 (0.16)	0.07 (0.07)	0.30 (0.03)

Variability between Individual Bumblebees

Looking at the parameters of the mixture of two normal distributions, estimated for $\rho(v_y)$ for different bumblebees, we found that there are strong variations between individuals. This is interesting as heterogeneous populations have been proposed as one mechanism by which anomalous diffusion at the population level can be generated even if the individual behaviour is normal [26, 25, 24].

Surprisingly, by comparing the best fits to these distributions for the different stages of the experiment, we could not detect any differences in the velocity distributions between the spider-free stage and the stages where artificial spider models were present, as is shown in Table 2.2. The parameters of the Gaussian mixture vary between individual bumblebees, as can be seen in figure 2.8, but there is no systematic change due to the presence of predators. The same is true for the distributions of v_x and v_z. However the observations by the experimentalists suggested that the behaviour of the bumblebees changes when threatened by predators. This means that the changes should be measurable by observables other than the velocity distributions.

Since the bumblebees vary in their weight and size as measured by their thorax width, it would be reasonable to assume that their size differences are the reason for their different speeds. Yet, no such relation has been found in this experiment as can be seen, e.g., in Fig. 2.9. This is consistent with a previous analysis of the experimental data, which found no effects of bee size and age on flight parameters

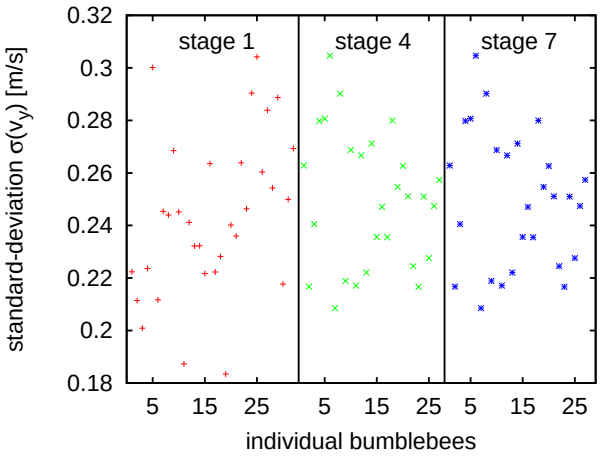

Figure 2.8: Individual variation of standard deviations of the distribution of v_y between bumblebees. No systematic change between experimental stages is observed for any velocity distribution.

and learning and memory of the bumblebees [1]. This analysis of a possible size-dependence used only data outside the flower zones (see section A.2.2) to exclude any complications due to variations in the time spent near flowers. Including the data near flowers also did not show any dependence.

Quantile-Quantile Plots

The information criteria only give the preference between the candidate distributions. However, they do not inform us if the best of the candidates is actually a good model: if all of the candidates are far off the real distribution, the Akaike weights (and Bayesian weights) could highlight one of them as the best of the poor fits. As a supplementary qualitative test to which extent the estimated distribution with the largest Akaike weight deviates from the data over the whole range variables, we use Quantile-Quantile (Q-Q) probability plots.

By using 20 surrogate data sets of the same size as the real data, generated by drawing i.i.d. random numbers from the estimated distribution, we looked for deviations from the model larger than those expected because of stochastic variations due to the finite quantity of the data. Figure 2.10 shows the typical

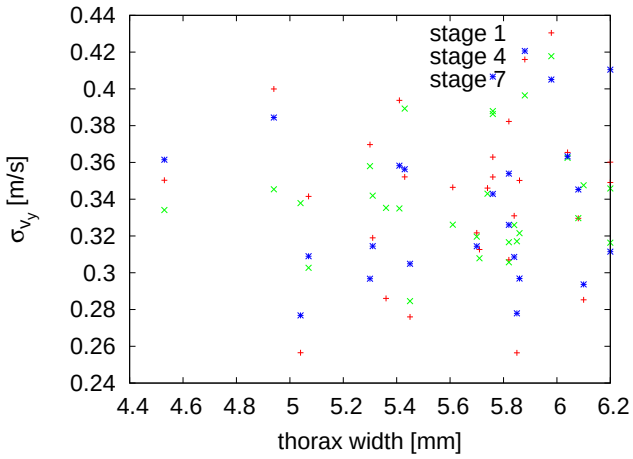

Figure 2.9: Standard deviations of v_y-distributions for each bumblebee depending on their thorax widths for flights outside the flower zones. No dependence of the velocities on the thorax widths has been found.

result for a single bumblebee: the fluctuations of the QQ-Plot of the data lie in the typical range one would expect for the given amount of data.

For comparison, figure 2.11 shows a Quantile-Quantile plot for a non-matching distribution, in this case a normal distribution with the correctly estimated parameters, for the aggregated data of all bumblebees. The distributions for the bumblebees have been normalized by the standard deviation before aggregation as a result of individual differences (see below).The strong departure from the diagonal indicates a clear mismatch of the normal distribution and the experimental data.

2.2.3 Local Behavioural Changes under Threat

One way to examine the effect of the presence of predation threats is via the probability of a bumblebee to fly directly in front of a flower. The change in the bumblebees' behaviour can be analysed by computing the difference between the position densities at stage (1) and (4) as a function of the positions parallel to and

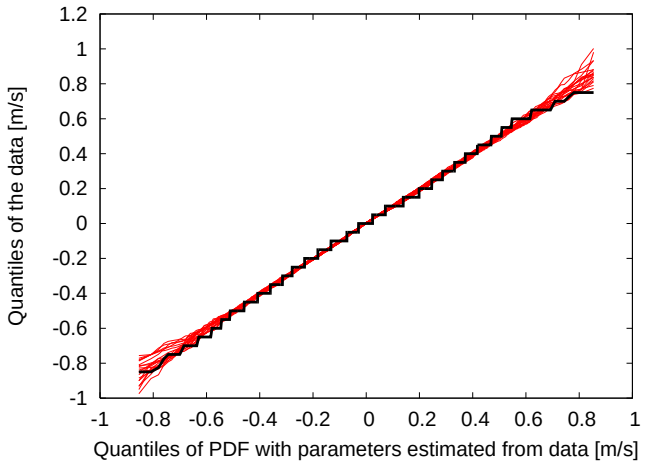

Figure 2.10: Quantile-Quantile plot of v_y against a Gaussian mixture as the best fit: quantiles of v_y (in $^{\mathrm{m}}/_{\mathrm{s}}$) of a single bumblebee against quantiles of an estimated mixture of two Gaussians. An ideal match would yield a straight line. The steps in the real data are a discretisation effect of the data. The red lines show 20 surrogate data sets of the same size.

near ($x < 5\,\mathrm{cm}$) the flower wall

$$\Delta\rho_p(y,z) = \rho_p^{(4)}(y,z) - \rho_p^{(1)}(y,z)\,. \tag{2.4}$$

Figure 2.12 shows that near the flowers, the position-PDF decreases when introducing a predation threat.

For a more detailed analysis of the local effects, it is useful to switch to relative coordinates where the origin is always the position of the flower which is closest to the bumblebee. Here we want to focus on the change in the behaviour from threat-less foraging to flights under predation threat and not on the preferences between individual flowers, we therefore treat all flowers as equivalent:

$$\Delta\rho_p(y_{rel},z_{rel}) = \rho_p^{(4)}(y_{rel},z_{rel}) - \rho_p^{(1)}(y_{rel},z_{rel})\,, \tag{2.5}$$

where the positions (y_{rel}, z_{rel}) are relative to the nearest flower centre. The changes thus extracted from the experimental data are shown in Fig. 2.13. Here data from all the individual bumblebees was accumulated: differences between individuals

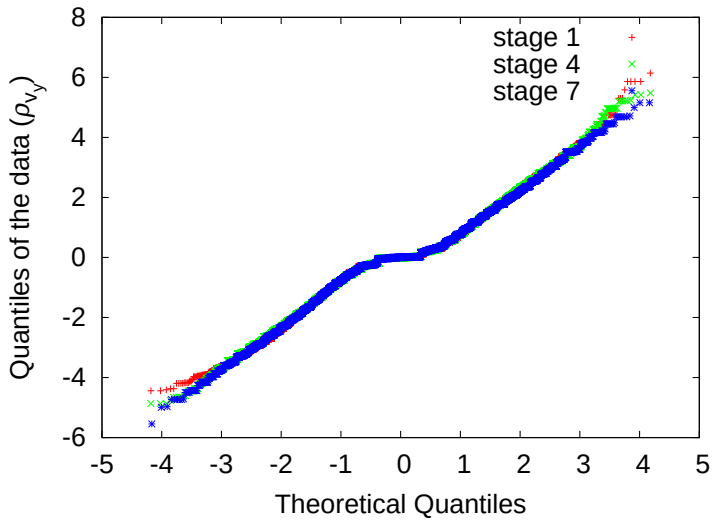

Figure 2.11: Quantile-quantile plots of v_y against a normal distribution using data of all bumblebees. The distributions for the individual bumblebees have been normalized by the standard deviation before aggregation. The mismatch of data and distribution is visualized by the departure from a diagonal and disqualifies the normal distribution as a valid model. The introduction of predation threat in stages (4) and (7) has no effect on velocity distributions of the bumblebees.

have been only found in the velocity distributions, but we did not find a strong variation in the position-PDFs. Two different types of behaviour can be seen here: First, there is a small increase in the amount of hovering, i.e. inspection flights near the flower platform when a spider model is present [93, 97], which is consistent with Ref. [1]. This increased hovering occurs only at flowers occupied by spiders: see below for an analysis of data in front of spider-free flowers (in Fig. 2.14). Second and more important is the local minimum representing the avoidance of flowers infected by spiders. This effect is strongest in an area around 3 cm above the flowers, because the flowers are predominantly approached from above.

While the increased hovering occurs only on those flowers in stage (4) which have spiders on them, the avoidance behaviour in stage (4) is also present in front of spider-less flowers. This can be seen in figure 2.14, where the comparison of the position-PDF between stages (1) and (4) is restricted to data in front of flowers

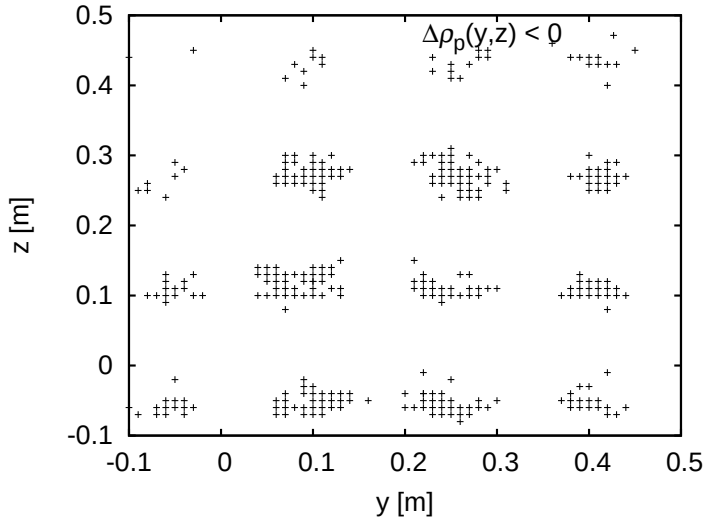

Figure 2.12: Regions in position space avoided under predation threat (stage (4)) relative to stage (1), i.e., where $\Delta\rho_p(y,z) < 0$. The marked regions correspond to positions near flowers.

without spiders. As stage (1) is spider-less anyway, this restriction only affects data from stage (4). The decrease of the position-PDF near spider-less flowers in stage (4) indicates that the bumblebees adapted to the predation threat by learning from the training stage (2) before. A purely instantaneous reaction to perceived spiders could only explain changes at the flowers with spiders. Notice that the experiment does not tell whether the adapted avoidance behaviour is completely new: it could be that the bumblebees just learned that this patch of flowers is dense in spiders and therefore switch to a more careful search mode, which already existed prior to the learning.

Even while the bumblebees have been shown above to reduce the time spent above *all* flowers after learning of the predation threat in the training phase (2), the avoidance is strongest for spider-occupied flowers. This avoidance behaviour affects not only flights near the flower wall but can still be detected further away from it. For stage (4), figure 2.15 compares the differences in the probabilities $\rho_p^{4,safe}$, $\rho_p^{4,spider}$ to fly in front of safe flowers without artificial spiders to those in front of flowers with spiders as a function of the distance to the flower wall. This

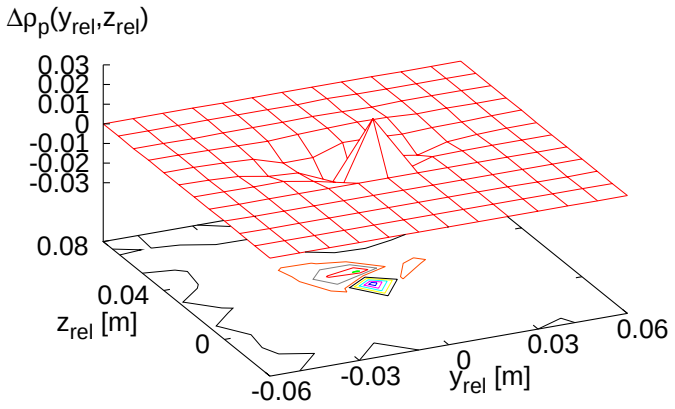

Figure 2.13: Predator avoidance of bumblebees at flowers, Eq. (2.5), extracted from the experimental data. Increased hovering behaviour (from stage (1) to stage (4)) in front of a flower is represented by the positive spike directly at the flower centre, while the negative region behind this spike reflects the general avoidance of flying near flowers in the presence of threats (stage (4)).

relative change is given by

$$\mathrm{change}_{rel}(x) = \frac{\rho_p^{4,spider}(x) - \rho_p^{4,safe}(x)}{\rho_p^{4,safe}(x)}. \tag{2.6}$$

Even up to 30 cm away from the flower wall the bumblebees are observed as less likely to be in front of flowers with spiders than in front of spider-free flowers. For larger distances there is not enough data available, making the comparison less and less reliable.

Figure 2.16 directly compares the corresponding histograms over the distance x to the flower wall. With the exception of small x, where relatively less flying space is available since the bumblebees' movement is restricted by the flower platforms, the histograms show a roughly exponential decay away from the flower wall, and for $10\,\mathrm{cm} < x < 30\,\mathrm{cm}$ the preference of flying in front of spider-free flowers is again visible.

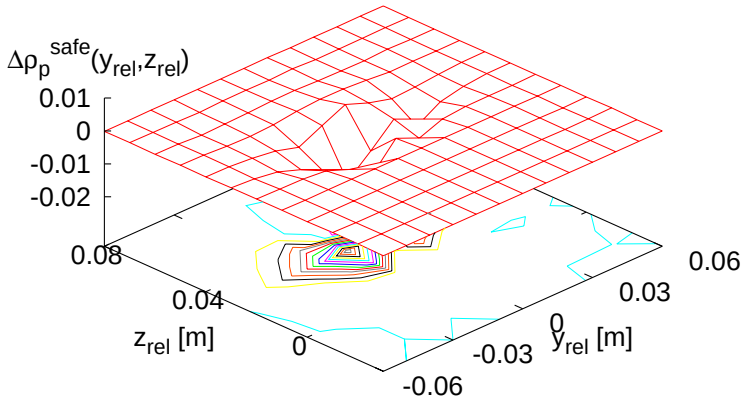

Figure 2.14: Avoidance of spider-free flowers in stage (4) in comparison to stage (1), as extracted from the experimental data. While not as strong as in front of flowers with spiders, the avoidance of spider-free flowers points towards an adaptation of the bumblebees to the predation threat by learned behaviour. No increased hovering is found here – this behaviour occurs solely near flowers occupied by spiders (compare Fig. 2.13).

In total, while other effects occur due to predator presence (e.g. the hovering detected above), the dominant effect on the bumblebee positions is the relative flower avoidance, which, while strongest at flowers with spiders, also affects the behaviour at spider-free flowers by learning.

2.2.4 Velocity Autocorrelations

As the velocity distributions were not affected by the environmental change, we also examined the autocorrelation function $v^{ac}(\tau)$ of the flight velocities

$$v^{ac}(\tau) = \frac{\langle (v(t) - \mu)(v(t+\tau) - \mu) \rangle}{\sigma^2} \tag{2.7}$$

for flights from flower to flower. The autocorrelation has been computed by averaging over all bumblebees and over time in all flights that are complete from

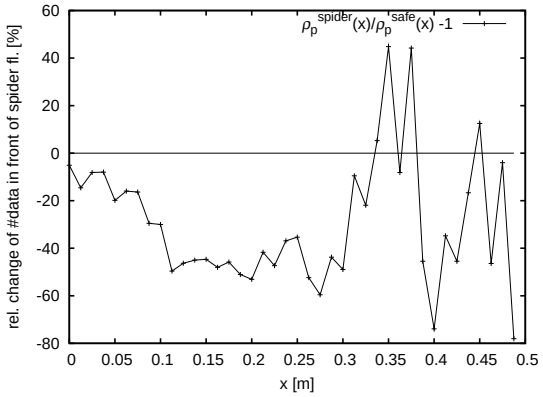

Figure 2.15: Relative difference of position x-PDFs (both in Stage (4), see Fig. 2.13) in front of flowers with vs. without spiders: $\mathrm{change}_{rel}(x)$ is given relative to the probability density function in front of flowers without spiders (see Eq. (2.6)). It shows the avoidance of flowers with predators even up to 30 cm away from the flower wall.

starting on one flower to landing on the next. We exclude flights containing too long gaps (see section A.2.3) and weighted with the total amount of data available for each time interval.

Figures 2.17 and 2.18 show the velocity autocorrelation in the x- and y-directions for different stages of the experiment. In the x-direction (Fig. 2.17) perpendicular to the flower wall the velocity autocorrelation has no qualitative dependence on the predation risk: It is always anti-correlated for times around $0.5\,\mathrm{s}$, which is due to the tendency of the bumblebees to quickly return to the flower wall. For longer times this effect of returning to the flower wall still induces some anti-correlation, although it gets quickly weaker for larger τ.

However, the flights with long durations between flower visits become more frequent for stages (2) and (3) where the bumblebees were exposed to predation risk compared with stage (1) (inset of Fig. 2.17). This is also reflected in a small shift of the global minimum in the correlations for stages (2) and (3) away from the origin.

In the vertical z-direction the autocorrelation is similar to the one in the x-direction. This relation is due to the correlations between x and z for the starting

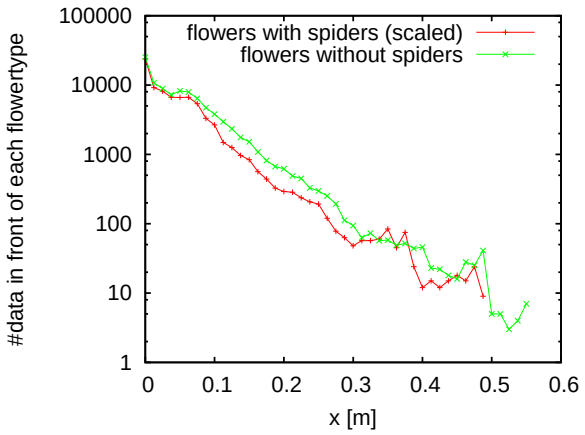

Figure 2.16: Histogram of bumblebee positions under predation threat (stage (4)) in x-direction. Shown is the position histogram in front of dangerous flowers (with spiders, red) and safe flowers (without spiders, green). The former is scaled by a factor of 3 for comparison, to account for the different number of dangerous and safe flowers.

and landing phases of the flights: bumblebees leaving a flower lead to a simultaneous increase in both position coordinates as can be seen in Fig. 2.5.

An important result is that v_y (parallel to the flower wall) is anti-correlated in the presence of spiders for $0.7\,\mathrm{s} < \tau < 2.8\,\mathrm{s}$, while for the spider-free stage it remains positive up to $1.7\,\mathrm{s}$ (Fig. 2.18). The autocorrelation function varies between bumblebees due to the limited amount of data and/or due to differences between individuals. We therefore re-sampled the result by leaving the data of each single bumblebee out (jackknifing). The re-sampling (inset of Fig. 2.18) confirms that the differences in the autocorrelation of v_y are due to the presence of spiders.

The velocity autocorrelations are consistent with a more careful search: When no threat of predators is present, the bumblebees forage more systematically with more or less direct flights from flower to flower, arching away from the flower wall. Under threat the trajectories become longer and the bumblebees change their direction more often in their search for food sources, rejecting flowers with artificial spiders. This reversing of directions generates the anti-correlations in the y-direction. By looking at the flight time distributions, i.e. the distribution of

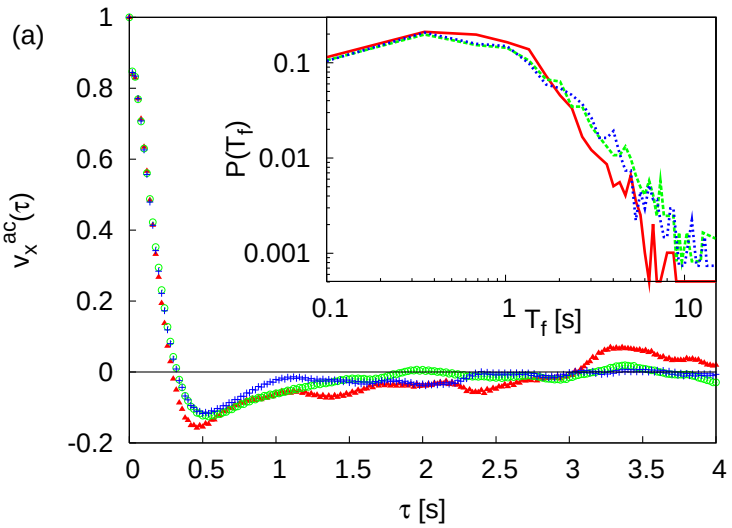

Figure 2.17: Autocorrelation of the velocities at different experimental stages: without spiders (red triangles), under threat of predation (green circles), and under threat a day after the last encounter with the spiders (blue crosses). In the x-direction the velocities are anti-correlated for small times ($\approx 0.5\,\mathrm{s}$) due to short flights from one flower to a nearby flower back at the flower wall. Inset: the distribution of flight-durations T_f for each stage shows a corresponding maximum for these short jumps. Under threat of predation (dotted) long flights become more frequent.

time intervals between starting on one flower and landing on another, one can rule out the possibility that the main features of the correlation functions are induced by the boundedness of the flight arena: in the inset of Fig. 2.17, all flight time distributions display maxima around $T_f \approx 0.5\,\mathrm{s}$ suggesting that times below $\simeq 2\,\mathrm{s}$ are primarily related to flights between flowers. Boundary effects are only evident for flight times that fall within the tail of the distributions. The anti-correlations in the y-direction thus cannot be induced by the walls but are generated by a reversal of directions at flowers under predatory threat. For the x- and y-direction, the return to the flower wall is responsible for the anti-correlation at small delay times, not the opposite wall, which is too far away to have a significant effect.

A simple model describing this mechanism is given in section 2.2.5 below. For the x-direction, the return to the flower wall is responsible for the anti-correlation

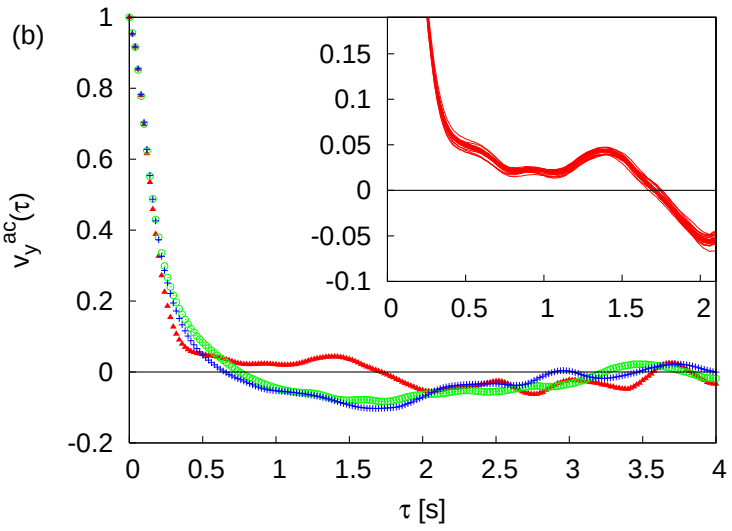

Figure 2.18: The autocorrelation of v_y shows the effect of the presence of spiders on the flight behaviour of the bumblebees. The inset shows the re-sampled autocorrelation for the spider-free stage in the region where the correlation differs from the stages with spider models, which confirms that the differences are due to predatory threat.

at small delay times.

The distributions of the flight durations T_f themselves also change under threat of predation: the inset of Fig. 2.17 shows the roughly exponential tails of $P(T_f)$ and that the flights with long durations between flower visits become more frequent for stages (4) and (7) compared to the pre-training stage (1).

2.2.5 An Effective Potential Model for the Dynamics of Threatened Bumblebees

The avoidance of spider-infected flowers seen in sections 2.2.3 and 2.2.4, together with the spatial switching of flight modes (see section 2.2.2), can be modelled by

the Langevin Equation:

$$\begin{aligned} \frac{d\mathbf{r}}{dt}(t) &= \mathbf{v}(t) \\ \frac{d\mathbf{v}}{dt}(t) &= -\eta\mathbf{v}(t) - \nabla U(\mathbf{r}(t)) + \boldsymbol{\xi}(\mathbf{r},t)\,, \end{aligned} \tag{2.8}$$

where η is a friction coefficient and $\boldsymbol{\xi}$ a vector of white Gaussian noise with standard deviation depending on the flight mode as a function of the position, $\boldsymbol{\xi}(\mathbf{r},t) = \chi_{\mathrm{fz}}(\mathbf{r})\boldsymbol{\xi_1}(t) + (1-\chi_{\mathrm{fz}}(\mathbf{r}))\boldsymbol{\xi}_2(t)$. Here $\mathbf{r} = (x,y,z)^\top$ is the position of the bumblebee at time t, $\chi_{\mathrm{fz}}(\mathbf{r})$ is the indicator function of the feeding zone, which is equal to one whenever the bumblebee is in the cube around a flower as defined before, and $\boldsymbol{\xi}_i\,,\ i=1,2$ is Gaussian noise with two different variances. The potential U models an interaction between bumblebee and spider in the form of a repulsive force exerted by the spider onto the bumblebee, for which we assume that the potential maxima are located near infected flowers.

When the mechanism generating the correlation functions shown in figures 2.17, 2.18 is not the focus of the investigation, it suffices to consider a reduced version of Eqs. (2.8) in the form of the *effective* Langevin equation

$$\frac{d\mathbf{r}}{dt} = \chi_{\mathrm{fz}}(\mathbf{r})\boldsymbol{\zeta}_1(t) + (1-\chi_{\mathrm{fz}}(\mathbf{r}))\boldsymbol{\zeta}_2(t)\,. \tag{2.9}$$

This equation describes the spatially varying hovering and search modes by using noise $\boldsymbol{\zeta}_i\,,\ i=1,2$, which models the impact of the potential U together with the noise $\boldsymbol{\xi}$. Further data analysis shows that excluding hovering has no significant impact on the velocity autocorrelations, which are dominated by the search flights. This is in full agreement with figures 2.17, 2.18, where the time scale for the predator-induced anti-correlation (Fig. 2.18) is larger than the time scale for flights between adjacent flowers (Fig. 2.17). Hence, we model $\boldsymbol{\zeta}_1(t)$ as a vector of Gaussian white noise with the smaller variance σ_1^2 given in Table 2.1 which describes the hovering. The search flights from flower to flower are reproduced by the correlated Gaussian noise vector $\boldsymbol{\zeta_2}(t)$ with variance σ_2^2 and the autocorrelations $v_i^{ac}(\tau)\,,\ i=x,y$ shown in Figs. 2.17, 2.18. While this model is a quite simplistic phenomenological model, the components we arrive at are directly based on our data analysis.

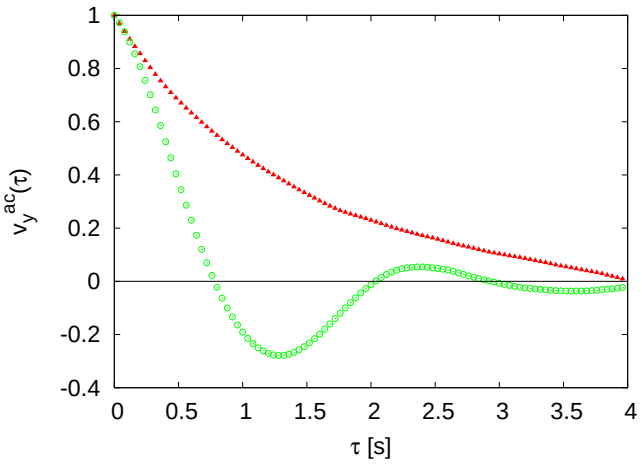

Figure 2.19: v_y-Autocorrelation for a model with a repulsive potential. Eqs. (2.10),(2.11) model the predation threat by different strengths of the repulsion. Shown are results from computer simulations without ($u = 0$; red triangles, upper line) and with predation threat ($u = 0.5\,\mathrm{m}^2/\mathrm{s}^2$; green circles, lower line). These results should be qualitatively compared with the experimental findings Fig. 2.18.

Simple Model explaining Anti-Correlations

We now focus on the different aspect of understanding the biophysical mechanism that generates the anti-correlations of the velocities parallel to y shown in Fig. 2.18. Starting from the full model Eqs. (2.8), since we have seen above that the velocity autocorrelations are dominated by the search flights it suffices to select that mode by setting $\boldsymbol{\xi}(\mathbf{r}, t) = \boldsymbol{\xi}_2(t)$ thus neglecting any spatial variations of the noise. This yields the Langevin equation

$$\frac{dv_y}{dt}(t) = -\eta v_y(t) - \frac{\partial U}{\partial y}(y(t)) + \xi(t)\,, \tag{2.10}$$

for the y-velocity only. A rough approximation for the repulsive force is provided by a periodic potential with maxima at dangerous flowers,

$$U(\mathbf{r}) = u\cos\left(2\pi\frac{y}{y_0}\right)\,, \tag{2.11}$$

where y_0 is the mean distance between spiders and u the strength of the repulsion.

We integrated this Langevin equation via an Euler-Maruyama method (see section A.3) under variation of u by computing the autocorrelation function v_y^{ac} of the generated data. Figure 2.19 shows v_y^{ac} by increasing the repulsion strength u. The correlation function changes from positive correlations to anti-correlations in a range of delay times τ comparable to the changes in the correlation function of the experimental data of Fig. 2.18. This qualitatively reproduces our experimental findings from first principles. Note that the oscillations for higher τ in Fig. 2.19 would be suppressed in a higher-dimensional model. The other directions can be treated analogously, e.g., by including an x-dependent term in the potential for the attraction of the bumblebees to the flower wall. A stochastic analysis of Langevin equations with periodic potentials can be found, e.g., in Ref. [38]. The effect of the harmonic potential on the creation of negative velocity correlations can also be calculated analytically [38].

We emphasize that our model Eqs. (2.10), (2.11) provides only a qualitative description of the biophysical mechanism generating the change in the correlations of the bumblebee velocities under predation threat. For a quantitative comparison to the experimental data a much more detailed model would be necessary, which needs to include the random positioning of the spiders and the general attractive force exerted by the flowers onto the bumblebees. Modelling the three-dimensional nature of the potential would also be important: notice, e.g., the local maximum of v_y^{ac} around $\tau \simeq 2.5$ which is an artefact of the one-dimensional modelling of spider avoidance. However, as it is difficult to reliably estimate the parameters of the potential from the given experimental data, we do not attempt such a quantitative comparison here. Instead of focussing on the local interactions with the food sources and predators, we will develop a more general model for the bumblebee behaviour during search flights from flower to flower in chapter 3.

2.3 Connection to the Lévy Hypothesis

The motivation for the introduction of Lévy walks into foraging from a theorists' point of view were the existence of a generalized central limit theorem as discussed in section 1.3.2 and optimality claims [77, 5], due to being scale-free (see

section 1.4.2). We emphasize that the experiment analysed in this chapter does not match the conditions of the Lévy flight hypothesis [5] (see section 1.4.2). We therefore cannot directly support or reject the biological Lévy hypothesis (section 1.4.2) with the given data, but can only argue indirectly over its applicability. Lévy flights and Lévy walks predict scale-free probability distributions [74] and generate trivial functional forms for the velocity correlations [89, 90]. Accordingly, experiments testing this hypothesis have focused on probability distributions, not on correlation decay [6, 84, 10, 68]. However, our results (see section 2.2.4) demonstrate that velocity autocorrelations can contain crucial information for understanding foraging dynamics, here in the form of a highly non-trivial correlation decay emerging from an interaction between forager and predator. Identifying such an emergent property in contrast to adaptive behaviour (see section 1.4.2), as we do with our simple model, has been highlighted as a crucial problem in foraging dynamics [68]. In addition, we observe a spatial variation of the velocity distributions (see section 2.2.2). These findings illustrate the presence of different flight modes governing the foraging dynamics on different scales of time and space. Our results thus indicate that taking scale-free distributions as a paradigm beyond the conditions of validity of the mathematical Lévy flight hypothesis might be too restrictive an approach in order to capture complex foraging dynamics of animals. This is consistent with our discussion in the previous chapter: in real application in the context of foraging animals, a variety of mechanisms may naturally lead to much more complicated distributions, e.g., individuality of animals [25, 24, 26], an intermittent switching between quasi-ballistic persistent dynamics and localised search modes [88, 48], or quantities over which one has averaged like time of day [68]. It is easy to mistake non-trivial velocity or step length distributions as a sign of the presence of Lévy walks, while a more detailed data analysis can reveal that seemingly heavy-tailed distributions are only an effect of, e.g., failing to distinguish different movement modes or ignoring a food source inhomogeneity (compare sections 1.3.5, 1.2.2, and 1.4.2). The variety of complications which can arise highlight the need to better understand, and more carefully analyse, the interplay between forager and environment, which will yield crucial information for constructing better mathematical foraging models. From that point of view Lévy walks are probably not a good starting point for

the analysis of foraging data.

2.4 Summary

In this analysis of an experiment on foraging bumblebees our main concerns were how the search behaviour the bumblebees is influenced by the presence of predators, in which statistical properties the influence manifests itself, and how and if the results are related to the Lévy hypothesis.

We found that the velocity distributions of the bumblebees can be described by Gaussian mixtures which can be interpreted as a switching process between two flight modes depending on the position of the bumblebee. As expected by the boundedness of the experiment, we did not find any evidence supporting the Lévy hypothesis (see section 1.4.2): the velocity distributions and also the flight-durations show no power law behaviour. The non-trivial velocity autocorrelations also do not match with a Lévy walk model (see section 1.3.2). In particular, the observed changes in the autocorrelation due to the introduction of a predatory threat would be difficult to include in such a model. The presence of different flight modes and their impact on the velocity distributions, and the changes in the velocity autocorrelations due to environmental changes show that bumblebee foraging is governed by different dynamics on different scales of time and space. We therefore argue (in section 2.3) that scale-free models such as Lévy flights might thus be a too simplistic approach to foraging.

Regarding the memory of the bumblebees and their adaptation to predation threat we confirmed an increased hesitation behaviour in front of flowers with spiders on them (section 2.2.3, [93]). More importantly we found a tendency to avoid flying near/above both types of flowers, dangerous and safe, under predation threat. Although not as strong as in front of flowers with spider models, the effect was also present at flowers without spiders. It is therefore not a direct reaction to seeing a spider model, but instead an adaptation to the general threat of predation. We described this learned flower avoidance behaviour by a repulsive potential in a flight model in section 2.2.5, for which we then discussed a simplified one-dimensional projection resulting in a qualitative modelling of the predation-induced velocity anti-correlations which we observed in section 2.2.4.

While we here concentrated on the effects of a predatory threat and the interaction of the bumblebees with the flowers, we will take a closer look at the search behaviour of the bumblebees outside the flower zones in the following chapter.

Chapter 3

Modelling Bumblebee Flights

In chapter 2 we analysed experimental data of foraging bumblebees with a focus on understanding the adaption of the bumblebees to their environment and the effect of artificial predators on the flight behaviour. In this chapter we want to step away from a detailed description of the local interaction with flowers and predators. Instead our aim is to find a good stochastic model for the bumblebee flights away from the flower wall. As an important aspect of the observed bumblebee movement is the directional persistence, we use a generalization of the reorientation model (see section 1.3.3) similar to the generalized Langevin equation in section 1.3.4 as our model class. The goal is a biologically and physically plausible model whose statistical properties should be similar to those of the experimental data.

In section 3.1 we describe the general set-up of our model, which we then construct in section 3.2 from the experimental data. With the estimated model parameters and interdependencies, we then validate the model in section 3.3 by simulation of its stochastic differential equations (section 3.3.2) and a comparison of the resulting simulated data with the experimental data in section 3.3.2. We conclude the chapter with a discussion of the differences of our model and reorientation models in section 3.2.8 and a summary in section 3.4. A brief presentation of the main results of this chapter can be found in [3].

3.1 Set-up and Assumptions

In the following we will present a flight model in terms of generalized Langevin Equations (see section 1.3.4), which we then use to analyse experimental bumblebee data. In this chapter we will focus on the horizontal bumblebee movement. By neglecting the slower vertical movements, which are of more interest when analysing the starting and landing behaviour near flowers (see section 2.2.3), we thus restrict ourselves to a two-dimensional model. Since we are not interested in the interaction with the food sources, we exclude flights near the flowers, i.e. in the flower zones (see section A.2.2) from the data of the experiment [1, 92] described in section 2.1.

Given movement data of flying bumblebees available with a constant time step Δt, the step length is determined by the speed $s(t) = |v(t)|$ of the animal. As we will be looking at a flying insect in a data recording which uses a small time step, we may expect to have a deterministic persistence due to the animals momentum. A reorientation model would assume that s and β are drawn i.i.d., which is sensible if Δt is large enough. However, for small time steps it cannot be excluded that the decision of the animal to turn left or right takes longer than the time step (or persists over a longer time time), which can correlate the turning-angles $\beta(t)$ over a number of time steps. If one wants to arrive at a better stochastic model for bumblebee flights than the simple reorientation model in section 1.3.3, one therefore has to capture the dynamics of the turning-angle and the speed in addition to their distributions. We model the changes in speed and turning-angle via two coupled generalized Langevin equations (under It$\bar{\text{o}}$-interpretation),

$$\frac{d\beta}{dt}(t) = h(\beta(t), s(t)) + \tilde{\xi}_s(t) \tag{3.1}$$

$$\frac{ds}{dt}(t) = g(\beta(t), s(t)) + \psi(t), \tag{3.2}$$

where we distinguish between the deterministic parts h and g and stochastic terms ψ and $\tilde{\xi}_s$ (whose speed dependence will be discussed in section 3.2.5). We assume that the noise processes are stationary with autocorrelation functions which may be non-trivial, and we make no further assumptions for the shape of their

stationary distributions.

While Eqs. (3.1, 3.2) represent a time-continuous description, the turning angle β yields the change of the direction α (in a non-comoving frame) according to the fixed time resolution Δt. That is, $\beta(t)$ relates to a time-continuous angular velocity γ of α via $\beta(t) = \int_{t-\Delta t}^{t} \gamma(\tau) d\tau$. The animals' position $\mathbf{r}(t) = (x(t), y(t))$ is then given by $dx/dt = s\cos(\alpha(t))$, $dy/dt = s\sin(\alpha(t))$ and $d\alpha/dt = \gamma(t)$. The numerical analysis is done with time-discrete data where the measured turning angle is given by $\beta(t) = \measuredangle(\mathbf{v}(t), \mathbf{v}(t-\Delta t))$, where $\mathbf{v}(t) = (\mathbf{r}(t+\Delta t) - \mathbf{r}(t))/\Delta t$ at times $t = n\Delta t$, $n \in \mathbb{N}$.

3.2 Model Construction

The generalized Langevin equations (3.1, 3.2) are an approach to model the bumblebee movement, which tries to separate deterministic parts of the dynamics from stochastic ones. The stochastic terms are not assumed to be originating only from outside influences, e.g. turbulences. Instead they will also represent the non-deterministic decision processes of the animal.

In this section we will first look at how one can, assuming stationarity and Markovianity (see section 3.2.1), extract coefficient functions of a Langevin Equation from data via a description by a Fokker-Planck Equation in section 3.2.3. In section 3.2.4 we will then extract and discuss the deterministic terms of our bumblebee model from experimental data. After a discussion of the interdependencies of turning-angles β and the speed s in section 3.2.5, we will then determine a stochastic description of β and s.

3.2.1 Stationary and Markov Processes

Two properties we will have to assume of our data, if we want to estimating Fokker-Planck coefficients, are stationarity and Markovianity.

Let $\mathbf{X}(t)$, $t \in \mathbb{R}^+$ be a d-dimensional stochastic process and $w(\mathbf{X}, t)$ be the time-dependent probability densities in phase space given a fixed initial distribution $w(\mathbf{X}, 0)$.

The *joint probability density* $w(\mathbf{X}_n, t_n; \mathbf{X}_{n-1}, t_{n-1}; \ldots; \mathbf{X}_0, t_0)$ is the probability density to be at $\mathbf{X}_i$ at time t_i for all $i = 0, 1, \ldots, n$. All the joint probability densities together specify a stochastic process [38].

The process $\mathbf{X}(t)$ is called *stationary* if all joint probability densities (for all $n \geq 0$) are independent of any time shift T:

$$w(\mathbf{X}_n, t_n; \mathbf{X}_{n-1}, t_{n-1}; \ldots; \mathbf{X}_0, t_0) = w(\mathbf{X}_n, t_n{+}T; \mathbf{X}_{n-1}, t_{n-1}{+}T; \ldots; \mathbf{X}_0, t_0{+}T). \tag{3.3}$$

The *conditional probability density* $p(\mathbf{X}_n, t_n|\mathbf{X}_{n-1}, t_{n-1}; \ldots; \mathbf{X}_0, t_0)$[1] is the probability density to be in state $\mathbf{X}_n$ at time t_n if the system was in $\mathbf{X}_i$ at t_i for all i,

$$p(\mathbf{X}_n, t_n|\mathbf{X}_{n-1}, t_{n-1}; \ldots; \mathbf{X}_0, t_0) := \frac{w(\mathbf{X}_n, t_n; \mathbf{X}_{n-1}, t_{n-1}; \ldots; \mathbf{X}_0, t_0)}{w(\mathbf{X}_{n-1}, t_{n-1}; \ldots; \mathbf{X}_0, t_0)}. \tag{3.4}$$

We call the process $\mathbf{X}(t)$ a *Markov process* if the conditional probability density of the process has the *Markov property*:

$$p(\mathbf{X}_n, t_n|\mathbf{X}_{n-1}, t_{n-1}; \mathbf{X}_{n-2}, t_{n-2}; \ldots) = p(\mathbf{X}_n, t_n|\mathbf{X}_{n-1}, t_{n-1}). \tag{3.5}$$

This means that the time evolution of the probability density function depends only on one previous time step. Eq. (3.5) holds for arbitrary $t_n - t_{n-1}$.

3.2.2 The Fokker-Planck Equation

While Langevin equations are stochastic descriptions of Markov processes, the *Fokker-Planck equation* is a deterministic way to describe these systems.[2] Instead of analyzing the dynamics of the observables directly, we now change our point of view on stochastic processes by looking at the dynamics of the probability density of the observables in phase space instead. The Fokker-Planck equation is an advection diffusion equation for the probability density function $w(\mathbf{X}, t)$ of

[1] where $t_{m+k} > t_m$ if $k > 0$

[2] Of course the Fokker-Planck equation is only deterministic in providing the deterministic dynamics of the probability density function and not of realizations of the process.

$\mathbf{X}(t)$:

$$\begin{aligned}\frac{\partial}{\partial t}w(\mathbf{X},t) = & - \sum_{i=1}^{d}\frac{\partial}{\partial X_i}\left(D_i^{(1)}(\mathbf{X},t)w(\mathbf{X},t)\right) \\ & + \frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial X_i \partial X_j}\left(D_{i,j}^{(2)}(\mathbf{X},t)w(\mathbf{X},t)\right) \end{aligned} \tag{3.6}$$

where w is the probability density function, $\mathbf{D^{(1)}}$ is the drift vector, $\mathbf{D^{(2)}}$ is the diffusion tensor, $\mathbf{X}$ is the vector in phase space and d is the dimension of the system.

The diagonal elements in the diffusion tensor describe the strength of the normal diffusion in the different directions in phase space while the off-diagonal elements measure cross diffusion.

We can rewrite the right hand side of Eq. (3.6) by introducing the *Fokker-Planck operator* $\mathcal{L}_{FP}$ so that the Fokker-Planck equation reads:[3]

$$\frac{\partial}{\partial t}w(\mathbf{X},t) = \mathcal{L}_{FP}(\mathbf{X},t)w(\mathbf{X},t). \tag{3.7}$$

We can integrate the Fokker-Planck equation to compute the evolution of a given probability density as an initial condition. A special case of an initial condition for a Fokker-Planck equation are δ-peaks as discussed in section 3.2.3 and leads to a method to extract the Fokker-Planck coefficients from sample paths.

3.2.3 Estimating the Drift- and Diffusion Coefficients

Let us assume that we know the exact state $\mathbf{X^s}$ of a Markov process at time t^s. From the point of view of a Fokker-Planck equation this means that the probability density function at time t^s is a δ-peak:

$$w(\mathbf{X},t^s) = \delta(\mathbf{X}-\mathbf{X}^s). \tag{3.8}$$

[3]The same can be done in the case of the Kramers-Moyal expansion giving the Kramers-Moyal operator $\mathcal{L}_{KM}$.

The conditional probability density at a time $t > t^s$ is then just the probability density function at that time:

$$p(\mathbf{X}, t|\mathbf{X}^s, t^s) = w(\mathbf{X}, t). \tag{3.9}$$

As the Fokker-Planck equation is an advection diffusion equation we know that for short times $\tau = t - t^s$ the probability density function $w(\mathbf{X}, t)$ is a multivariate normal distribution with a mean of $\mathbf{X}^s + \tau\mathbf{D}^{(\mathbf{1})}(\mathbf{X}^s, t^s)$ and a variance (covariance matrix) of $\tau\mathbf{D}^{(\mathbf{2})}(\mathbf{X}^s, t^s)$ in the first order of τ [38]. This means that if we have a large enough ensemble of sample paths of lenght τ starting from position $\mathbf{X}^s$ at time t^s, we can estimate the mean and the variance of $w(\mathbf{X}, t^s + \tau)$ to calculate the Fokker-Planck coefficients at position $\mathbf{X}^s$ at time t^s.

If the process is stationary we can estimate the time-independent drift- and diffusion coefficients from only one sample path $\widetilde{\mathbf{X}}(t)$ by [38]:

$$\mathbf{D}^{(\mathbf{1})}(\mathbf{X}) = \lim_{\tau \to 0} \frac{1}{\tau} \left\langle \widetilde{\mathbf{X}}(t+\tau) - \mathbf{X} \right\rangle \Big|_{\widetilde{\mathbf{X}}(t)=\mathbf{X}} \tag{3.10}$$

$$\mathbf{D}^{(\mathbf{2})}(\mathbf{X}) = \lim_{\tau \to 0} \frac{1}{\tau} \left\langle \left(\widetilde{\mathbf{X}}(t+\tau) - \mathbf{X} \right) \cdot \left(\widetilde{\mathbf{X}}(t+\tau) - \mathbf{X} \right)^{\top} \right\rangle \Big|_{\widetilde{\mathbf{X}}(t)=\mathbf{X}} \tag{3.11}$$

where $\widetilde{\mathbf{X}}(t)$ is a realization of the Markov process and $\langle \cdots \rangle|_{\widetilde{\mathbf{X}}(t)=\mathbf{X}}$ is the conditional time average[4] over all t for which $\widetilde{\mathbf{X}}(t) = \mathbf{X}$.

For a numerical estimate from a limited amount of data, the conditional average has to be taken over all t for which $\widetilde{\mathbf{X}}(t) \in U(\mathbf{X})$ where $U(\mathbf{X})$ is a neighbourhood of $\mathbf{X}$ because we should have enough data to get a reliable average.

The limit for τ in Eq. (3.10) and Eq. (3.11) means that for data with a finite sampling rate, an approximation is needed in order to get estimates for the Fokker-Planck coefficients. An approach which is applicable to real data is given below.

[4]Ergodicity, which guarantees that the time average is the same as the space average is used here, too. For stochastic processes this follows from the stationarity with the exception of a pathological process which has a phase space which is split by infinite potential walls. In that case there would not even be a unique invariant density.

Connection of the Fokker-Planck Equation and the Langevin Equation

The relation of the coefficients of a Langevin equation (see Eq. (1.9) in section 1.3.4) to those of a Fokker-Planck equation depends on the choice of a stochastic integral, because, if we fix the process by the Fokker-Planck equation, the deterministic terms of the Langevin equation depend on its interpretation (see section 1.3.4).

For the *Itō interpretation* of a Langevin equation the drift coefficient equals the deterministic part of the Langevin equation:

$$\mathbf{D^{(1)}}(\mathbf{X}, t) = \mathbf{f}(\mathbf{X}, t). \tag{3.12}$$

The *Stratonovich interpretation* of the Langevin equation gives a more complicated result[5]:

$$\mathrm{D}_i^{(1)}(\mathbf{X}, t) = f_i(\mathbf{X}, t) + \frac{1}{2}\sum_{j,l} k_{j,l}(\mathbf{X}, t)\frac{\partial k_{i,l}}{\partial X_j}(\mathbf{X}, t). \tag{3.13}$$

The additional term is called *spurious drift* and is a consequence of the integration scheme. It is induced by and depends only on the *non-deterministic* part of the Langevin equation.

The relation between the diffusion and the stochastic part of the Langevin equation is the same for both interpretations of the stochastic integral:

$$\mathbf{D^{(2)}}(\mathbf{X}, t) = \mathbf{k}^2(\mathbf{X}, t). \tag{3.14}$$

With Eq. (3.14) we can easily get $\mathbf{D^{(2)}}$ from $\mathbf{k}$ by a matrix multiplication. The inverse transformation is not as direct: in order to compute $\mathbf{k}$ from $\mathbf{D^{(2)}}$ we have to compute a root of the diffusion matrix. A root $\mathbf{R}$ of a diagonalizable matrix $\mathbf{M}$ is a matrix satisfying $\mathbf{R}^2 = \mathbf{M}$. With a diagonalization $\mathbf{M} = \mathbf{PAP^{-1}}$, where $\mathbf{P}$ is an invertible matrix and $\mathbf{A}$ is a diagonal matrix, a root $\mathbf{R}$ can be computed by[6]:

$$\mathbf{R} = \mathbf{P}\left(\sqrt{\mathbf{A}}\right)\mathbf{P}^{-1}. \tag{3.15}$$

[5] See Eq. (1.12).

[6] The root of a diagonal matrix is just the matrix of the roots of all entries.

The root of a matrix is not unique as we have the choice of $\mathbf{P}$: for every orthonormal matrix $\mathbf{O}$ the matrix $\widetilde{\mathbf{P}} := \mathbf{OP}$ gives a diagonalization $\mathbf{M} = \widetilde{\mathbf{P}}\widetilde{\mathbf{A}}\widetilde{\mathbf{P}}^{-1}$, too. It follows that for every Fokker-Planck equation we have a corresponding family of Langevin equations for different $\mathbf{O}$ with $\mathbf{k}(\mathbf{X}, t) = \widetilde{\mathbf{k}}(\mathbf{X}, t)\mathbf{O}$. This normally does not lead to complications as these Langevin equations specify the same Markov process which can be explained by the fact that the rotated (or mirrored) d-dimensional Gaussian white noise process $\widetilde{\mathbf{\Gamma}} := \mathbf{O\Gamma}$ is the same as the process $\mathbf{\Gamma}$.

The Fokker-Planck equation and the Langevin equation are two equivalent descriptions for the same class of processes. Equations (3.12), (3.13) and (3.14) give the means to transform one description to the other and back providing, e.g., the possibility to look at probability density functions and to compute invariant probability densities from the Fokker-Planck equations and integrate the Langevin equations to get sample paths.

Finite Time Corrections for Diffusion Coefficients

In general time series which originate from measured data have a finite sampling rate. This means we cannot go to the limit of $\tau \to 0$ but we have to use the smallest τ available.[7] Due to this approximation we have to correct the diffusion term for the finite time effects induced by the drift term giving [98, 99, 100]:

$$\begin{aligned}\mathbf{D^{(2)}}(\mathbf{X}) = \tau^{-1}\Big\langle \quad & \Big(\widetilde{\mathbf{X}}(t+\tau) - \mathbf{X} - \tau\mathbf{D^{(1)}}(\mathbf{X})\Big) \\ & \cdot \Big(\widetilde{\mathbf{X}}(t+\tau) - \mathbf{X} - \tau\mathbf{D^{(1)}}(\mathbf{X})\Big)^{\top}\Big\rangle\Big|_{\widetilde{\mathbf{X}}(t)=\mathbf{X}} .\end{aligned} \tag{3.16}$$

The drift term is approximated by:

$$\mathbf{D^{(1)}}(\mathbf{X}) = \frac{1}{\tau}\left\langle \widetilde{\mathbf{X}}(t+\tau) - \mathbf{X}\right\rangle\Big|_{\widetilde{\mathbf{X}}(t)=\mathbf{X}} . \tag{3.17}$$

This correction of the diffusion coefficients due to the finite time τ is only the

[7] Another way would be to compute the Fokker-Planck coefficients for different τ and then extrapolate them to $\tau = 0$.

first term for a full correction.[8] A more elaborate correction and a discussion of different corrections can be found in [100]. In the case of $\tau \to 0$ this is not necessary as the mean is $\tau \mathbf{D^{(1)}}(\mathbf{X}) = 0$ and all corrections of higher order vanish, too. Another aspect of a finite amount of data is that any transient dynamics at the beginning of the time series should be discarded. Otherwise stationarity would be broken.

3.2.4 Determining Deterministic Dynamics of Flight Data

We now examine the bumblebee flight data by treating it as data generated by a stationary Markov process with states $\mathbf{X} = (\beta, s)^\top \in [-\pi, \pi] \times \mathbb{R}^+$, and numerically estimate [98, 100, 101, 102] the components of the drift vector field (drift coefficients) $\mathbf{D^{(1)}}(\beta, s)$ of the corresponding Fokker-Planck equation using Eq. (3.17).

Since we interpret our model of Langevin equations (3.1, 3.2) under Itō-interpretation, the drift coefficients of the Fokker-Planck equation are the deterministic terms of the Langevin equations we were looking for (see section 3.2.3):

$$\mathbf{D^{(1)}}(\beta, s) = (g(\beta, s), h(\beta, s))^\top. \tag{3.18}$$

This estimation of the drift terms is based on a Markov approximation: only those parts of the dynamics which match to a Markovian description in the state space variables β and s have their deterministic terms reflected in $\mathbf{D^{(1)}}(X)$. Any other parts of the flight dynamics – stochastic as well as deterministic but not Markovian in β and s – are captured by the stochastic terms of Eqs. (3.1, 3.2).

By looking at the drift coefficients we can examine the mean behaviour of the turning-angle and the speed s. The drift vector field (normalised for better visibility) in Figure 3.1 shows that the drift is quite well-behaved: the drift vectors quickly push the turning-angle β towards 0, while the dynamics in the speed s is much slower. We therefore find a timescale separation: the deterministic part of the dynamics can be reduced to a regular and fast relaxation of the turning-angle β and slow dynamics in s. The nearly horizontal vectors with minimal curl

[8] The diffusion estimation with this correction is a kind of 'inverse algorithm' of the Euler-Maruyama approximation (see section A.3).

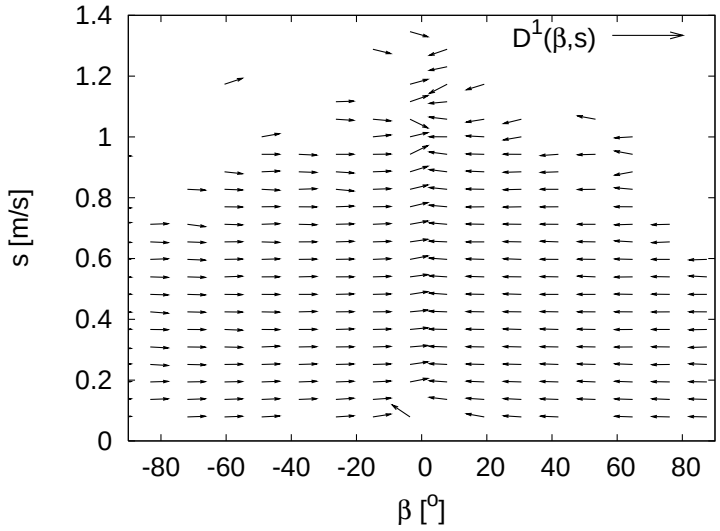

Figure 3.1: **Normalised drift vector field $\mathbf{D^{(1)}}(\beta, s)$**corresponding to the deterministic terms of the Langevin equations (3.1, 3.2) estimated via Eq. (3.10). The regular structure shows the quick relaxation to small angles and the absence of strong cross-dependencies in the drift.

demonstrate that the cross-dependencies $h(s)$ and $g(\beta)$ are weak; in our model we will neglect them completely.

By projection on the turning-angle β resp. on the velocity v we examined the drift of each variable separately: $g(s)$ and $h(\beta)$.

Examining the drift $h(\beta)$ of the turning angle in Fig. 3.2 reveals that the drift term seems linear in β — indeed we find numerically that its slope $-k$ matches exactly to a decay of the turning angle to 0 in a single observation time step Δt by $k \approx 1/\Delta t$, disregarding the noise term. This means that by integrating Eq. (3.1) over a time Δt and approximating the drift $h(\beta)$ for small Δt by $\int_t^{t+\Delta t} h(\beta(\tau))d\tau \approx h(\beta(t))\Delta t$, we have

$$\beta(t+\Delta t) - \beta(t) = -k\beta(t)\Delta t + \int_t^{t+\Delta t} \tilde{\xi}_s(\tau)d\tau = -\beta(t) + \int_t^{t+\Delta t} \tilde{\xi}_s(\tau)d\tau. \tag{3.19}$$

With $\xi_s(t) := \int_{t-\Delta t}^{t} \tilde{\xi}_s(\tau)d\tau$ and Eq. (3.19), the time scale separation in the β-Langevin equation due to the very fast relaxation means that we can simplify

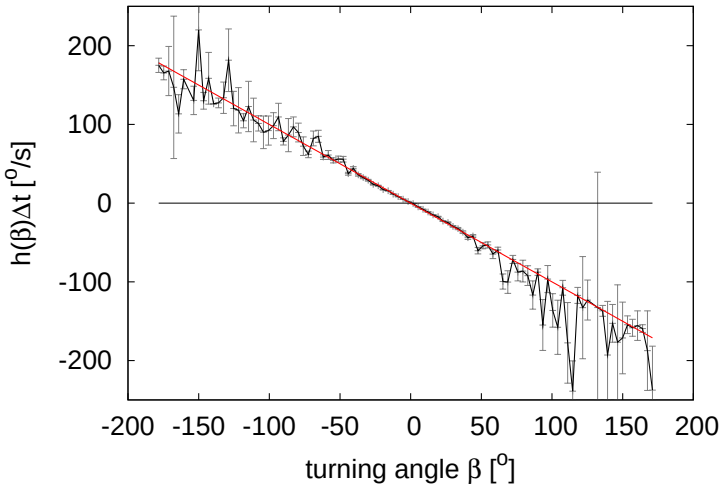

Figure 3.2: **Drift coefficient of the turning-angle.** The deterministic drift $h(\beta)$ as estimated from data (black, 95% confidence intervals in grey) is in good approximation linear (Stokes-like) in β (red diagonal).

Eqs. (3.1, 3.2) to:

$$\beta(t) = \xi_s(t), \tag{3.20}$$

$$\frac{ds}{dt}(t) = g(s(t)) + \psi(t). \tag{3.21}$$

While this reduction of dynamics from $d\beta/dt$ to β makes the model resemble the simple reorientation model (section 1.3.3), the turning angles are still correlated, as we will see in section 3.2.6. Since the turning angles are smaller for high velocities it would be tempting to use βs as a scaled turning angle to simplify the geometry of the system.[9] However, as we will see in section 3.2.5, the speed-dependence of the turning angle is more complex.

The speed-drift $g(s)$ displayed in Fig. 3.3 shows that the deterministic part of the speed-Langevin equation alone is non-linear and would have a stable fixed point around $s_0 = 0.27\,\mathrm{m/s}$. Comparing the slopes above and below s_0 reveals that for $s < s_0$ the force towards s_0 is stronger than for $s > s_0$. This is biologically

[9]This small-angle approximation would assume a purely geometric dependence of s on β.

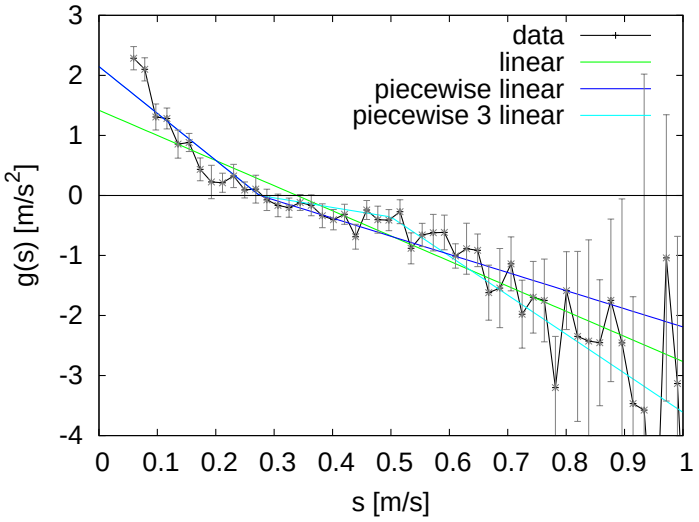

Figure 3.3: **Drift coefficient of the speed.** The experimental deterministic drift coefficient $g(s)$ (black, 95% confidence intervals in grey) has been approximated by piecewise linear functions from one to three pieces (blue,green,cyan). The data shows the tendency to quickly increase low speeds. However, speeds above 0.27 m/s decrease more slowly, except for the rare high speeds.

plausible if one interprets s_0 as a preferred speed: if the bumblebee is slower it accelerates, but if it is faster it does not rush to decelerate as it would give up the energy spent to reach a high velocity. For very high velocities (over 0.55 m/s) the slope of $g(s)$ increases again. This might be caused by the limited space available to the bumblebee in the flight arena. For our model we approximated $g(s)$ by a piecewise linear function:

$$g(s) \approx (s - s_0) \times \begin{cases} -d_1 & \text{for} \quad s < s_0 \\ -d_2 & \text{for} \quad s \geq s_0 \end{cases}, \tag{3.22}$$

where $d_1 > d_2 > 0$. As the very high velocities are rare, it made no difference in our model whether we used Eq. (3.22) or a piecewise linear function with three pieces.

Beyond Deterministic Bumblebee Dynamics

With the deterministic drift terms (see section 3.2.4) estimated from the experimental data as described in section 3.2.3, we could now go on with an estimation of the diffusion tensor as described in section 3.2.3. However, we know that the assumption of the Markov property (see section 3.2.1) is actually not valid for the analysed data. For example, anti-correlations, as observed in the velocities v_x and v_y in x- and y-direction in section 2.2.4, cannot be generated by a Fokker-Planck equation with the drift vector field shown in Fig. 3.1 together with uncorrelated diffusion terms. Our approach is therefore the following: we estimated the drift terms assuming that the process is Markovian as described above. Therefore the drift terms only capture the mean behaviour, and all parts of the dynamics which are not described by the drift have now to be treated as noise. Notice that the full flight dynamics has been projected on the turning-angle β and the speed s – should there be other relevant variables, with our modelling approach their dynamics will contribute to the noise terms even if their dynamics was actually deterministic. This means that, in order to get a useful description of the data, we have to allow autocorrelations in the noise terms of the Langevin equation. Apart from this coloured noise, another reason not to use the estimation of the Fokker-Planck diffusion tensor as described above, is the complication of a dependence between turning-angle and speed. In the following we will at first quantify this dependence in section 3.2.5 and then discuss the coloured noise terms $\xi_s(t)$ and $\psi(t)$ of β and s separately in section 3.2.6 and 3.2.7 respectively.

3.2.5 Dependencies of Turning-Angle and Speed

The turning-angle of an animal and its speed are often assumed to be independent for simplicity. Given that the force a bumblebee can use to change directions is finite, the largest turning-angles have to be smaller when flying with high speeds (see Fig. 3.4). In our case, this is consistent with the absence of simultaneously having high speed and large turning-angle in the data, as is evident, e.g., from the data gaps in Fig. 3.1 in section 3.2.3. However, animals can counteract this geometric dependence by varying the forces used for changing direction with the speed.

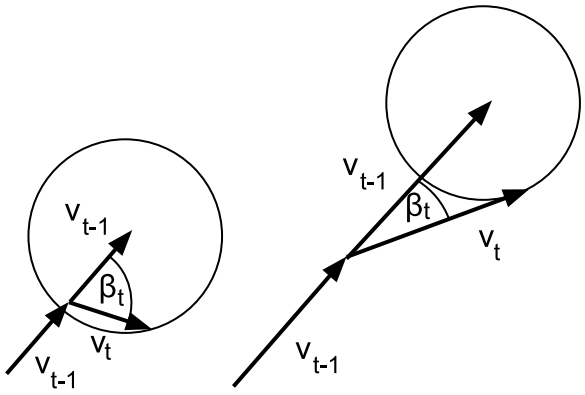

Figure 3.4: **Schematics of the dependence of β on speed s.** Assuming a constant maximal force (circle) available to the bumblebee to accelerate during a time step, the distribution of the turning-angle β depends on the previous speed $s_{t-1} = |\mathbf{v}_{t-1}|$. Illustrated is the change from large angles for low speeds (left) to a stronger concentration around 0° for higher speeds (right).

In this section we will first discuss an alternative model in which the accelerations of the organism are assumed to be independent of its speed before discussing the experimental data. In models in which the momentum of the animal is not important for the observed directional persistence, this cross-dependence is often neglected [44].

Turning-Angles in a Model with Speed-independent Accelerations

A simple model showing a dependence of the turning-angles on the speed (see Fig. 3.4) is given in the following. Given the velocity $\mathbf{v}(t) = (v_1(t), v_2(t))^\top$ of an animal for each time step Δt, assume that the distribution of acceleration vectors $\mathbf{a}(t) = \frac{\mathbf{v}(t)-\mathbf{v}(t-\Delta t)}{\Delta t}$ is invariant under rotation with variance σ^2 in all directions, and the random accelerations are drawn i.i.d. from a binormal distribution and independent of the speed.

Using the comoving frame of the animal at time t, i.e. centered at the animals position $\mathbf{x}(t) = (x_1(t), x_2(t))^\top = 0$ and oriented in the (old) direction $\mathbf{v}(t-\Delta t) = (s(t-\Delta t), 0)^\top$ for step lengths $s(t) = |\mathbf{v}(t)|$, the position at time $t + \Delta t$ is distributed as

$$\tilde{\rho}(\mathbf{x}(t+\Delta t)) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x_1(t)-s(t-\Delta t))^2+x_2(t)^2}{2\sigma^2}}. \tag{3.23}$$

Changing a volume element $d\mathbf{x}$ to polar coordinates $(s(t), \beta(t))$ with the new step length $s(t)$ and turning-angle $\beta(t)$ between $\mathbf{v}(t-\Delta t)$ and $\mathbf{v}(t)$ results in the probability $\rho(s(t), \beta(t))dsd\beta := \tilde{\rho}(\mathbf{x}(t+\Delta t))d\mathbf{x}$ via

$$\rho(s(t), \beta(t))ds(t)d\beta(t) = \frac{1}{2\pi\sigma^2} e^{-\frac{s(t)^2 + s(t-\Delta t)^2 - s(t)s(t-\Delta t)\cos(\beta)}{2\sigma^2}} s(t)ds(t)d\beta(t). \tag{3.24}$$

The turning-angle β then depends on the quotient $\eta(t) := \frac{s(t-\Delta t)}{\sqrt{2}\sigma}$ between the former speed and the noise strength σ. Integrating out $s(t)$ the distribution $\rho(\beta)$ of the turning-angle is given by:

$$\rho(\beta) = \frac{e^{-\eta^2}}{2\pi} + \frac{e^{-\eta^2 \sin^2(\beta)}}{2\sqrt{\pi}} \eta\cos(\beta)(1 + \operatorname{erf}(\eta\cos(\beta))) \tag{3.25}$$

for $-\pi \leq \beta \leq \pi$. With vanishing relative speed $\eta(t) = 0$ the first term gives a uniform distribution as expected, and for $\eta(t) \to \infty$ the distribution sharply peaks at $\beta = 0$ with its variance σ_β approaching 0, similar to the behaviour in the simpler case of a von Mises distribution [42, 103].

Experimental Speed Dependence of Turning-Angles

Analysing the experimental data we find a strong dependence of the turning-angle on the speed. Figure 3.5 shows the standard $\sigma_\beta(s)$ of the turning-angle distribution as a function of the bumblebee speed s. The dependence of σ_β on s is robust over data collected from the different experimental stages: the variations seen for high speeds s in Fig. 3.5 are statistical errors due to a lack of sufficiently many data points for high speeds s.

The experimental bumblebee data does not show a decay of σ_β to 0 but to a finite positive value. Therefore the simple geometric model with constant accelerations in section 3.2.5 does not hold: the accelerations have to be modelled as speed-dependent.

While Fig. 3.5 shows that distinguishing between exact functional forms for $\sigma_\beta(s)$ is difficult, the double-logarithmic plot in Fig. 3.6 suggests that the decay of $\sigma_\beta(s)$ to a constant offset is roughly exponential. The given confidence intervals are calculated based on the χ^2-distributed variance (see section A.1.2). Given the

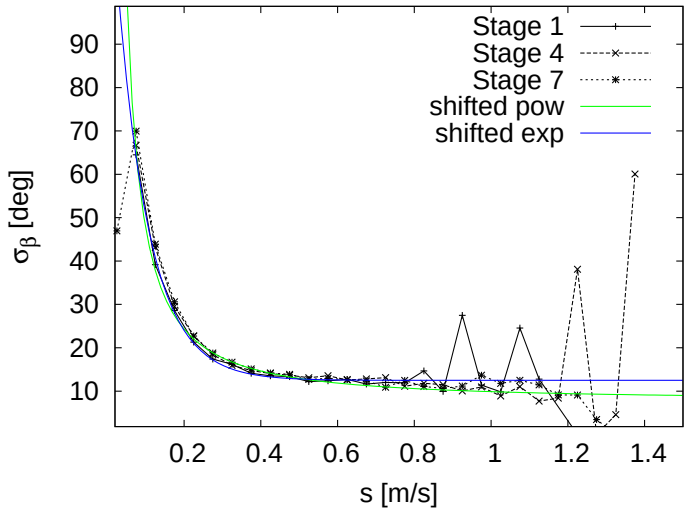

Figure 3.5: **Speed-dependence of the turning-angle.** The standard deviation σ_β of the turning-angle distribution is shown as a function of the speed as estimated from data (black dots) and approximated by shifted power-law (green) and shifted exponential (blue). The dependence is robust over the different experimental stages (solid, dashed, dotted).

amount of data, the possibility of a power-law decay with the same constant offset cannot be reliably excluded, however for large speeds s the exponential tail is a better match to the data.

3.2.6 Stochastic Description of Turning-Angles

As seen in section 3.2.5, the distribution for the turning-angles depends on the speed of the bumblebee. In theory one would have to estimate its shape for each range of speeds separately to get a good description of the turning-angle. However, this would significantly limit the number of usable data points for the estimation. For simplicity we therefore approximated the distribution of the turning-angles $\rho_s(\beta)$ for each given speed s by a normal distribution. This approximation works best for low speeds, as can be seen from the estimated kurtosis[10] of $\rho_s(\beta)$

[10] $\mathrm{Kurtosis}[X] = \frac{E[X^4]}{E[X^2]^2}$ where $E[X^4]$ is the 4^{th} central moment and $E[X^2]$ is the variance, since the mean of X is 0.

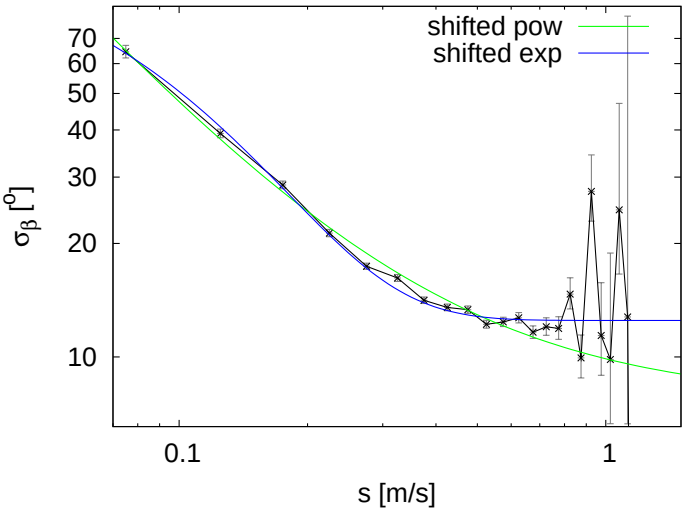

Figure 3.6: **Log-log plot demonstrating the speed-dependence of the turning-angle distribution.** The standard deviation σ_β of the turning-angle is shown as a function of the speed as estimated from data from all experimental stages (black) and approximated by a shifted power-law (green) and a shifted exponential (blue). 95% confidence intervals for σ_β based on a χ^2-distribution are shown in grey.

shown in figure 3.7. For higher speeds the kurtosis is consistently higher than the 3 expected for a normal distribution. While there are deviations from Gaussianity, we did not find a reliable fit of a better model for the whole distribution due to the limited amount of data available. For our model we made the simplifying assumption of Gaussian noise.

In total, we therefore model the turning-angles as speed-dependent Gaussian noise: $\xi_s(t) \sim \mathcal{N}(0, \sigma_\xi(s)))$ with $\sigma_\xi(s) = c_1 e^{-c_2 s} + c_3$ as estimated above in section 3.2.5. The offset c_3 could either be an effect of the boundedness of the flight arena, since the bumblebee has to turn more often to avoid walls when flying fast. Or it could be that the bumblebees use stronger forces for turning during fast flights to maintain their manoeuvrability. It would be interesting to examine free-flight data to check for the cause. For the two stochastic parts of the Langevin equations, we estimated the autocorrelation functions from the data. The turning-angle autocorrelation is approximated by a power-law as seen in Fig. 3.8, which

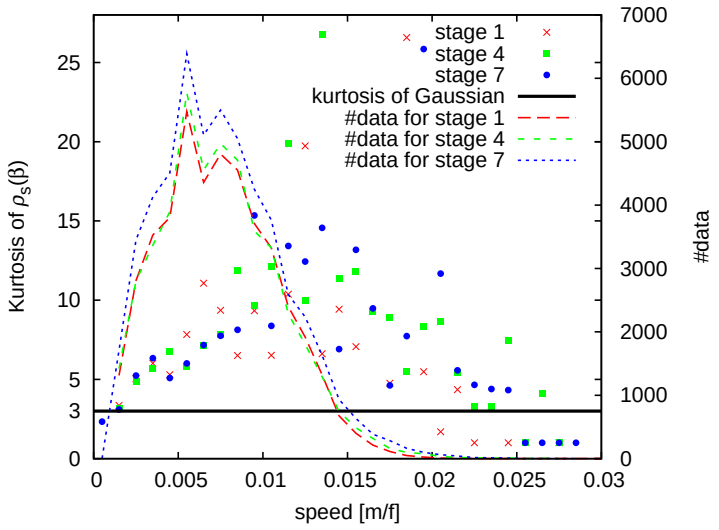

Figure 3.7: **Kurtosis of the turning-angle distribution.** The kurtosis of $\rho_s(\beta)$ is given as a function of s for all experimental stages. The kurtosis of a normal distribution and the number of available data points are shown for reference.

in this case is preferable to the alternative fit by a simple exponential decay.

3.2.7 Stochastic Description of Speed

In addition to the drift term $g(s)$, the dynamics of the speed is governed by a stochastic term, which we analyse here. By subtraction of our approximation for the deterministic term $g(s)$ from the observed speed changes ds/dt in Eq. (3.21) we can estimate the distribution and autocorrelation of the acceleration noise term $\psi(t) = ds(t)/dt - g(s(t))$.

Strength of the Acceleration Noise Term

The noise term $\psi(t)$ is well approximated by Gaussian noise, however the strength of the noise has to be corrected for discretisation effects.

In order not to overestimate the noise term, discretisation errors of an approximate size of $\Delta x/\Delta t^2$ due to the finite resolution $\Delta x = 10^{-3}\,\mathrm{m}$ of the cameras have been accounted for. The calculation of the discretisation error is a one-

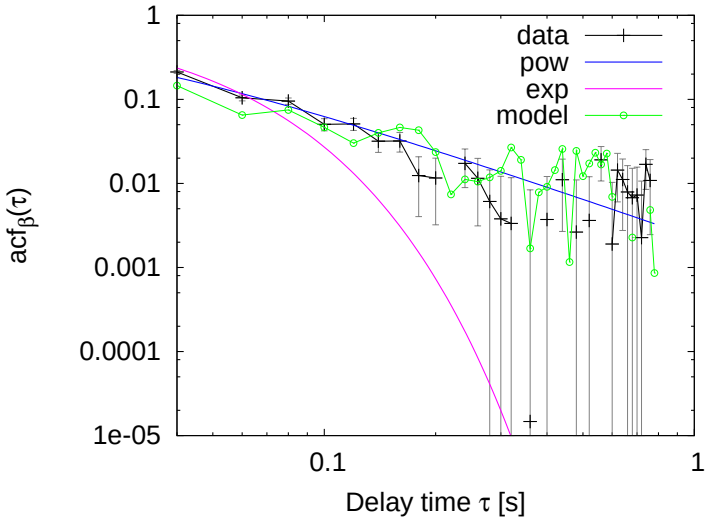

Figure 3.8: **Log-log plot of the autocorrelation of turning-angles β.** The experimental data (black crosses) together with an exponential (magenta) and a power-law (blue) fit is shown with the large-lag standard error (grey). The green circles show the autocorrelation extracted from the simulated data.

dimensional approximation of the error in accelerations given discretised position data. The real one-dimensional bumblebee positions x_t for $t \in 1 \ldots n$ can be described by $x_t = \tilde{x}_t + u_t$, where $\tilde{x}_t$ is the measured discretised position (i.e. the center of a bin) and u_t is uniformly distributed between $-\Delta x/2$ and $\Delta x/2$ and is assumed to be drawn i.i.d., representing the uncertain relative position inside a discretisation bin. With $x''_t = \frac{1}{\Delta t^2}(x_{t+\Delta t} - 2x_t + x_{t-\Delta t})$ as an approximation for the real accelerations, the measured accelerations $\tilde{x}''_t = \frac{1}{\Delta t^2}(\tilde{x}_{t+\Delta t} - 2\tilde{x}_t + \tilde{x}_{t-\Delta t})$ have a variance of

$$Var(\tilde{x}''_t) = Var(x''_t) + Var(u_{t+\Delta t} - 2u_t + u_{t-\Delta t}) = Var(x''_t) + \frac{\Delta x^2}{2\Delta t^4}. \quad (3.26)$$

Therefore the standard deviation of the real accelerations is given by:

$$Stdev(x''_t) = \sqrt{Var(\tilde{x}''_t) - \frac{\Delta x^2}{2\Delta t^4}}. \quad (3.27)$$

As the bumblebee flights are modelled in two dimensions, Δx has been scaled

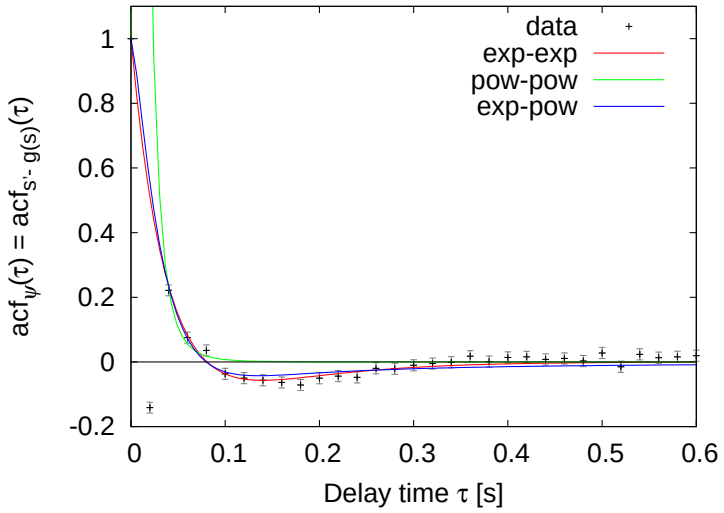

Figure 3.9: **Autocorrelation of the non-deterministic speed changes $\psi(t)$.** The autocorrelation function of $\psi(t) = ds/dt(t) - g(s(t))$ estimated from the experimental data (dots) with two times the large-lag standard error (grey) and three fitted approximations: difference of 2 exponentials (red), difference of 2 power-laws (green), difference of exponential and power-law (blue).

by $\sqrt{2}$, giving a rough estimate for the strength of the discretisation-induced noise added to $\psi(t)$. Since Δx is quite small in the experimental data, this has been good enough for our modelling purposes (see section 3.3.2). A full 2-dimensional treatment of the discretisation effects on the accelerations and especially on turning-angles would be more cumbersome.

Auto-Correlations of the Acceleration Noise Term

Figure 3.9 shows the autocorrelation function of the noise term $\psi(t)$ and a set of fitted functional shapes. The anti-correlations of $\psi(t)$ can be approximated e.g. by $\mathrm{acf}_{\psi}^{e-e}(\tau) = ae^{-\lambda_1\tau} + (1-a)e^{-\lambda_2\tau}$. While an autocorrelation function of the shape of $\mathrm{acf}_{\psi}^{p-p}(\tau) = b(\tau+1)^{-p_1} + (1-b)(\tau+1)^{-p_2}$ can be exluded, a difference between an exponential and a power-law $\mathrm{acf}_{\psi}^{e-p}(\tau) = ce^{-\lambda_3\tau} + (1-c)(\tau+1)^{-p_2}$ is not significantly worse than $\mathrm{acf}_{\psi}^{e-e}$. For our model we chose the simple difference of exponentials acf^{e-e}.

As the observed anti-correlation between delays of $0.1\,\mathrm{s} > \tau > 0.3\,\mathrm{s}$ happens on a time scale which is too short to be an effect of the boundedness of the experiment or of residual effects of the presence of the foraging wall [2], it is unclear where the anti-correlation comes from. One could speculate that it might be the result of a stabilising mechanism in the bumblebee dynamics.

3.2.8 The Complete Flight Model

The full set of parameters estimated from the data set which was used for the simulation is given below. For reference, the model equations and the equations describing all terms are also collected here. Where not specified otherwise, the parameters of the functional shapes were estimated with a least-squares fit (Marquardt-Levenberg algorithm[104, 105]). The (simplified) Langevin equations (3.20, 3.21) are,

$$\begin{aligned} \beta(t) &= \xi_s(t) \\ \frac{ds}{dt}(t) &= g(s(t)) + \psi(t). \end{aligned}$$

The parameters for the standard deviation $\sigma_\xi(s) = c_1 e^{-c_2 s} + c_3$ of the angle noise $\xi_s(t) \sim \mathcal{N}(0, \sigma_\xi(s)))$ are $c_1 = 126°$, $c_2 = 12\,\mathrm{s/m}$, $c_3 = 12.5°$ and its autocorrelation is given by $\mathrm{acf}_\beta(\tau) = (\tau + 1)^{-1.5476}$ (see sections 3.2.5, 3.2.6). For the deterministic drift of the speed (see Eq. (3.22))

$$g(s) \approx (s - s_0) \times \begin{cases} -d_1 & \text{for} \quad s < s_0 \\ -d_2 & \text{for} \quad s \geq s_0 \end{cases},$$

the change of slope is at $s_0 = 0.275\,\mathrm{m/s}$ while the slopes are $d_1 = 0.16$ and $d_2 = 0.06$. The non-deterministic changes $\psi(t)$ of the speed (see section 3.2.7) are assumed to be normally distributed with standard deviation $\sigma_\psi = 3.52\,\mathrm{m/s^2}$ and autocorrelated according to $\mathrm{acf}_\psi^{e-e}(\tau) = a e^{-\lambda_1 \tau} + (1 - a) e^{-\lambda_2 \tau}$ where $a = 1.44$, $\lambda_1 = 25.5$ and $\lambda_2 = 10.7$.

Comparison to the Reorientation Model

Given the complete flight model we can now compare it to a reorientation model (see section 1.3.3). In section 3.2.4 we found that the drift term of the turning angle is trivial, which is in accord with the reorientation model. If one is not interested in the bumblebee movement on short time scales, i.e. under $\approx 0.3\,\mathrm{s}$ where the autocorrelations of the noise terms $\xi_s(t)$ and $\psi(t)$ have not fully decayed yet, an approximation by a reorientation model would look feasible. However, the resulting simplification would still include a deterministic speed-drift $g(s)$ which induces correlations for the speed s, and the dynamics of the turning angle and the speed are still dependent via $\xi_s(t)$. Therefore the resulting model should be understood as a variation of active Brownian particle models instead (see section 1.3.4).

3.3 Model Validation

With the information gathered in section 3.2 from the experimental data, we have now completed our two-dimensional model of bumblebee flights. We can now simulate it to generate artificial sample trajectories. In this section we will describe the details of the simulation and compare the resulting flight paths to the experimental data in order to validate our findings.

3.3.1 Generating Correlated Noise

For the simulation of the bumblebee model above, we need to be able to generate noise, whose distribution and autocorrelation function match those estimated for the noise terms $\xi_s(t)$ and $\psi(t)$. There are two main ways to generate coloured noise: the first is to find some stochastic process, whose autocorrelation function $\mathrm{acf}(\tau)$ and probability density function $\rho(x)$ coincides with those of the desired noise, and numerically integrate that process. A variety of different algorithms have been used for the simulation of correlated noise [106, 107]. The success of this approach depends critically on the ability to find a suitable stochastic process with the desired properties.

The second main source of algorithms to generate coloured noise comes from the idea to use uncorrelated noise samples, and correlate them by shaping their

spectra in the frequency domain. A variety of more sophisticated algorithms have been developed to account for the distortions of the autocorrelations due to sampling and windowing effects, see e.g. [106]. In this work we used a straightforward method of correlating noise as described below. While the method is not the most accurate, it is sufficient in our case: The autocorrelation functions, which we estimated from the experimental bumblebee data, have associated measurement errors (see Fig. 3.8 and 3.9) which are much larger than the small imprecisions due to the inaccuracy of the algorithm below.

The algorithm works in the following way: We start by generating an i.i.d. noise sample $x_1, \ldots, x_n$ of the desired probability density function $\rho(x)$. In case of the turning-angle and speed noises used for the simulation of bumblebee flights, $\rho(x)$ is chosen to be a Gaussian — the sample can be generated e.g. by the *Box-Muller method* [104]. The sample is then transformed to the frequency domain with a *discrete Fourier transform* into the sequence:

$$X_k = \sum_{j=1}^{N} x_j e^{-2\pi i \frac{k}{n} j} \ . \tag{3.28}$$

The modulus of X_k is the spectral amplitude which we want to shape — in case of uncorrelated (white) noise it is the constant 1 — and the modulus squared is called the *power spectral density*. The *Wiener-Khinchin theorem* [38] states that for (weakly) stationary processes, the power spectral density is the *Fourier transform* of the autocorrelation function $acf(\tau)$:

$$\mathrm{psd}(k) = \widehat{\mathrm{acf}}(k) = \int_{-\infty}^{\infty} \mathrm{acf}(\tau)\, e^{-2\pi i k \tau} d\tau. \tag{3.29}$$

Therefore if we take the power spectral density $\mathrm{psd}(k)$ corresponding to the desired autocorrelation and multiply X_k by $\sqrt{\mathrm{psd}(k)}$ we arrive at the desired coloured noise in the frequency domain:

$$\widetilde{X}_k = X_k \sqrt{\mathrm{psd}(k)} \quad \text{for } 0 < k < n/2. \tag{3.30}$$

Note that X_0 is not scaled as X_0/n is the mean of the x_i, and for $k > n/2$ the coefficients have to be kept in symmetry: $\widetilde{X}(k) = \widetilde{X}(n-k)^*$. The result is then

transformed back to the time domain with the *inverse discrete Fourier transform*:

$$\widetilde{x}_j = \frac{1}{n}\sum_{k=1}^{N} \widetilde{X}_k e^{2\pi i \frac{k}{n} j}. \tag{3.31}$$

For the speed noise ψ the autocorrelation function is approximated by

$$\mathrm{acf}_{\psi}^{e-e}(\tau) = a e^{-\lambda_1 \tau} + (1-a) e^{-\lambda_2 \tau} \tag{3.32}$$

as shown in figure 3.9, which corresponds to a power spectral density of:

$$\mathrm{psd}_{\psi}(k) = \frac{2\lambda_1 a}{(\lambda_1^2 + 4\pi^2 k^2)} - \frac{2\lambda_2 (a-1)}{(\lambda_2^2 + 4\pi^2 k^2)}. \tag{3.33}$$

For the turning-angle noise ξ the autocorrelation function is approximated by $\mathrm{acf}_{\xi}(\tau) = (\tau+1)^L$. As the corresponding power spectral density does not have a particularly pleasant expression[11] we calculated $\mathrm{psd}_{\xi}(k)$ numerically with a discrete Fourier transform of $\mathrm{acf}_{\xi}(\tau)$.

3.3.2 Simulation of the Bumblebee Model

Given the complete model specification in section 3.2.8 we can now generate artificial bumblebee flight trajectories. To simulate the bumblebee model, we integrate its Langevin equations (3.20, 3.21) using the correlated noise terms $\xi_s(t)$ and $\psi(t)$ from section 3.3.1 and the estimated drift $g(s)$ from section 3.2.5. These Langevin equations are the stochastic differential equations of an *Itō process*. If instead a Stratonovic interpretation of the SDEs had been used, the estimation of of the drift terms would have needed a correction due to an induced *spurious drift*, as described e.g. in [38]. The numerical integration for the Langevin equations therefore has to be done with an Itō scheme. One of the most basic Itō integration schemes is the *Euler-Maruyama-scheme* (see section A.3), which we use here. Writing the Langevin equations (3.20, 3.21) of our model in Itō form,

[11] $\mathrm{psd}_{\xi}(k) = \frac{2\,{}_1F_2(1;1-\frac{L}{2},\frac{3}{2}-\frac{L}{2};-\pi^2k^2)}{L-1} + \frac{(2\pi)^L |k|^{L-1} \csc(\pi L) \sin(\frac{\pi}{2}(L+4|k|))}{\Gamma(L)}$
where ${}_1F_2$ is a generalized hypergeometric function and $\Gamma(L)$ is the gamma function.

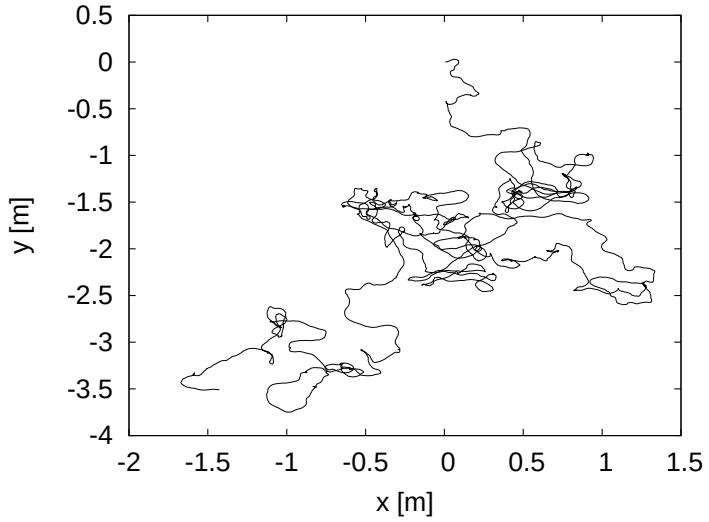

Figure 3.10: **Simulated trajectory of a bumblebee.** The complete model (see section 3.2.8) is simulated for 200 s ($= 10^5$ time steps) with an Euler-Maruyama scheme using the noise samples for ξ and ψ correlated beforehand (see section 3.3.1).

and discretising time with a time step Δt gives the following integration scheme:

$$\beta(t + \Delta t) = \xi_s(t), \tag{3.34}$$

$$s(t + \Delta t) = s(t) + g(s(t))\Delta t + \Delta\psi(t). \tag{3.35}$$

Notice that the noise strength of $\Delta\psi(t)$ has to be scaled in consistence with the time-step Δt used for the integration. In our simulations we used the time resolution $\Delta t = 0.02\,\mathrm{s}$ of the experimental data as a time delay for the integration, as it is already small enough. In rare cases where the Gaussian noise $\psi(t)$ would lead to a negative speed despite the positive drift $g(s)$ for $s < s_0$, we enforce a non-negative speed by setting $s(t) = 0$.

The dependence of the turning-angle distribution on the speed $s(t)$ adds a complication to the simulation. While the acceleration term $\psi(t)$ can be simply added to the speed in each time step of the integration, the turning-angle noise $\xi_s(t)$ is speed-dependent and cannot be generated in advance as described in section 3.3.1.

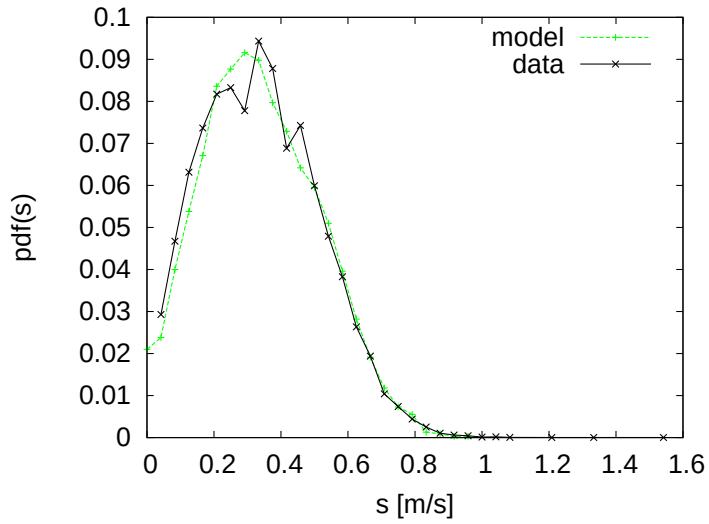

Figure 3.11: **Comparison of the speed-distributions.** The green (dashed) line shows the probability density $\mathrm{pdf}(s)$ extracted from the simulated data, the black (solid) line shows the experimental data of all bumblebees (≈ 45000 data points).

Instead the correlated Gaussian noise has to be scaled by its s-dependent standard deviation $\sigma_\beta(s)$ for each step (see section 3.2.6). As this happens after correlating the noise, this does not reproduce the autocorrelation of the turning-angle exactly. However the error made is acceptable in our case, as it is less than the errors from the estimation of acf_β. The resulting scheme for the turning-angle β can be written as

$$\beta(t + \Delta t) = \xi_s(t) = \sigma_\beta(s(t))\xi(t) \tag{3.36}$$

where $\xi(t)$ is the unscaled correlated Gaussian turning-angle noise.

A sample trajectory of a bumblebee simulated for 200 s using 10^5 time steps is shown in Fig. 3.10. The trajectory shows the typical switching of flight patterns between localized flights with low velocity and large turning angles, and faster movement with low sinuosity. Using the generated data we can now check the validity of the model by comparison to the experimental data of all bumblebees.

Model Comparison to Experimental Data

Figure 3.11 compares the probability density function $\mathrm{pdf}(s)$ of the speed extracted from the simulated data with the corresponding probability density function from the experimental data. Despite the fact that we made quite a lot of simplifications when building the bumblebee model, e.g. ignoring the influence of vertical movement, and that we used rather simple approximations for the estimated properties, e.g. the speed drift $\mathrm{g}(s)$, the distribution of speeds in the model matches the experimental data rather well.

Apart from the correct variance, the resulting turning angle distribution $\mathrm{pdf}(\beta)$ does not match the experimental data very well. This is not astonishing since the model simplified the turning angle noise $\xi_s(\beta)$ by assuming that its distribution is normal for all values of s. As discussed in section 3.2.6, the Gaussian approximation is only valid for low speeds. While this could have an effect for the short term dynamics, over a few time steps the accumulated sum of the turning angles becomes normal again due to the central limit theorem (see section 1.3.2).

The autocorrelation function of the turning-angle is shown in figure 3.8. There is a good (and not so astonishing) agreement between the autocorrelation of β in the experiment and in the model, which mostly shows that the generation of coloured noise works.

The autocorrelation $\mathrm{acf}_s(\tau)$ of the speed s, which is shown in figure 3.13, has to be looked at in more detail. While the model is in nice agreement with the autocorrelation from experimental stage (4) as can be seen in figure 3.12, it differs from stage (7) and especially stage (1), which show a stronger (positive) correlation for larger delay times τ. This difference can be explained by the following: the dynamics in stage (1), meaning without predation threat at the flowers, is dominated by short systematic flights between flowers. There the regular flights mean that the speed is autocorrelated over longer times. Under threat of predation in stage (4) the bumblebee reacts to the predators by breaking the regular search pattern, leading to a quicker decay of the autocorrelation. The data from stage (7) lies in the middle between the other two, since the bumblebees were trained, but have already partially forgotten about it. For evidence of the flower avoidance see section 2.2.3 and for the predator-induced anti-correlations see sections 2.2.4 and

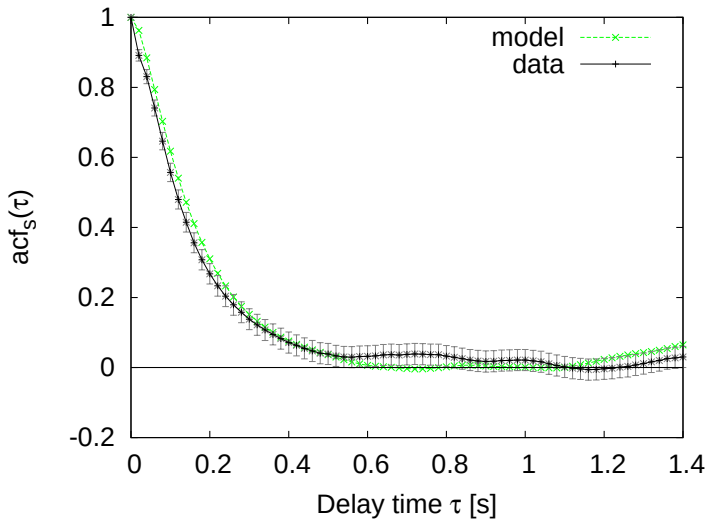

Figure 3.12: **Autocorrelation of bumblebee speed.** The green (dashed) line shows the autocorrelation extracted from the simulated data, the black (solid) line from the experimental data (from stage (4)) with two times the large-lag standard error (grey).

2.2.5. Since we are in this chapter not so much interested in the systematic flights from flower to flower, but rather on the free search flights away from the flowers, which are more pronounced in stage (4), we are happy with the matching of the data from stage (4) and the model. With regard to an application of the model to a bumblebee in a natural habitat, the presence of predators is also the default state.

Mean Square Displacement

Figure 3.14 shows the mean squared displacement (MSD) of the bumblebee position as determined from each experimental stage and from the simulated model data. While both, the experimental results and the simulation, show well-matching ballistic dynamics for low delay times $\tau < 0.5\,\mathrm{s}$, the diffusive behaviour for large τ, as emphasized by a linear fit, can only be seen in the model. Instead, the experimental data shows a saturation of the MSD around $0.08\,\mathrm{m}^2$ for $\tau > 1.6\,\mathrm{s}$. The saturation is to be expected, since the movement of the bumblebees is hindered by the flower zones and bounded by the walls. In the model the bumblebee flights

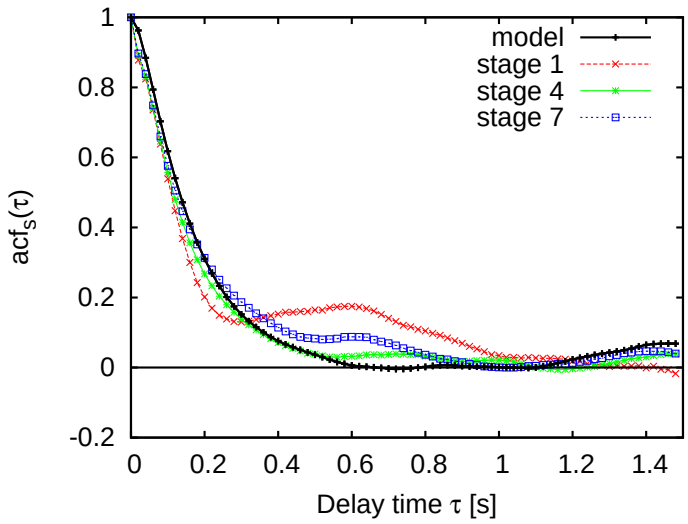

Figure 3.13: **Autocorrelation function of the bumblebee speed for different experimental stages.** The black line shows the autocorrelation extracted from the simulated data, the other lines show its estimates from the experimental data in stages (1),(4),(7) in red, green and blue respectively.

are however not constrained, leading to the deviation of the MSD for large delays.

3.4 Summary

The family of reorientation models has been often used to describe the correlated random walk of animals. We therefore tried to examine their suitability as a model to describe the foraging behaviour of bumblebees in consistency with the observed data. Instead of looking only at the distributions of the turning angle β and the speed s, we generalized the reorientation model by explicitly modelling their dynamics via generalized Langevin equations. Analysing movement data of the bumblebee experiment, we extracted information on the deterministic and stochastic terms of Eqs. (3.1, 3.2). We examined a deterministic part of the dynamics of (β, s) using a Markov approximation by estimating the drift coefficients of the Fokker-Planck equation corresponding to the Langevin equation. Any effects not captured by this drift term contributed to the correlated noise terms (see

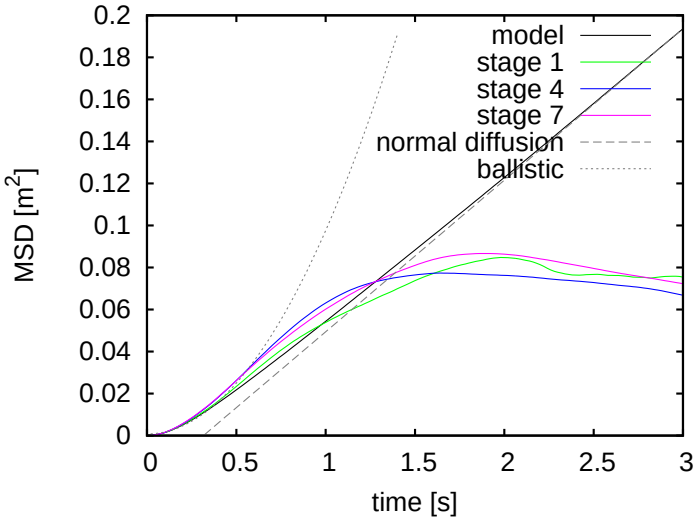

Figure 3.14: **Mean squared displacement.** The measured mean squared displacement of a bumblebee is shown for different experimental stages: stage (1) (green), stage (4) (blue), stage (7) (violet). In addition the MSD of the simulated data is shown (black), together with a linear fit (dashed grey) of its diffusive behaviour for large delays and a quadratic fit (dotted grey) of its ballistic short term behaviour.

sections 3.2.6, 3.2.7) in the resulting model (section 3.2.8). With the estimation of the turning angle drift $h(\beta)$ we found that while the usual assumption of i.i.d. turning angles is not valid in our case, the lack of a non-trivial drift and the weak autocorrelation of ξ_s are consistent with the usual reorientation model (see section 3.3.2). However, our generalized model exhibits significant differences in the non-trivial deterministic part $g(s)$ of the speed change ds/dt and the speed dependence of the turning angles. In terms of active Brownian particle models ([23, 58], see section 1.3.4) we described the two-dimensional bumblebee movement by a particle with a non-linear friction term $g(s)$ depending and acting only on the speed, driven by multiplicative coloured noise with different correlations for the angle component and the speed component of the velocity. While this combination of complications might make it difficult to treat the system analytically, progress in this direction has been made [108, 57].

To validate the bumblebee flight model, we simulated it by stochastic integra-

tion using correlated noise in section 3.3. The comparison to the data has shown that the resulting model agrees very well with the experimental data despite the approximations we made for the model.

Given that the experiment which yielded our data is rather small and provided the bumblebees with an artificial environment, it would be interesting to apply our new model to free-flying bumblebees to reveal how much the results depend on the specific set-up. This would clarify whether the flight behaviour seen in the laboratory experiment survives as a flight mode for foraging in a patch of flowers in an intermittent model, with an additional flight mode for long flights between flower patches. The analysis of data from other flying insects and birds by using our model could be interesting in order to examine whether the piecewise linear nature of the speed drift and the trivial drift of the turning angle are a common feature. In view of understanding the small-scale bio-mechanical origin of flight dynamics, our model might serve as a reference point for any more detailed dynamical modelling. That is, we would expect that any more microscopic model should reproduce our dynamics after a suitable coarse graining over relevant degrees of freedom.

Chapter 4

Fluctuation Relations

In this chapter we will at first briefly motivate fluctuation relations (see section 4.1). We then summarize a study of the interplay of fluctuation dissipation relations and fluctuation relations for a concrete class of stochastic processes in section 4.2. The main results of this study were done by A. V. Chechkin and R. Klages [4], I contributed with work on spectral densities of autocorrelation functions as shown in section 4.3, and some discussions. Specifically I focused on the non-negativity of spectral densities for given autocorrelation functions.

While the aim of this study — examining *anomalous fluctuation relations*, that is deviations from (normal) fluctuation relations — is not related to foraging, the set-up has similarities to our data analysis in chapters 1 and 3 and some related foraging models: a stochastic process described by a generalized Langevin equation albeit a different generalization than discussed here, see section 1.3.4 and a non-trivial autocorrelation function where anti-correlations (see e.g. section 2.2.4) and heavy tails play a decisive role. In contrast to the previous chapters, here the Langevin equation will be generalized through a friction kernel (Eq. (4.3) in section 4.2) giving another way to model autocorrelated processes.

4.1 Introduction to Fluctuation Relations

For isolated systems, the *second law of thermodynamics* states that the entropy of the system cannot decrease. While the second law is fundamental, the restric-

tions on its applicability made it desirable to find more general and more precise relations for the entropy production. From statistical mechanics the second law has been understood to be strictly valid only in the thermodynamic limit. However, given any real systems of finite size the law has to be interpreted probabilistically: for large sizes the probability of a violation of the second law decreases and becomes effectively negligible. However, with the increased interest in smaller systems, e.g. single macromolecules, those probabilities become important.[1] Therefore there was a need to find laws which hold for small systems, too, and preferably these laws would be also applicable in non-equilibrium situations since many systems of interest, e.g. all biological systems, are not isolated but externally driven systems.

Consequently the probability density function $p(\xi_t)$ of the entropy production ξ_t over a time interval t starting from an equilibrium state and ending in a non-equilibrium steady state has been studied in more detail. For large classes of systems [109, 110, 111, 112] it has been shown that the second law can be generalized to the so-called *fluctuation relation* [113, 114, 115]

$$\ln \frac{p(\xi_t)}{p(-\xi_t)} = \xi_t \ , \quad \text{for } \xi_t \geq 0 \ . \tag{4.1}$$

This means that negative entropy productions are less likely than positive productions by a factor $\exp(\xi_t)$. The expectation of the entropy production is therefore still non-negative which then leads to the second law in the thermodynamic limit:

$$< \xi_t >= \int_{-\infty}^{\infty} \xi_t p(\xi_t) d\xi_t = \int_0^{\infty} \xi_t p(\xi_t) \left(1 - \exp^{-\xi_t}\right) d\xi_t \geq 0 \ . \tag{4.2}$$

While the quantity of interest is often the entropy production, similar laws exist for related quantities, e.g., the accumulated work to switch between two equilibrium steady states [114]. Since the fluctuation relation gives quite detailed information about $p(\xi_t)$, it is interesting for which processes it holds under which conditions. We therefore examine its validity for the class of Gaussian stochastic

[1] A simple example is an ideal gas of n particles with uniformly drawn initial positions in a cubical container. Without external forcing the probability to find all particles in the left half of the cube after a time t converges to 2^{-n} for $t \to \infty$, which is only insignificant for large n.

processes in the next section.

4.2 Fluctuation Relations in Gaussian Stochastic Processes

Let us look at a *Gaussian stochastic process* given by a *generalized Langevin equation* [116]:

$$\int_0^t \dot{x}(t')\gamma(t-t')dt' = \frac{F}{m} + \xi(t) \tag{4.3}$$

with a constant force F, mass m and a friction kernel $\gamma(t)$. The noise $\xi(t)$ should be stationary, Gaussian and have an ensemble average $\langle \xi(t) \rangle = 0$ for all t, but it does not have to be uncorrelated (white), i.e. coloured noise is allowed. The question of interest is, under which conditions the work $W := Fx$ obeys the (normal) *transient work fluctuation relation* [113]:

$$\ln \frac{p(W,t)}{p(-W,t)} = \frac{W}{k_B T} \tag{4.4}$$

where $p(W,t)$ is the probability density function of the work, k_B is the Boltzmann constant and T is the temperature of the system. This depends critically on the autocorrelation function[2] $\mathrm{acf}(\tau) := \langle \xi(t)\xi(t+\tau) \rangle_t$ of the noise and if and how it relates to the friction kernel. For these Gaussian stochastic processes it can be shown that:

$$\ln \frac{p(W,t)}{p(-W,t)} = \frac{2\langle x(t) \rangle}{F\sigma_x^2(t)} W \tag{4.5}$$

which means that whether the fluctuation relation (Eq. (4.4)) holds depends on the mean displacement $\langle x(t) \rangle$ and the mean square displacement (MSD) $\sigma_x^2(t)$. In cases where the MSD does not scale linearly in time t, the diffusion is called *anomalous*.

The fluctuation relation (4.4) does not hold in general without further restrictions on $\mathrm{acf}(\tau)$ and $\gamma(\tau)$. For *internal noise* the source of the friction and the noise are the same, which gives rise to the *fluctuation-dissipation relation of the*

[2] In contrast to the other chapters, $\mathrm{acf}(\tau)$ is not normalized here.

second kind (FDR II):

$$\mathrm{acf}(\tau) = \frac{k_B T}{m}\gamma(\tau) \; . \tag{4.6}$$

In [4] we show that the fluctuation relation follows in this situation, and which implications exist between the transient fluctuation relation and the fluctuation-dissipation relations. A special case of a model with internal noise is *Brownian motion*, where the friction is proportional to a delta function: $\gamma(\tau) = 2\gamma\delta(\tau)$ with the friction constant γ, and the noise is white: $\mathrm{acf}(\tau) = (2\gamma k_B T/m)\delta(\tau)$.

For *external noise* there is no FDR II, which means that in general one cannot expect the fluctuation relation to hold for arbitrary autocorrelation $\mathrm{acf}(\tau)$ and friction $\gamma(\tau)$. But even for a simple friction $\gamma(\tau) = 2\gamma\delta(\tau)$ there is a variety of behaviour in the MSD which we investigated. The behaviour can lead to normal and anomalous diffusion and varying validity of the fluctuation relation: we give an overview of the results in section 4.4 (see [4] for details). The results depend on the properties of the autocorrelation function: power law tails are needed for anomalous diffusion and the MSD and the fluctuation relation critically depend on its exponent. The behaviour also depends on whether the autocorrelation is *persistent* (decay to 0 from above for large τ) or *anti-persistent* (anti-correlation and convergence to 0 from below), and in the case of anti-persistence whether it is pure (see section 4.3.2). In the following section we therefore look at suitable classes of autocorrelation functions.

4.3 Spectral Densities of Autocorrelation Functions

For an autocorrelation function $\mathrm{acf}(\tau)$ of a stochastic process the corresponding *spectral density*:

$$S(\omega) := \int_{-\infty}^{\infty} e^{-i\omega\tau}\mathrm{acf}(\tau)d\tau \tag{4.7}$$

has to be non-negative for all $\omega \geq 0$ for consistency [117]. Since we want to construct examples for different classes of stochastic processes by choosing the distribution of the noise and the autocorrelation, we checked the non-negativity for a few classes of autocorrelation functions. In this case we are interested in anti-correlations and/or heavy-tailed correlations (see [4]), which we investigate

in the following.

4.3.1 Power-Law Decay

A simple class of autocorrelation functions with power-law decay and a finite value $c_\beta = \mathrm{acf}_\beta(0)$ is:

$$\mathrm{acf}_\beta(\tau) = \frac{c_\beta}{\left(1 + \frac{|\tau|}{\delta}\right)^\beta} \tag{4.8}$$

where $0 < \beta < \infty$, $c_\beta \geq 0$, $\delta > 0$. The corresponding spectral density is:

$$S_\beta(\omega) = 2c_\beta \delta^\beta \omega^{\beta-1} \left(\cos(\omega\delta) \int_{\omega\delta}^{\infty} \frac{\cos x}{x^\beta} dx + \sin(\omega\delta) \int_{\omega\delta}^{\infty} \frac{\sin x}{x^\beta} dx \right). \tag{4.9}$$

In order to test the non-negativity of the corresponding spectral density the constants c_β and δ (which is effectively just a scaling factor for the frequencies ω) are not important. For convenience we set $c_\beta = \frac{1}{2}$ and $\delta = 1$ giving:

$$S_\beta(\omega) = \omega^{\beta-1} \left(\cos(\omega) \int_{\omega}^{\infty} \frac{\cos x}{x^\beta} dx + \sin(\omega) \int_{\omega}^{\infty} \frac{\sin x}{x^\beta} dx \right). \tag{4.10}$$

As $\omega \geq 0$ it suffices to examine $I_\beta(\omega) := \omega^{1-\beta} S_\beta(\omega)$ which then can be simplified as follows:

$$\begin{aligned}
I_\beta(\omega) &= \cos\omega \int_{\omega}^{\infty} \frac{\cos x}{x^\beta} dx + \sin\omega \int_{\omega}^{\infty} \frac{\sin x}{x^\beta} dx && (4.11)\\
&= \int_{\omega}^{\infty} \frac{\cos x \cos\omega + \sin x \sin\omega}{x^\beta} dx && (4.12)\\
&= \int_{\omega}^{\infty} \frac{\cos(x-\omega)}{x^\beta} dx && (4.13)\\
&= \int_{0}^{\infty} \frac{\cos x}{(x+\omega)^\beta} dx && (4.14)\\
&= \frac{\sin(x)}{(x+\omega)^\beta}\bigg|_0^\infty + \beta \int_0^\infty \frac{\sin x}{(x+\omega)^{\beta+1}} dx && (4.15)\\
&= \beta \int_0^\infty \frac{\sin x}{(x+\omega)^{\beta+1}} dx && (4.16)
\end{aligned}$$

which is positive because the denominator is strictly increasing: the integral over each positive part from root $2\pi k$ to root $2\pi k + \pi$ is always larger than the next negative part from $2\pi k + \pi$ to $2\pi(k+1)$ for all natural k, as shown below:

$$I_\beta(\omega) = \beta \int_0^\infty \frac{\sin x}{(x+\omega)^{\beta+1}} dx \tag{4.17}$$

$$= \beta \sum_{k=0}^{\infty} \left(\int_{2\pi k}^{2\pi k+\pi} \frac{\sin x}{(x+\omega)^{\beta+1}} dx + \int_{2\pi k+\pi}^{2\pi(k+1)} \frac{\sin x}{(x+\omega)^{\beta+1}} dx \right) \tag{4.18}$$

$$= \beta \sum_{k=0}^{\infty} \left(\int_{2\pi k}^{2\pi k+\pi} \frac{\sin x}{(x+\omega)^{\beta+1}} dx + \int_{2\pi k}^{2\pi k+\pi} \frac{\sin(x+\pi)}{(x+\pi+\omega)^{\beta+1}} dx \right) \tag{4.19}$$

$$= \beta \sum_{k=0}^{\infty} \int_{2\pi k}^{2\pi k+\pi} \left(\frac{\sin x}{(x+\omega)^{\beta+1}} + \frac{-\sin x}{(x+\pi+\omega)^{\beta+1}} \right) dx \tag{4.20}$$

$$= \beta \sum_{k=0}^{\infty} \int_{2\pi k}^{2\pi k+\pi} \sin x \left(\frac{1}{(x+\omega)^{\beta+1}} - \frac{1}{(x+\pi+\omega)^{\beta+1}} \right) dx > 0 \tag{4.21}$$

since $\sin x \geq 0$ for $x \in [2\pi k, 2\pi k + \pi]$ and $\frac{1}{(x+\omega)^{\beta+1}} > \frac{1}{(x+\pi+\omega)^{\beta+1}}$. This means the spectral density is non-negative for all β.

4.3.2 Anti-Correlation

An example for autocorrelation functions which show anti-correlation is given by:

$$\mathrm{acf}(\tau) = ke^{-a|\tau|} - (k-1)e^{-b|\tau|} \tag{4.22}$$

for $a > b > 0$ and $k > 1$. Since the autocorrelation function is an even function, the integral of the product with the odd function $\sin$ yields 0, so the corresponding spectral density is:

$$S(\omega) = \mathcal{F}(\mathrm{acf})(\omega) = 2 \int_0^\infty \cos(\omega\tau) \left(ke^{-a\tau} - (k-1)e^{-b\tau} \right) d\tau \ . \tag{4.23}$$

Using $\int_0^\infty e^{-px} \cos(qx) dx = \frac{p}{p^2+q^2}$ for all $p > 0$:

$$S(\omega) = 2k \frac{a}{a^2+\omega^2} - 2(k-1) \frac{b}{b^2+\omega^2} \ . \tag{4.24}$$

$S(\omega) \geq 0$ iff

$$2k\frac{a}{a^2+\omega^2} \geq 2(k-1)\frac{b}{b^2+\omega^2} \tag{4.25}$$

$$\Leftrightarrow \quad \frac{a}{b} \geq \frac{k-1}{k}\frac{a^2+\omega^2}{b^2+\omega^2}, \tag{4.26}$$

which means that $S(\omega) \geq 0$ for all ω iff

$$\frac{a}{b} \geq \max_{\omega}\left(\frac{k-1}{k}\frac{a^2+\omega^2}{b^2+\omega^2}\right) \tag{4.27}$$

$$\Leftrightarrow \quad \frac{a}{b} \geq \frac{k-1}{k}\max_{\omega}\left(1+\frac{a^2-b^2}{b^2+\omega^2}\right) \tag{4.28}$$

$$\Leftrightarrow \quad \frac{a}{b} \geq \frac{k-1}{k}\left(1+\frac{a^2-b^2}{b^2}\right) \tag{4.29}$$

$$\Leftrightarrow \quad \frac{a}{b} \geq \frac{k-1}{k}\frac{a^2}{b^2} \tag{4.30}$$

$$\Leftrightarrow \quad \boxed{\frac{a}{b} \leq \frac{k}{k-1}}. \tag{4.31}$$

In this case the consistency of the corresponding stochastic process depends on a proper choice of parameters a, b and k. The condition is quite restrictive; a positive example is: $a = 84$, $b = 67/2$, $k = 158/153$. However, the integral over the autocorrelation $\int_0^\infty \mathrm{acf}(\tau)d\tau$ is only non-negative for $\frac{a}{b} \geq \frac{k}{k-1}$, which means the only valid case is the *pure anti-persistent* case with $\int_0^\infty \mathrm{acf}(\tau)d\tau = 0$ where $\frac{a}{b} = \frac{k}{k-1}$.

4.3.3 Anti-Correlation and Power-Law Tail

Similar in shape (starting positive, then with anti-correlation which converges to 0 from below) but with a power-law tail is the following autocorrelation function (Fig. 4.1):

$$\mathrm{acf}(\tau) = (1+\alpha)e^{-|\tau|/\delta} - \alpha(1+|\tau|)^{-\beta} \tag{4.32}$$

where $\delta > 0$, $\alpha > 0$ and $\beta \geq 1 + \frac{\alpha}{(1+\alpha)\delta}$ to ensure that $\int_0^\infty \mathrm{acf}(\tau)d\tau \geq 0$.

In this case we did not find an analytical proof for the non-negativity of the spectral density but could only show it numerically for a wide range of parameters.

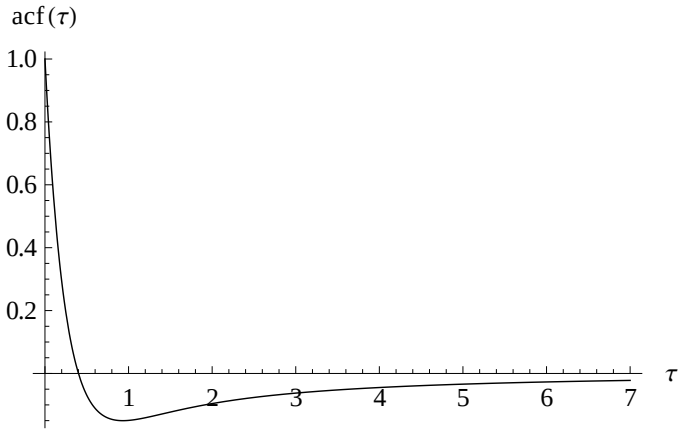

Figure 4.1: **Autocorrelation with anti-correlation and power-law tail.** Shown is an example for the class of autocorrelations with an anti-correlated heavy tail given in Eq. (4.32). The chosen parameters are: $\alpha = \frac{1}{2}, \beta = \frac{3}{2}, \delta = \frac{1}{4}$.

A closed form expression of the spectral density can still be computed, e.g. by Mathematica, but it leads to an expression which is not easy to analyze:

$$I(\omega) = \frac{(1+\alpha)\delta}{1+w^2\delta^2} + \frac{\alpha\, {}_1F_2(1;\, 1-\frac{\beta}{2}, \frac{3}{2}-\frac{\beta}{2};\, -\frac{w^2}{4})}{1-\beta} - \frac{\alpha\pi w^{-1+\beta}\sin(w+\frac{\pi\beta}{2})}{\Gamma(\beta)\sin(\pi\beta)}. \tag{4.33}$$

Here Γ is the *gamma function* $\Gamma(z) := \int_0^\infty t^{z-1}e^{-t}\,\mathrm{d}t$ and ${}_1F_2$ is a *generalized hypergeometric function* [118] given by:

$${}_1F_2(a_1, a_2; b; x) := \sum_{x=0}^{\infty} \frac{(a_1)_k (a_2)_k}{(b)_k} \frac{x^k}{k!} \tag{4.34}$$

with the *Pochhammer symbol* $(a)_k := \Gamma(a+k)/\Gamma(a)$.

Plotting $I(\omega)$ for a many different parameters (Fig. 4.2) gives evidence for the non-negativity of the spectral density. While this is no proof that the autocorrelation function is suitable for all possible parameters, it gives some indication that this choice of $\mathrm{acf}(\tau)$ is of use in building stochastic models with anti-persistence and anomalous diffusive behaviour.

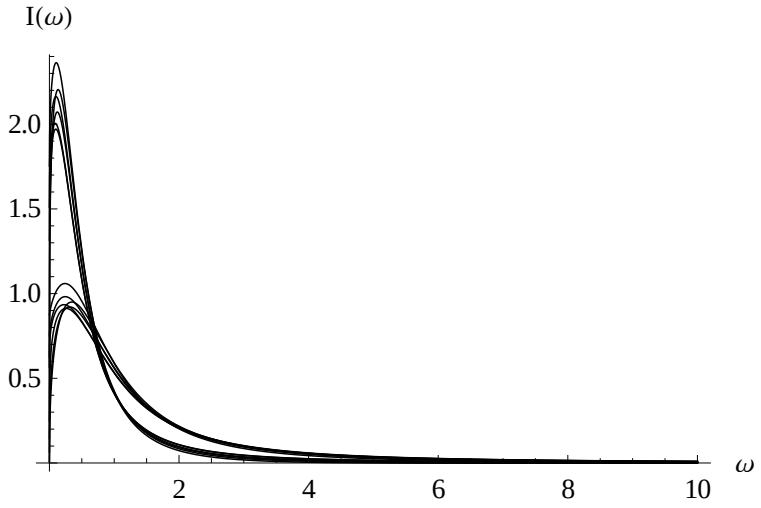

Figure 4.2: **Numerical evidence for the non-negativity of $I(\omega)$.** Plotted are the spectral densities (see Eq. (4.33) corresponding to the anti-correlated autocorrelations with power-law tail given by Eq. (4.32) for a variety of parameters: searching through many valid combinations of parameters we consistently found that $I(\omega) \geq 0$ for all ω.

4.4 Fluctuation Relations and MSD for External Noise

After checking the consistency of different classes of autocorrelation functions via their spectral density in section 4.3, we can now finish this chapter by discussing the validity of the transient work fluctuation relation for Gaussian stochastic processes given by the Langevin equation (4.3) for the case of *external noise* (see section 4.2). Here we restrict ourselves to friction kernels $\gamma(\tau) = 2\gamma\delta(\tau)$ without memory – the analysis would be severely more complicated by simultaneously considering friction kernels and noise autocorrelations which are unrelated and both non-trivial. This section is an overview of the results in [4], where a more detailed discussion can be found. Notice that the main results of this paper were **not** derived by me: they are included below only to complete the discussion in section 4.2. For external noise, interesting anomalies of the diffusion can occur due to autocorrelation functions $\mathrm{acf}(\tau)$ whose tail can be approximated by a power-law.

We therefore examine autocorrelation functions of the following shape:

$$\mathrm{acf}(\tau) = \begin{cases} \mathrm{acf}_1(|\tau|) & \text{for } |\tau| \leq \Delta \\ \mathrm{acf}_2(|\tau|) & \text{for } |\tau| > \Delta \end{cases}, \tag{4.35}$$

with $\mathrm{acf}_1(0) > 0$, $\Delta > 0$, $\mathrm{acf}_1(\Delta) = \mathrm{acf}_2(\Delta)$, and tails

$$\mathrm{acf}_2(\tau) = C_\beta \left(\frac{|\tau|}{\Delta} \right)^{-\beta}. \tag{4.36}$$

The autocorrelations discussed in section 4.3.1 and 4.3.3 are examples with the same asymptotic behaviour for the persistent case with $C_\beta > 0$ and $0 < \beta < \infty$, and for the anti-persistent case with $C_\beta < 0$ and $1 < \beta < \infty$ respectively.

The diffusive behaviour, i.e. the mean squared displacement, and the validity of the fluctuation relation for $t \gg \Delta$ now depend on the sign of C_β and on β. The right hand side of the fluctuation relation (Eq. 4.4) and the MSD can both be calculated explicitly [4]. The results, which are discussed below, are summarized in Table 4.1 using the following constants as abbreviations: $D := \frac{1}{\gamma^2} \int_0^\infty \mathrm{acf}(\tau) d\tau$, $T_{\text{eff}} := \frac{mD}{k_B \gamma}$, and $R := \int_0^\infty \tau \mathrm{acf}(\tau) d\tau$.

For persistent external noise with fast enough correlation decay ($\beta > 1$) the process exhibits normal diffusion and a *generalized fluctuation relation* holds, where the temperature T is replaced in Eq. (4.4) with an *effective temperature* T_{eff}. For $\beta \leq 1$ the process is instead superdiffusive and the fluctuation relation is *anomalous*, i.e. it does not hold.

Given anti-persistent external noise, an exponent $\beta \leq 1$ would be inconsistent with $D \geq 0$ (compare section 4.3.3), and is therefore excluded. Additionally, if D is strictly positive, then the MSD shows normal diffusion for large t and a generalized fluctuation relation as above. For the remaining purely anti-persistent cases (with $D = 0$) the fluctuation relation does not hold and the diffusive behaviour ranges from subdiffusion for $1 < \beta < 2$ to localization, i.e. an asymptotically constant MSD, for $\beta > 2$.

The transition points between different types of behaviour, i.e. $\beta = 1$ for persistent and $\beta = 2$ for anti-persistent noise, show additional logarithmic terms in

Table 4.1: **Work fluctuation relation and MSD for Gaussian stochastic processes with external noise.** Shown are the mean squared displacement and the right hand side of the fluctuation relation (4.1) for $t \gg \Delta$ depending on the exponent β of the autocorrelation decay for persistent and purely anti-persistent noise.

<table>
<tr><th></th><th colspan="2">persistent ($C_\beta > 0$)</th><th colspan="2">pure anti-persistent ($C_\beta < 0, D = 0$)</th></tr>
<tr><td>β</td><td>MSD</td><td>$\ln \frac{p(W,t)}{p(-W,t)}$</td><td>MSD</td><td>$\ln \frac{p(W,t)}{p(-W,t)}$</td></tr>
<tr><td>$0 < \beta < 1$</td><td>$\sim t^{2-\beta}$</td><td>$\sim \frac{W}{t^{1-\beta}}$</td><td colspan="2" rowspan="2">This regime does not exist.</td></tr>
<tr><td>$\beta = 1$</td><td>$\sim t \ln\left(\frac{t}{\Delta}\right)$</td><td>$\sim \frac{W}{\ln(t/\Delta)}$</td></tr>
<tr><td>$1 < \beta < 2$</td><td rowspan="3">$\sim 2Dt$</td><td rowspan="3">$\sim \frac{W}{k_B T_{\text{eff}}}$</td><td>$\sim t^{2-\beta}$</td><td>$\sim W t^{\beta-1}$</td></tr>
<tr><td>$\beta = 2$</td><td>$\sim \ln\left(\frac{t}{\Delta}\right)$</td><td>$\sim \frac{Wt}{\ln(t/\Delta)}$</td></tr>
<tr><td>$\beta > 2$</td><td>$\sim \frac{2}{\gamma^2}\,|R|$</td><td>$\sim \frac{\gamma}{m|R|} Wt$</td></tr>
</table>

both the MSD and the fluctuation relations. In [4] it is shown that these terms also appear when looking at processes with exponents β near the transition points, if one does not consider the asymptotic behaviour for large t, but instead examines the MSD or the fluctuation relation at intermediate times. This is of relevance to the analysis of experimental data, since the time scales which can be explored experimentally are typically restricted. This means that the asymptotic regime might be out of reach and only behaviour on intermediate time scales are accessible. The topic of experimental observations of anomalous fluctuation relations – including logarithmic corrections – is discussed in more detail in [4].

In summary, it has been shown that a large variety of diffusive behaviours can occur for Gaussion stochastic processes given by the generalized Langevin equation (4.3). We discussed the role of the autocorrelation function of the external noise for the flucutation relation and for the mean squared displacement. We also checked the non-negativity of corresponding spectral densities in order to find consistent classes of autocorrelations, which can be used as examples for Gaussian stochastic processes showing different kinds of anomalous behaviour.

Thesis Summary

In this thesis we investigated how the movement of foraging bumblebees can be analysed in terms of stochastic models. Specifically, we examined the effect the presence of predators has on the search patterns of the bumblebees, and the applicability of a generalised reorientation model to describe experimental flight data.

A large part of the difficulty in the field of foraging is the multitude of different biological factors which influence foraging behaviour. We therefore started our discussion in chapter 1 with a general introduction to the most relevant factors. As a variety of stochastic models has been proposed for the analysis of animal movements, we also described the most common classes of foraging models. On this background we discussed the idea of optimal foraging – in particular the mathematical Lévy flight hypothesis. We argued that the strict conditions of the hypothesis on the specific foraging situation under investigation make it doubtful whether much evidence for the hypothesis should be expected when examining experimental data. The validity of the biological Lévy flight hypothesis, i.e., whether real animals perform Lévy walks on their food search, has been an influential question in the last years. We ended the chapter with the conclusion that, while the Lévy flight hypothesis has been influential in spurring the cooperation between the research communities on foraging and the stochastic processes, its usefulness as a paradigm under which foraging animals are studied is limited.

In chapter 2 and 3 we analysed data of a laboratory experiment on bumblebee search flights from two different points of view. In chapter 2, our focus was the influence that the threat of predation has on the movement behaviour of foraging bumblebees. Interestingly we found that the predatory threat affects the bumblebees' movement, which already showed two different flight modes before introducing any threats. While we found local changes in the behaviour near the food

sources due to the introduction of artificial spiders as predators, its effect on the search flights from flower to flower was more interesting: we found that change of behaviour was primarily visible in the velocity autocorrelation functions, which we explained by a simple stochastic model, and not in velocity distributions, as we originally expected. We also discussed the connection of the importance of the autocorrelation in this case study with the Lévy flight hypothesis.

We switched our point of view in chapter 3 to the question of how to develop a stochastic model for the bumblebee movements from the experimental data. We approached the problem by generalising a reorientation model and extracting the coefficients of its generalised Langevin equations from the data. After a discussion of the main differences of the resulting model — which is a variation on active Brownian particles — and simpler reorientation models, we validated our model by simulation and comparison to the observed data.

After the analysis of foraging animals and especially bumblebees in the first chapters, the second theme of generalised Langevin equations reappeared in chapter 4. Here we gave a brief introduction to fluctuation relations and discussed them for Gaussian stochastic processes given by a different generalisation of the Langevin equation. In this context we came back to the analysis of autocorrelation functions, that is to say, we checked specific examples of functions for their validity as autocorrelation functions of Gaussian stochastic processes by examining the corresponding spectral densities, and we finished the chapter with a short discussion of fluctuation relations for external noise.

Appendix A

Appendix

A.1 Error Analysis

A.1.1 Standard Error of the Mean

Let $x_1, \ldots, x_n$ be n random variables all drawn independently from one distribution with population mean μ, population standard deviation σ and variance σ^2. Unbiased estimators of μ and σ^2 are the *sample mean* $m = \sum_{i=1}^{n} x_i/n$ and the *sample variance* $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - m)^2$, giving the (biased) *sample standard deviation* s. The *standard error* SE of a statistic is the standard deviation of the statistic, and describes the size of the error made when estimating the underlying statistical parameter by the statistic. In the case of the estimation of the population mean μ by the sample mean m the standard error SE_m can be approximated by

$$SE_m = \frac{s}{\sqrt{n}},$$

which is related to the standard deviation σ_m of the sample mean. The standard error can be used to derive confidence intervals for the mean, e.g. the 95% confidence intervals $CI = [m - 1.96SE_m, m + 1.96SE_m]$, where 1.96 is approximately the 0.975-quantile of a Gaussian distribution.

A.1.2 Confidence Intervals for Standard Deviations

To compute 95% confidence intervals for the sample standard deviation s (see section A.1.1) we first compute the confidence intervals for the sample variance s^2. Under the assumption that the random variables $x_1, \ldots, x_n$ are drawn from a normal distribution, the sample variance s^2 has a χ^2 distribution with $n-1$ degrees of freedom

$$s^2 \sim \chi^2_{n-1} = \frac{1}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)} x^{\frac{n-1}{2}-1} e^{-\frac{x}{2}}$$

with quantiles q_i. For a confidence level of $\alpha = 0.95$ the confidence interval of s^2 is then given by $CI_{s^2} = [s^2(n-1)/q_{(1+\alpha)/2}, s^2(n-1)/q_{(1-\alpha)/2}]$ and the corresponding interval for s is given by taking the square root of the boundaries.

A.1.3 Large-Lag Standard Error of Autocorrelation Functions

When computing confidence intervals for an autocorrelation function

$$\mathrm{acf}(\tau) = \frac{\langle (x_t - m)(x_{t+\tau} - m) \rangle}{s^2}$$

at time-lag $\tau = k\Delta t$ for a given time series $x_0, x_\tau, \ldots, x_{(n-1)\tau}$ with mean m and variance s^2, it is important to realize that the coefficients for different time-lags are not necessarily independent – and neither are their errors. Instead of assuming an uncorrelated time-series for the error analysis, which is unrealistic for many applications, the underlying assumption for the *large-lag standard error* for a time-lag τ is, that the autocorrelation coefficients for higher lags are negligible while the ones for lower lags might not be zero. This assumption [119] gives the following approximation $SE_{\mathrm{acf}(\tau)}$, called *large-lag standard error* [120], of the error of the autocorrelation function at lag τ:

$$SE_{\mathrm{acf}(\tau)} = SE_{\mathrm{acf}(k\Delta t)} = \sqrt{\frac{1}{n}\left(1 + 2\sum_{l=0}^{k-1} \mathrm{acf}\,(l\Delta t)^2\right)}.$$

A.2 Data Cleaning

The experimental flight data contained various artefacts, which had to be accounted for. The recorded time series have been visually inspected to check for obvious recording errors, e.g., single data points which lie far from the otherwise smooth trajectory immediately before and after the outlier. These errors have been marked as invalid and treated as a gap in the data. Other sources of artefacts could be dealt with automatically, e.g., the times when the bumblebees were not flying but crawling on a surface (see section A.2.1), and gaps in the recorded data (see section A.2.3).

A.2.1 Exclusion of Crawling

For the analysis of the bumblebee flights, parts of the experimental data had to be excluded: the cameras tracking the bumblebees not only recorded the flight trajectories, but also the crawling of the bumblebees on the objects in the flight arena: the walls of the arena and especially the artificial flowers (see section 2.1). Therefore the data has been filtered: all recorded positions of bumblebees within 1 cm of the flowers, including the landing platform and the mechanical traps, have been excluded from any analysis to capture bumblebee flights only. The size of this boundary is based on the size of the bumblebees, which have a height of approximately 1 cm. While a smaller cut-off would not exclude all crawling behaviour, the cut-off can be increased robustly within reasonable bounds. We have checked that, e.g. a 2 cm cut-off does not have any influence on any of the analysed quantities, as the amount of the data which would be excluded in addition is very small.

A.2.2 Flower Zones

Analysing the experimental bumblebee flight data, a distinction had to be made between the space near foraging flowers and the space away from them in the rest of the flight arena (see section 2.1). For that reason *flower zones* were defined as the following: for each flower a rectangular box with width and height $w = h = 9\,\mathrm{cm}$ is centred on the flower. The back side is at the foraging wall, while the

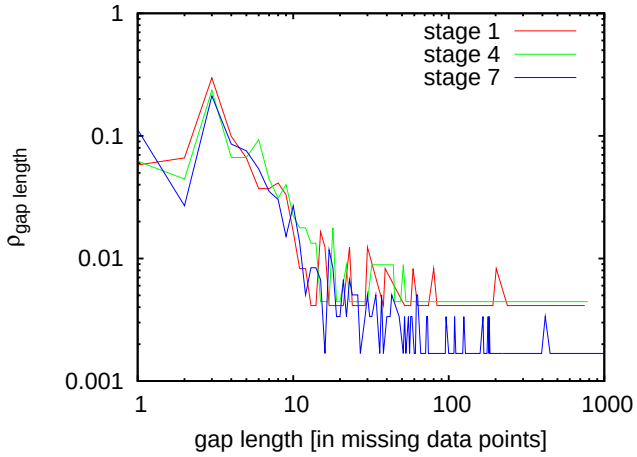

Figure A.1: **Distribution of gap-lengths in the experimental data.** The gaps with small gap-lengths (counted in missing data points/frames) can be interpolated to gain more complete trajectories.

front extends into the arena with a depth $d = 7\,\mathrm{cm}$, including the whole foraging platform and some space above. For details, see [1]. Notice that, while nearly all crawling behaviour (see section A.2.1) of the bumblebees happened on the flower platforms, this data is *not* included when we speak of data "in flower zones".

This separation of data is used when distinguishing different kinds of behaviour in section 2.2, while in chapter 3 all data inside the flower zones is excluded in order to focus on the "free" flight behaviour instead of on the flower-bumblebee interaction. This reduces the available data for chapter 3 to ≈ 49000 data points in a single experimental stage.

A.2.3 Gaps in the Experimental Data

The experimental bumblebee flight data contains quite a large number of gaps due to e.g. measurement errors and bumblebees leaving the region observed by cameras: a small region near the wall opposite to the flower wall was not captured by the cameras. For calculating quantities which depend on the availability of seamless time series, e.g. autocorrelations, small gaps in the time series have been interpolated linearly, instead of splitting the trajectory into two independent

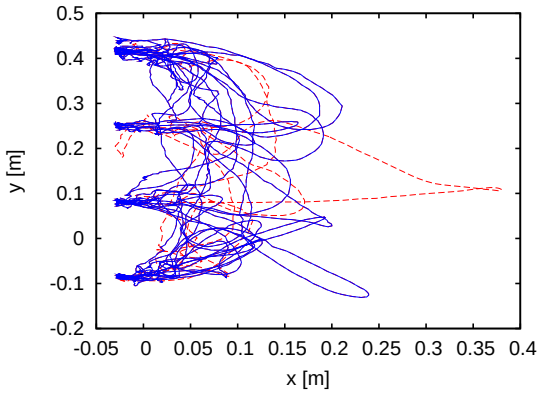

Figure A.2: **Additional data after gap interpolation.** Complete flower-to-flower trajectories of a single bumblebee in stage (7) without interpolation of gaps (blue), and additional complete trajectories after gap filling (red, dashed).

parts. As the number of gaps was small the correlations for short times were not affected, however, the interpolation increased the usable data for long time delays. Trajectories were split at larger gaps, for example when entering a flower zone, to exclude correlations induced by flower visits.

Fig. A.1 shows that most of the gaps have a short duration, which means that a conservative approach of interpolating only gaps no longer than 5 time steps ($= 0.1\,\mathrm{s}$) already gives most of the benefit in making more complete trajectories available (see Fig. A.2). For the interpolation of longer gaps a more sophisticated algorithm would have to be used, but the gain would be much less than that of the filling of small gaps done here.

A.3 The Euler-Maruyama Approximation

The Euler-Maruyama approximation is a simple time discrete approximation of an Itō process. It is the first and simplest strong Taylor approximation [121].

Though it is possible to use variable time steps in the Euler-Maruyama approximation, we only consider a fixed time step Δ so that the discretisation of the

time interval $[t_0, t_N]$ is

$$\{t_n\} = \{t_0 + n\Delta t : n = 1 \dots N\}. \tag{A.1}$$

For an Itō process defined by a Langevin equation:

$$d\mathbf{X}(t) = \mathbf{f}(\mathbf{X}(t), t)dt + \mathbf{k}(\mathbf{X}(t), t)d\mathbf{W}(t) \tag{A.2}$$

the scheme of the approximation in the one dimensional case for an initial value $X_{t_0} = X_0$ is:

$$X(t_{n+1}) = X(t_n) + \Delta t f(X(t_n), t_n) + k(X(t_n), t_n)(W(t_{n+1}) - W(t_n)) \tag{A.3}$$

where W is the Wiener process. This means that for each time step we have to compute the increment $W(t_{n+1}) - W(t_n)$ which is Gaussian white noise with a standard deviation of $\sqrt{\Delta t}$.[1]

In the D-dimensional case each component i of the scheme has the form:

$$X_i(t_{n+1}) = X_i(t_n) + \Delta t f_i(\mathbf{X}(t_n), t_n) + \sum_{j=1}^{D} k_{i,j}(\mathbf{X}(t_n), t_n)(W_j(t_{n+1}) - W_j(t_n)) \tag{A.4}$$

where each component W_j of the vector $\mathbf{W}$ is an independent Wiener process. In vector form:

$$\mathbf{X}(t_{n+1}) = \mathbf{X}(t_n) + \Delta t \mathbf{f}(\mathbf{X}(t_n), t_n) + \mathbf{k}(\mathbf{X}(t_n), t_n)(\mathbf{W}(t_{n+1}) - \mathbf{W}(t_n)). \tag{A.5}$$

In the special case of $\mathbf{k}(\mathbf{X}, t) \equiv 0$ the Euler-Maruyama approximation reduces to the Euler scheme for deterministic differential equations.

The order β of weak convergence[2] for the Euler-Maruyama approximation is $\beta = 1$ given some conditions on $\mathbf{f}$ and $\mathbf{k}$.[3]

[1] See section 1.3.1.

[2] For time discrete approximations $X_{\Delta t}$ *weak convergence* to a process Y means that ensemble averages of nice enough functionals (e.g. moments) of the process converge at each time t: $\lim_{\Delta t \to 0} \langle g(Y(t)) \rangle - \langle g(X_{\Delta t}(t)) \rangle = 0$. See [121, p. 327].

[3] For details on the conditions see [121, p. 457ff].

A.4 Index of Common Variable Names

$\mathcal{N}(\mu, \sigma)$	normal distribution with mean μ and variance σ^2
$\mathbf{W}(t)$	d-dimensional Wiener process
τ	time delay
Δt	discrete time step, temporal resolution of data
$\mathbf{D^{(1)}}(\mathbf{X})$	drift coefficient vector of a Fokker-Planck equation
$\mathbf{D^{(2)}}(\mathbf{X})$	diffusion coefficient tensor of a Fokker-Planck equation
$\mathbf{f}(\mathbf{X}), \mathbf{k}(\mathbf{X})$	deterministic and stochastic terms of a Langevin equation
$\mathrm{acf}(\tau)$	autocorrelation function
x, y, z	bumblebee position in foraging arena: distance to flower wall, position horizontal parallel to flower wall, height
β	horizontal turning angle
s	horizontal bumblebee speed
$g(s), h(\beta)$	deterministic drift of speed and turning angle
$\xi_s(t)$	speed-dependent noise term of turning angle β
$\psi(t)$	non-deterministic (noise) term in changes of speed s

Bibliography

[1] T. C. Ings and L. Chittka. Speed-accuracy tradeoffs and false alarms in bee responses to cryptic predators. *Current Biology*, 18(19):1520–1524, Oct 2008.

[2] F. Lenz, T. C. Ings, L. Chittka, A. V. Chechkin, and R. Klages. Spatiotemporal dynamics of bumblebees foraging under predation risk. *Phys. Rev. Lett.*, 108:098103, Mar 2012.

[3] F. Lenz, A. V. Chechkin, and R. Klages. Constructing a stochastic model of bumblebee flights from experimental data. *PLoS ONE*, 8(3):e59036, 03 2013.

[4] A. V. Chechkin, F. Lenz, and R. Klages. Normal and anomalous fluctuation relations for gaussian stochastic dynamics. *J. Stat. Mech.*, 2012(11):L11001, 2012.

[5] G. M. Viswanathan, Sergey V. Buldyrev, Shlomo Havlin, M. G. E. da Luz, E. P. Raposo, and H. Eugene Stanley. Optimizing the success of random searches. *Nature*, 401(6756):911–914, Oct 1999.

[6] G. M. Viswanathan, V. Afanasyev, S. V. Buldyrev, E. J. Murphy, P. A. Prince, and H. E. Stanley. Lévy flight search patterns of wandering albatrosses. *Nature*, 381(6581):413–415, May 1996.

[7] S. Benhamou. How many animals really do the Lévy walk? *Ecology*, 88(8):1962–1969, AUG 2007.

[8] A. M. Edwards, R. A. Phillips, N. W. Watkins, M. P. Freeman, E. J. Murphy, V. Afanasyev, S. V. Buldyrev, M. G. E. da Luz, E. P. Raposo, H. E. Stanley, and G. M. Viswanathan. Revisiting Lévy flight search patterns of wandering albatrosses, bumblebees and deer. *Nature*, 449(7165):1044–1048, Oct 2007.

[9] A. M. Reynolds and M. A. Frye. Free-flight odor tracking in drosophila is consistent with an optimal intermittent scale-free search. *PLoS One*, 4:e354, 2007.

[10] D. W. Sims, E. J. Southall, N. E. Humphries, G. C. Hays, C. J. A. Bradshaw, J. W. Pitchford, A. James, M. Z. Ahmed, A. S. Brierley, M. A. Hindell, D. Morritt, M. K. Musyl, D. Righton, E. L. C. Shepard, V. J. Wearmouth, R. P. Wilson, M. J. Witt, and J. D. Metcalfe. Scaling laws of marine predator search behaviour. *Nature*, 451(7182):1098–1102, Feb 2008.

[11] D. W. Stephens, J. S. Brown, and R. C. Ydenberg. *Foraging: Behavior and Ecology*. University of Chicago Press, 2007.

[12] L. Giuggioli and F. Bartumeus. Linking animal movement to site fidelity. *J. Math. Biol.*, 64:647–656, 2012. 10.1007/s00285-011-0431-7.

[13] E. P. Raposo, F. Bartumeus, M. G. E. da Luz, P. J. Ribeiro-Neto, T. A. Souza, and G. M. Viswanathan. How Landscape Heterogeneity Frames Optimal Diffusivity in Searching Processes. *PLoS Comput. Biol.*, 7(11), NOV 2011.

[14] F. Bartumeus, J. Catalan, G. M. Viswanathan, E. P. Raposo, and M. G. E. da Luz. The influence of turning angles on the success of non-oriented animal searches. *J. Theor. Biol.*, 252(1):43–55, MAY 7 2008.

[15] P. A. Zollner and S. L. Lima. Behavioral tradeoffs when dispersing across a patchy landscape. *OIKOS*, 108(2):219–230, FEB 2005.

[16] W. M. Getz and D. Saltz. A framework for generating and analyzing movement paths on ecological landscapes. *Proc. Natl. Acad. Sci. USA*, 105(49):19066–19071, DEC 9 2008.

[17] O. Miramontes, D. Boyer, and F. Bartumeus. The effects of spatially heterogeneous prey distributions on detection patterns in foraging seabirds. *PLoS One*, 7(4):e34317, 04 2012.

[18] E.T. Kai, V. Rossib, J. Sudreb, H. Weimerskirch, C. Lopez, E. Hernandez-Garcia, F. Marsac, and V. Garcon. Top marine predators track Lagrangian coherent structures. *Proc. Natl. Acad. Sci. USA*, 106:8245–8250, 2009.

[19] F. Bartumeus, F. Peters, S. Pueyo, C. Marrase, and J. Catalan. Helical Lévy walks: Adjusting searching statistics to resource availability in microzooplankton. *Proc. Natl. Acad. Sci. USA*, 100:12771–12775, 2003.

[20] C. Beck and E. G. D. Cohen. Superstatistics. *Physica A*, 322(0):267 – 275, 2003.

[21] J. R. Krebs and N. B. Davies. *Behavioral Ecology An Evolutionary Approach*. Blackwell, Oxford, 1978.

[22] J. R. Krebs. Optimal foraging, predation risk and territory defense. *ARDEA*, 68(1-4):83–90, 1980.

[23] P. Romanczuk, M. Baer, W. Ebeling, B. Lindner, and L. Schimansky-Geier. Active Brownian particles from individual to collective stochastic dynamics. *Eur. Phys. J.-Spec. Top.*, 202(1):1–162, MAR 2012.

[24] S. Petrovskii and A. Morozov. Dispersal in a statistically structured population: Fat tails revisited. *American Naturalist*, 173(2):278–289, 2009.

[25] S. Hapca, J. W. Crawford, and I. M. Young. Anomalous diffusion of heterogeneous populations characterized by normal diffusion at the individual level. *J. R. Soc. Interface*, 6(30):111–122, 2009.

[26] C. Hawkes. Linking movement behaviour, dispersal and population processes: is individual variation a key? *J. Anim. Ecol.*, 78(5):894–906, 2009.

[27] E. Jongejans, O. Skarpaas, and K. Shea. Dispersal, demography and spatial population models for conservation and control management. *Perspectives in Plant Ecology, Evolution and Systematics*, 9(3-4):153–170, 2008.

[28] R. Kawai and S. Petrovskii. Multi-scale properties of random walk models of animal movement: lessons from statistical inference. *Proc. R. Soc. A*, 468(2141):1428–1451, MAY 8 2012.

[29] G. A. Nevitt, M. Losekoot, and H. Weimerskirch. Evidence for olfactory search in wandering albatross, Diomedea exulans. *Proc. Natl. Acad. Sci. USA*, 105(12):4576–4581, MAR 25 2008.

[30] J. Spaethe, J. Tautz, and L. Chittka. Visual constraints in foraging bumblebees: Flower size and color affect search time and flight behavior. *Proc. Natl. Acad. Sci. USA*, 98(7):3898–3903, MAR 27 2001.

[31] M. Lihoreau, L. Chittka, and N. E. Raine. Travel Optimization by Foraging Bumblebees through Readjustments of Traplines after Discovery of New Feeding Locations. *American Naturalist*, 176(6):744–757, DEC 2010.

[32] G. M. Viswanathan, E. P. Raposo, and M. G. E. da Luz. Lévy flights and superdiffusion in the context of biological encounters and random searches. *Physics of Life Reviews*, 5(3):133–150, SEP 2008.

[33] E. P. Raposo, S. V. Buldyrev, M. G. E. da Luz, G. M. Viswanathan, and H. E. Stanley. Lévy flights and random searches. *J. Phys. A: Math. Theor.*, 42(43):434003, 2009.

[34] J. G. Burns and J. D. Thomson. A test of spatial memory and movement patterns of bumblebees at multiple spatial and temporal scales. *Behav. Ecol.*, 17:48–55, 2005.

[35] M. Lihoreau, N. E. Raine, A. M. Reynolds, R. J. Stelzer, K. S. Lim, A. D. Smith, J. L. Osborne, and L. Chittka. Radar Tracking and Motion-Sensitive Cameras on Flowers Reveal the Development of Pollinator Multi-Destination Routes over Large Spatial Scales. *PLoS Biol.*, 10(9), SEP 2012.

[36] K. Ohashi, J. D. Thomson, and D. D'Souza. Trapline foraging by bumble bees: IV. Optimization of route geometry in the absence of competition. *Behavioral Ecology*, 18(1):1–11, JAN-FEB 2007.

[37] M. Lihoreau, N. E. Raine, A. M. Reynolds, R. J. Stelzer, K. S. Lim, A. D. Smith, J. L. Osborne, and L. Chittka. Unravelling the mechanisms of trapline foraging in bees. *Communicative and Integrative Biology: in press*, 2013.

[38] H. Risken. *The Fokker-Planck Equation : Methods of Solution and Applications*. Springer, Berlin New York, 1989.

[39] P. Lévy. *Théorie de l'addition des variables aléatoires*. Guathier Villars, Paris, 1937.

[40] B. V. Gnedenko and A. N. Kolmogorov, editors. *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, London, 1954.

[41] R. C. H. Cheng and W. B. Liu. A continuous representation of the family of stable law distributions. *J. Roy. Stat. Soc. B*, 59(1):137–145, 1997.

[42] E. A. Codling and N. A. Hill. Sampling rate effects on measurements of correlated and biased random walks. *J. Theor. Biol.*, 233(4):573–588, APR 21 2005.

[43] S. Benhamou. How to reliably estimate the tortuosity of an animal's path: straightness, sinuosity, or fractal dimension? *J. Theor. Biol.*, 229(2):209–220, JUL 21 2004.

[44] P. M. Kareiva and N. Shigesada. Analyzing Insect Movement as a Correlated Random-Walk. *Oecologia*, 56(2-3):234–238, 1983.

[45] V. M. Kenkre, E. W. Montroll, and M. F. Shlesinger. Generalized master equations for continuous-time random walks. *J. Stat. Phys.*, 9(45), 1973.

[46] C. Gardiner. *Handbook of Stochastic Methods*. Springer-Verlag, Berlin, Germany, 1985.

[47] N. Van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland, Amsterdam, The Netherlands, 2001.

[48] O. Bénichou, C. Loverdo, M. Moreau, and R. Voituriez. Intermittent search strategies. *Rev. Mod. Phys.*, 83(1):81, Mar 2011.

[49] M. Raghib, S. A. Levin, and I. G. Kevrekidis. Multiscale analysis of collective motion and decision-making in swarms: An advection-diffusion equation with memory approach. *J. Theor Biol.*, 264(3):893–913, JUN 7 2010.

[50] R. Klages, G. Radons, and I.M. Sokolov, editors. *Anomalous transport*. Wiley-VCH, Weinheim, July 2008.

[51] P. Bovet and S. Benhamou. Spatial-Analysis of Animals Movements Using a Correlated Random Walk Model. *J. Theor. Biol.*, 131(4):419–433, APR 21 1988.

[52] S. Benhamou. Detecting an orientation component in animal paths when the preferred direction is individual-dependent. *Ecology*, 87(2):518–528, FEB 2006.

[53] A. M. Reynolds. Bridging the gulf between correlated random walks and Lévy walks: autocorrelation as a source of Lévy walk movement patterns. *J. R. Soc. Interface*, 7(53):1753–1758, DEC 6 2010.

[54] D. S. Johnson, J. M. London, M. Lea, and J. W. Durban. Continuous-time correlated random walk model for animal telemetry data. *Ecology*, 89(5):1208–1215, MAY 2008.

[55] G. Wittemyer, L. Polansky, I. Douglas-Hamilton, and W. M. Getz. Disentangling the effects of forage, social rank, and risk on movement autocorrelation of elephants using Fourier and wavelet analyses. *Proc. Natl. Acad. Sci. USA*, 105(49):19108–19113, DEC 9 2008.

[56] G. M. Viswanathan, E. P. Raposo, F. Bartumeus, Jordi Catalan, and M. G. E. da Luz. Necessary criterion for distinguishing true superdiffusion from correlated random walk processes. *Phys. Rev. E*, 72(1):011111, Jul 2005.

[57] B. Lindner. Diffusion of particles subject to nonlinear friction and a colored noise. *New J. Phys.*, 12, JUN 15 2010.

[58] P. Romanczuk and L. Schimansky-Geier. Brownian motion with active fluctuations. *Phys. Rev. Lett.*, 106:230601, Jun 2011.

[59] T. Hart, R. Mann, T. Coulson, N. Pettorelli, and P. Trathan. Behavioural switching in a central place forager: patterns of diving behaviour in the macaroni penguin (Eudyptes chrysolophus). *Marine Biology*, 157(7):1543–1553, JUL 2010.

[60] M. A. Lomholt, T. Koren, R. Metzler, and J. Klafter. Lévy strategies in intermittent search processes are advantageous. *Proc. Natl. Acad. Sci. USA*, 105:11055–11059, Aug 2008.

[61] A. M. Reynolds. On the intermittent behaviour of foraging animals. *Europhys. Lett.*, 75(4):517–520, AUG 2006.

[62] Alla Mashanova, Tom H. Oliver, and Vincent A. A. Jansen. Evidence for intermittency and a truncated power law from highly resolved aphid movement data. *J. R. Soc. Interface*, 7(42):199–208, JAN 6 2010.

[63] Alka A. Potdar, Junhwan Jeon, Alissa M. Weaver, Vito Quaranta, and Peter T. Cummings. Human Mammary Epithelial Cells Exhibit a Bimodal Correlated Random Walk Pattern. *PLoS One*, 5(3), MAR 10 2010.

[64] Surya G. Nurzaman, Yoshio Matsumoto, Yutaka Nakamura, Kazumichi Shirai, Satoshi Koizumi, and Hiroshi Ishiguro. From Lévy to Brownian: A Computational Model Based on Biological Fluctuation. *PLoS One*, 6:e16168, 2011.

[65] J. M. Morales, D. T. Haydon, J. Frair, K. E. Holsiner, and J. M. Fryxell. Extracting more out of relocation data: Building movement models as mixtures of random walks. *Ecology*, 85(9):2436–2445, SEP 2004.

[66] O. Benichou, M. Coppey, M. Moreau, P. H. Suet, and R. Voituriez. Optimal search strategies for hidden targets. *Phys. Rev. Lett.*, 94(19), MAY 20 2005.

[67] E. A Codling, M. J Plank, and S. Benhamou. Random walk models in biology. *J. R. Soc. Interface*, 5(25):813–834, 2008.

[68] N. E. Humphries, N. Queiroz, J. R. M. Dyer, N. G. Pade, M. K. Musyl, K. M. Schaefer, D. W. Fuller, J. M. Brunnschweiler, T. K. Doyle, J. D. R. Houghton, G. C. Hays, C. S. Jones, L. R. Noble, V. J. Wearmouth, E. J. Southall, and D. W. Sims. Environmental context explains Lévy and Brownian movement patterns of marine predators. *Nature*, 465:1066–1069, 2010.

[69] M. J. Plank and E. A. Codling. Sampling rate and misidentification of Lévy and non-Lévy movement paths. *Ecology*, 90(12):3546–3553, DEC 2009.

[70] C. Darwin. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray, 1859.

[71] G. H. Pyke, H. R. Pulliam, and E. L. Charnov. Optimal foraging: A selective review of theory and tests. *Q. Rev. Biol.*, 52(2):137–154, 1977.

[72] G. H. Pyke. Optimal Travel Speeds of Animals. *American Naturalist*, 118(4):475–487, 1981.

[73] A. James, J. W. Pitchford, and M. J. Plank. Efficient or inaccurate? Analytical and numerical modelling of random search strategies. *Bull. Math. Biol.*, 72:896–913, 2009.

[74] G. M. Viswanathan, M. G. E. da Luz, E. P. Raposo, and H. E. Stanley. *The Physics of Foraging*. Cambridge University Press, Cambridge, 2011.

[75] M. G. E. da Luz, A. Grosberg, E. P. Raposo, and G. M. Viswanathan. The random search problem: trends and perspectives. *J. Phys. A: Math. Theor.*, 42(43):430301, OCT 2009.

[76] G. Ramos-Fernández, J. Mateos, O. Miramontes, G. Cocho, H. Larralde, and B. Ayala-Orozco. Lévy walk patterns in the foraging movements of spider monkeys (ateles geoffroyi). *Behav. Ecol. Sociobiol.*, 55:223–230, 2004.

[77] Michael F. Shlesinger and J. Klafter. Lévy walk vs. Lévy flights. *On Growth and Form*, page 279, 1985.

[78] R. S. Schick, S. R. Loarie, F. Colchero, B. D. Best, A. Boustany, D. A. Conde, P. N. Halpin, L. N. Joppa, C. M. McClellan, and J. S. Clark. Understanding movement data and movement processes: current and emerging directions. *Ecology Letters*, 11(12):1338–1350, DEC 2008.

[79] F. Bartumeus. Behavioral intermittence, Lévy patterns, and randomness in animal movement. *Oikos*, 118(4):488–494, APR 2009.

[80] P.E. Smouse, S. Focardi, P.R. Moorcroft, J.G. Kie, J. D. Forester, and J. M. Morales. Stochastic modelling of animal movement. *Phil. Trans. R. Soc. B*, pages 2201–2211, 2010.

[81] R. H. MacArthur and E.R. Pianka. On optimal use of a patchy environment. *American Naturalist*, 100:603–609, 1966.

[82] A. James, M. J. Plank, and A. M. Edwards. Assessing Lévy walks as models of animal foraging. *J. R. Soc. Interface*, 8(62):1233–1247, SEP 7 2011.

[83] F. Bartumeus, M. G. E. Da Luz, G. M. Viswanathan, and J. Catalan. Animal search strategies: A quantitative. random-walk analysis. *Ecology*, 86(11):3078–3087, NOV 2005.

[84] A. M. Reynolds, A. D. Smith, R. Menzel, U. Greggers, D. R. Reynolds, and J. R. Riley. Displaced honey bees perform optimal scale-free search flights. *Ecology*, 88(8):1955–1961, AUG 2007.

[85] David W. Sims, Nicolas E. Humphries, Russell W. Bradford, and Barry D. Bruce. Lévy flight and Brownian search patterns of a free-ranging predator reflect different prey field characteristics. *J. Anim. Ecol.*, 81(2):432–442, 2012.

[86] D. W. Sims, D. Righton, and J. W. Pitchford. Minimizing errors in identifying Lévy flight behaviour of organisms. *J. Anim. Ecol.*, 76(2):222–229, 2007.

[87] A. M. Edwards. Using likelihood to test for Lévy flight search patterns and for general power-law distributions in nature. *J. Anim. Ecol.*, 77(6):1212–1222, 2008.

[88] P. Dieterich, R. Klages, R. Preuss, and A. Schwab. Anomalous dynamics of cell migration. *Proc. Natl. Acad. Sci. USA*, 105:459–463, 2008.

[89] T. Geisel, J. Nierwetberg, and A. Zacherl. Accelerated diffusion in Josephson junctions and related chaotic systems. *Phys. Rev. Lett.*, 54:616–619, 1985.

[90] G. Zumofen and J. Klafter. Power spectra and random walks in intermittent chaotic systems. *Physica D*, 69:436–446, 1993.

[91] E. A. Codling and M. J. Plank. Turn designation, sampling rate and the misidentification of power laws in movement path data using maximum likelihood estimates. *Theoretical Ecology*, 4(3):397–406, AUG 2011.

[92] Thomas C. Ings and Lars Chittka. Predator crypsis enhances behaviourally mediated indirect effects on plants by altering bumblebee foraging preferences. *Proc. R. Soc. B-Biol. Sci.*, 276(1664):2031–2036, JUN 7 2009.

[93] T. Ings, M.-Y. Wang, and L. Chittka. Colour-independent shape recognition of cryptic predators by bumblebees. *Behavioral Ecology and Sociobiology*, pages 1–10, 2011. 10.1007/s00265-011-1295-y.

[94] H. Akaike. A new look at the statistical model identification. *IEEE Trans. Autom. Control*, 19(6):716–723, dec 1974.

[95] G. Schwarz. Estimating the dimension of a model. *Ann. Statist.*, 6(2):461–464, 1978.

[96] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, 1965.

[97] T. Yokoi and K. Fujisaki. Hesitation behaviour of hoverflies sphaerophoria spp. to avoid ambush by crab spiders. *Naturwissenschaften*, 96(2):195–200, FEB 2009.

[98] R. Friedrich and J. Peinke. Description of a turbulent cascade by a Fokker-Planck equation. *Phys. Rev. Lett.*, 78:863–866, Feb 1997.

[99] R. Friedrich, J. Peinke, and C. Renner. How to quantify deterministic and random influences on the statistics of the foreign exchange market. *Phys. Rev. Lett.*, 84(22):5224–5227, MAY 29 2000.

[100] M. Ragwitz and H. Kantz. Indispensable finite time corrections for Fokker-Planck equations from time series data. *Phys. Rev. Lett.*, 87(25), DEC 17 2001.

[101] S Kriso, J Peinke, R Friedrich, and P Wagner. Reconstruction of dynamical equations for traffic flow. *Phys. Lett. A*, 299(2-3):287–291, JUL 1 2002.

[102] F. Lenz, D. Herde, A. Riegert, and H. Kantz. Bivariate time-periodic Fokker-Planck model for freeway traffic. *Eur. Phys. J. B*, 72(3):467–472, DEC 2009.

[103] E. Batschelet. *Circular Statistics in Biology*. Academic Press, London, 1981.

[104] William Press, Saul Teukolsky, William Vetterling, and Brian Flannery. *Numerical Recipes in C*. Cambridge University Press, Cambridge, UK, 2nd edition, 1992.

[105] Donald W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*, 11(2):431–441, 1963.

[106] N. J. Kasdin. Discrete Simulation of Colored Noise and Stochastic-Processes and 1/F(Alpha) Power-Law Noise Generation. *Proc. of the IEEE*, 83(5):802–827, MAY 1995.

[107] S. Primak. Generation of compound non-gaussian random processes with a given correlation function. *Phys. Rev. E*, 61:100–103, Jan 2000.

[108] F. Peruani and L. G. Morelli. Self-propelled particles with fluctuating speed and direction of motion in two dimensions. *Phys Rev Lett*, 99(1), JUL 6 2007.

[109] R. J. Harris and G. M. Schütz. Fluctuation theorems for stochastic dynamics. *J. Stat. Mech.*, 2007(07):P07020, 2007.

[110] J. Kurchan. Non-equilibrium work relations. *J. Stat. Mech.*, 2007(07):P07005, 2007.

[111] D. J. Evans and D. J. Searles. The fluctuation theorem. *Advances in Physics*, 51(7):1529–1585, 2002.

[112] L. Rondoni and C. Mejía-Monasterio. Fluctuations in nonequilibrium statistical mechanics: models, mathematical theory, physical mechanisms. *Nonlinearity*, 20(10):R1, 2007.

[113] D. J. Evans, E. G. D. Cohen, and G. P. Morriss. Probability of second law violations in shearing steady states. *Phys. Rev. Lett.*, 71:2401–2404, Oct 1993.

[114] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78:2690–2693, Apr 1997.

[115] G. E. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E*, 60:2721–2726, Sep 1999.

[116] A. V. Chechkin and R. Klages. Fluctuation relations for anomalous dynamics. *J. Stat. Mech.-Theory and Experiment*, Mar 2009.

[117] S. M. Rytov, I. U. A. Kravtsov, and V. I. Tatarskii. *Principles of Statistical Radiophysics*, volume 1. Springer-Verlag, Berlin, 1989.

[118] L. Pochhammer. Über die Differentialgleichung der allgemeineren hypergeometrischen Reihe mit zwei endlichen singulären Punkten. *Journal für die reine und angewandte Mathematik*, 1888(102):76 – 159, 1826.

[119] G. E. P. Box and G. M. Jenkins. *Time series analysis: forecasting and control*. Holden-Day series in time series analysis and digital processing. Holden-Day, 1976.

[120] O. D. Anderson. *Time series analysis and forecasting: the Box-Jenkins approach*. Time series and forecasting. Butterworths, 1977.

[121] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*. Springer, Berlin, 1999.