Analyzing the House Fly's Exploratory Behavior with Autoregression Methods

Hisanao TAKAHASHI[1]*, Naoto HORIBE[2],
Takashi IKEGAMAI[2], Masakazu SHIMADA[2]

[1]*interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology. Yokohama 226-8502*
[2]*Graduate School of Arts and Science, University of Tokyo. Tokyo 153-8902*

This paper presents a detailed characterization of the trajectory of a single housefly with free range of a square cage. The trajectory of the fly was recorded and transformed into a time series, which was fully analyzed using the autoregression model. A main discovery was that the fly switched styles of motion from a low dimensional regular pattern to a higher dimensional disordered pattern.

The exploratory behavior analysis is important from both the psychological and dynamic system's point of view, and it is characterized by the anomalous diffusion.

KEYWORDS: housefly, AR model, anomalous diffusion, exploration behavior, memory and learning

1. Introduction

Biological autonomy is one way to characterize life forms. It is the autonomous dynamics of a living system that show different behaviors in the same context, or the same behaviors in a different context. Quantitatively, a creature's spontaneous movement is worth studying as a primary index for biological autonomy. Creatures much simpler than the vertebrates, such as flies or even unicellular animals such as bacteria can show an internally generated behavior, which the most complex man-made robots fail to show. We thus designate the common fly as a test animal to investigate its spontaneous motion. An advantage of studying a fly's behavior is that a fly responds to the information of its environment by changing navigational patterns. Their responses are not simply reactive, for they behave differently even in the same context. If

*Previously, Graduate School of Arts and Science, University of Tokyo.

a fly's motion is simply random or always driven by the environmental context, we do not call it autonomous. We define and study autonomous motion as an interplay between internal and external dynamics, which was recently proposed as an embodied chaotic itinerancy (ECI).[3] Of course, we cannot directly see this interplay, but here we assume that the interplay of dynamics can be quantified by the navigational trails of a fly. In particular, switching navigational modes during foraging is what ECI considers to be a candidate of autonomous motion. Therefore, the characterization of the switching behavior is the main purpose of this paper.

Until recently, such spontaneous movements as a fly's exploratory movements have not been studied quantitatively, but one of the first serious observations was provided by Murdie and Hassell[8] in 1973. They analyzed the parent changes of the fly's walking between pre- and post-feeding phases by decomposing the fly's temporal motion into the angle turned and the distance (e.g. forwarding) components. Before feeding, there is considerable variation in both the angle turned and the distance movement. After feeding, there is a sharp increase in the mean angle turned and a sharp decrease in the distance moved. They also demonstrate that a fly's biased motion increases the probability of discovery of food sites.

We use an autoregression model (AR model)[2,10] for the first time to analyze the trajectory data. The AR model shows us different aspects of the time series patterns from the previous studies,[6,8] where their method is often called descriptive statistics. Although the AR model is one of the simplest models that examines time structure, it is effective in classifying the dynamics, as we show in this paper. In our analysis of a local stationary AR model, we found that the distance movement took a stationary distribution, which is rather inconsistent with Murdle and Hassell's results. We assume that the difference comes from the definition of "stationary" in our experiment. Our definition is appropriate if we regard a fly as an autonomous agent, so that when a fly changes its behavioral pattern, it must be correlated with the intrinsic value system. Also, by analyzing the difference between the fly's navigational and random walk, which has been discussed in terms of the foraging efficiency,[1,4,9,11,14] we will also characterize the fly's motion as an anomalous diffusion process.

2. Method

2.1 Experiment

We used a common housefly and a personal computer to record the trajectory of its walking in an acryl cage by digital video camera. Figure 1 shows the whole picture of our system. To begin, a fly is put in the acryl cage whose size is 47 cm square and 2.5 cm deep. From the cage above, the digital video camera takes a picture of the whole layout. Then, the place of the fly in the picture is transformed into a two-dimensional time series (on x, y coordinate) using a personal computer.

Small droplets of sugar solution (4 % source, 2 ml) were distributed on the floor of the cage. We also performed some experiments under the no-sugar solution condition for comparison. Figure 2 shows a trajectory of a fly for about 7 minutes.

2.2 Local stationary AR model

We use a local stationary AutoRegressive (AR) model to analyze the time series, which is computed by the following procedures. Let us denote the value of time series at time $t,\ t-1,\ t-2, ...$ by $x_t,\ x_{t-1},\ x_{t-2}....$ Also let $z_t,\ z_{t-1},\ z_{t-2}, ...$ be the deviations from the mean value of the time series μ, i.e. $z_t = x_t - \mu$. Then, the mth order AR model is defined as follows.

$$z_t = a_1\, z_{t-1} + a_2\, z_{t-2} + \cdots + a_m\, z_{t-m} + w_t \tag{1}$$

where w_t is the Gaussian white noise whose variance is σ^2. A standard Akaike Information Criteria (AIC) is used to decide the effective AR coefficients $a_1,\ a_2, ...,\ a_m$, and the variance (σ^2) of this model.[10] The number of the AR order m is called the AR dimension.

Next, a local stationary AR model is constructed by the recursive procedures as follows.

Step 1) Divide the time series into the well-defined small intervals which have the same segmental length L, and suppose the AR model is stationary in each interval.

Step 2) Using the AIC, we decide the AR coefficient and the variance in the first and second intervals and name these AIC values as AIC_1 and AIC_2, respectively. (See Fig. 3.)

Step 3) Third, we make the united interval from the two successive intervals, taking the starting point of the first interval as the starting point of the united one and the ending point of the second interval as the ending point of the united one. We decide the AR coefficient and the variance for this interval in the same way as before, and name this AIC value as AIC_{12}. If the inequality $AIC_1 + AIC_2 < AIC_{12}$ holds, then we assume the two intervals are driven by the different AR model and, otherwise, the two intervals are driven by the same AR model. In the former case, we think two intervals are separated segments, which means the walking pattern was changed between the two intervals. And, in the latter case, the two intervals are treated as one segment, which means the walking pattern was the same in the two intervals.

Step 4) For the next step, we rename the intervals as follows. When the inequality $AIC_1 + AIC_2 < AIC_{12}$ holds, we regard the second interval as a first interval and the third interval as a second interval, and repeat the same procedure as before. When the inequality is invalid, we regard the united interval as a first interval, and the third interval as a second interval and repeat the same steps. We repeat the same procedure until all the data set is visited.

If we find a united interval, we conclude the process is stationary in the interval. We only use the smallest size of a segment L as 100 in this paper. Practically, the number of free parameters estimated by AIC should be less than $2\sqrt{L}$.[10] Consequently, $L = 100$ is a sufficient length in this study.

2.3 Spectrum

We should be careful about applying the AR model. Even if the time series is separated by the different local AR models, we should not presume that the strategy of the fly's walk is changed at that separated point. The walk may be caused by the nonstationary or nonlinear effects of the fly's walk itself. So, we continue to use an AR spectrum that inversely generates the time series to compute the power spectrum of the reconstructed time series.

By generating the AR model such as,

$$x_n = \sum_{i=1}^{j} a_i\, x_{n-i} + w_n,$$

we then compute its power spectrum as follows.

$$p(f) = \sum_{k=-\infty}^{\infty} E(x_n x_{n-k}) e^{2\pi i k f} = \frac{\sigma^2}{\left| 1 - \sum_{j=1}^{m} a_j e^{2\pi i j f} \right|^2}.$$

Where E[xy] is the temporal average of two variables x and y.

2.4 Analysis

We decompose the motion behavior into velocity v_i and angular element θ_i components. Before applying the local stationary AR model, we define a velocity and angular difference such as $z_i = x_i - x_{i-1}$ for the target data set. Figure 4 shows the time series of the velocity's difference (a) and the angular difference (b), which we will analyze practically in this paper.

2.5 Deviation from the Brownian Motion

In order to characterize the difference between a fly's foraging motion and the Brownian motion (which is assumed as a random process), we study the two point correlations of the fly's navigational trajectory. A particular interest is the order in a fly's behavioral pattern which is quantified as an anomalous diffusion. Anomalous diffusion is observed in many kinds of exploratory behavior of organisms.[9,14] Here, we examine the diffusive speed of the fly's walk as

$$< (x(t+t_0) - x(t_0))^2 >_{t_0} \sim t^{\alpha}. \tag{2}$$

If the α is greater than one, it implies an anomalous diffusion. $\alpha = 1$ implies the Gaussian random walk.

The Levy flight might be the simplest explanation for this behavior. However, the appearance of the trajectory is different from the simulated Levy flight. The fly's exploratory pattern is relatively smooth, whereas the Lévy walks studied in [9, 14] are a combination of straight lines. We will show that the fly's walk can be decomposed into several walking patterns, which is classified by the AR model in this paper. Since it is possible to make an anomalous diffusion form the combination of other walking styles,[11] anomalous diffusion is what we expect here.

3. Results

A trajectory in Fig. 2 is recorded for about 7 minutes and the minimum time span of the data is set at 0.2 seconds. The density of the line of the trajectory is higher at some region and lower at the other. There are some droplets of sugar solution around the coordinate (440, 380) where the lines are crowded.

A fly in the cage reaches a sugar solution by walk, then sucks the solution, and finally leaves.

After a fly leaves the solution (from 171 steps to 570 steps), it walks around the solution for about one minute, where the line becomes crowded (Fig. 5 (a)). Also, it seems that the fly's walking pattern becomes circular, almost as if it has a virtual center. We assume that this biased motion is caused by memory capability of a fly, and also that it should be consistent with the anomalous diffusion discussed later.

Figure 5 (b) shows the trajectory after the pattern in Fig. 5 (a). The fly's walking pattern changes its navigational style; that is, the density of the trajectory becomes lower, and the fly walks into the wider area. After searching near the previous food area for a while, the fly seems to start to search in the wider area.

The results of the local AR model are given in Table I. Concerning the velocity's difference, two intervals from 171 steps to 370 steps are united as one interval by the local AR model with AR orders 4, and from 871 steps to 1070 steps are also united with AR orders 5. The unification of intervals is different between the time series of velocity data and that of the angular data. For the angular component time series, three intervals are picked up from 371 steps to 670 steps, and two from 671 steps to 870 steps, as for the same AR model. Because the angular and velocity motion of the fly do not change synchronously in the cases of 271, 471, 571, 771, and 971 steps, we assume that a fly can sometimes independently control its angular and velocity motion.

In the intervals from 171 to 670, the AR order is higher in regard to the angular difference. The hight AR order represents that the fly is using longer memory from the past to walk that might be necessary to wind around a certain point. The fly's walk in the interval from 171 to

470 seems to wind around the sugar solution it found (see Fig.5 (a)). It is true that we have to concern efferct or interaction both from angular differnce and velocity's difference, but if concerning under the restriction such as the velocity as constant, then the winding walk is necessary to be highter AR order.

In the intervals from 671 to 1170, and from 1271 to 1370, the AR order becomes lower and the elements of the longer period are higher in its spectrum, as we will mention later, and this fact corresponds to the fact that the fly's walk is becoming smoother (see Fig. 5(b)). In the case of the interval from 371 to 670, though the AR orders for the angular differences are high, the trajectory of the interval is smooth to some extent, so we have to consider the effect not only of the angular difference, but also of the velocity's difference.

If the local AR model shows that the two time series are different from each other, this might be the result of the non-stationarity or non-linearity of the time series, as we mentioned before in section 2.3. Therefore, we should carefully estimate the point where the fly's walking style changes, using the AR spectrum. Figure 6 shows the spectrum of the velocity's difference, and we show the corresponding intervals at the top of each figure. It is hard to find out the same regularity or rule in these figures. We hypothesize that the figures of (371-470), (571-670), and (671-770) have similar spectrum patterns, which suggest the same walking style of motion in those regions.

From the same discussion, we regard the intervals of (871-1070) and (1171-1270) as belonging to the same style of motion.

Figure 7 shows the spectrum of the angular difference. From the same reason explained above, the intervals (671-870), (871-970), (1071- 1170) and (1271-1370) imply the same styles of motion. If a fly is walking, while changing its angular component periodically and slowing the velocity component, then it explores a compact area. This happened at 171-470 steps. Even if a fly is walking while changing its angular component periodically and maintaining a moderate velocity, then it explores the wider area. This happened at 471-670 steps. If a fly walks moderately and changes its angular component smoothly, then it can explore a much wider area.

Figure 8 shows two different time series taken from two other individuals. Figure 8 (a) is the trajectory for about 14 minutes with sugar droplets, i.e. the only differences from the previous one were the positions of the droplets. Figure 8 (b) for about 27 minutes with no sugar droplets. In the case of the figure (a), the fly sucked several times at around (340,250) and left from there. The result of with sugar solution is similar with the above result. The fly walks around food, so that the line is crowded and zigzagged. After the fly left from the feeding area, the line gets smoother. In case of Fig. 8 (b)with a no-sugar condition, the line becomes smoother and more spread-out as in the case of Fig. 5 (b).

Figure 9 shows log-log plots for the eq. (2) and it shows that the points well fit from the time interval 0.4 sec to 8 sec. We have the exponent $\alpha = 1.47$ by using these data points. This result shows that the fly's walk is indeed an anomalous diffusion.

In this experiment, a fly was caged in a box, so that the anomalous diffusion is held only in the limited range.[5,11] We were able to verify that the fly's walk is characterized by anomalous diffusion in the range from mille-meters to centimeters. This diffusion is at the upper limit of the range at which we can verify diffusion in this experiment. This limit can be improved if we use a bigger cage, but we have to take into account a fly's ability to fly.

4. Concluding remarks

We found that the velocity and angular elements synchronously change concerning the AR dimension, but sometimes they change asynchronously. Based on our observation, a fly has a control system of walking movement as follows. There exist two main choices of navigational patterns: one for changing the velocity, and the other for changing the angular element. There seems to exist several channels for each choice, and a fly chooses one channel for each choice, which we can observe from the data set. Those channels are not prepared as a fixed pattern, but can be dynamically varied as a model simulates.[3]

We assume the time series of the fly's walk to be the local stationary time series, but we should carefully examine the nonstationary or nonlinear aspect of the time series. To solve the problem of the nonstationarity or nonlinearity of the fly's walk, we should use a totally different analysis, which will be left for future studies. It is true that a fly does not undertake long walks, rather stopping frequently and sometimes stop-walking and flying. These characteristics are complicating factors in this experiment. It seems that a fly undertakes flight when it cannot find anything interesting after a certain period of time. A big discrepancy found at about 500 steps in Fig. 4 (a) is caused by this reason. After flying, the fly changes its walking style if we compare it with the previous style. (But, it is not detected clearly yet.)

We only used the local stationary AR model for the fixed intervals, that is $L = 100$. But, the time series of a fly's walk does not change regularly at those points. One solution is to use the variable interval method, which requires an appreciable amount of CPU time. Thus, in order to compromise, we used only the local stationary AR method with the fixed length.

It is interesting to see the anomalous diffusion process in the experiment. The simplest explanation may be to assume random walking with a memory effect. But, we also interpret this behavior as a product of inter-play between the internal dynamics and the externally driven dynamics, which is theoretically studied as "embodied chaotic itinerancy (ECI)".[3] ECI shows randomly traveling dynamics among different pseudo-attractors deterministically. The experiment we undertook here can be explained by ECI.

Asynchronous dynamics between angular and velocity components provide the other evidence of such interplay. Recent experimental studies show that flies certainly have internal dynamics (e.g.van Swinderen B. et al.;[13] A. Maye et al.[7]). In Maye, A. et al.[7] showed that there is a deterministic chaos dynamics hidden in the intrinsic dynamics. A fly possesses some form of memory system, and the genetic studies of learning and memory for the fly are now being progressed.[12]

Acknowledgement

The part of this work is supported by The 21st Century COE (Center of Excellence) program (Research Center for Integrated Science) of the Ministry of Education, Culture, Sports, Science, and Technology, Japan, and the ECAgent project, sponsored by the Future and Emerging Technologies program of the European Community (IST-1940). We used the time series analysis and control program package (TIMSAC) for R to analyze the data.

References

1) Bond, A. B. (1982) The Geometry of Foraging Pattern: Components of Thoroughness in Random Searching. Available on Web.
2) Box, G. E. P., Jenkins, G. M., Reinsel, G. C.: *Time Series Analysis, 3rd ed.* (Printice-Hall International, Inc., New Jersey, 1994).
3) Ikegami, T. "Simulating Active Perception and Mental Imagery with Embodied Chaotic Itinerancy", Journal of Consciousness Studies 14 (2007) No.7 pp.111-125.
4) Klafter, J., Sokolov, J. M. (2005) Anomalous Diffusion Spreads its Wings, *Physics World* August 1-4.
5) Mantegna, Rosario N. and Stanley, H. Eugene (1994) Stochastic process with ultraslow convergence to a Gaussian: The truncated Lévy flight. *Phy. Rev. Lett.*, **73**, 2946-2949. See also Shlesinger, M. F. (1995) Comment on "Stochastic process with ultraslow convergence to a Gaussian: The truncated Lévy flight". *Phy. Rev. Lett.*, **74**, 4959.
6) Martin, J-R (2004) A portrait of locomotor behaviour in Drosophila determined by a video-tracking paradigm, *Behavioural Processes* **67**, 207-219.
7) May, A., Hsieh, C., Sugihara, G., and Brembs, B. "Order in Spontaneous Motion", PlosOne (2007) Issue 5. e443.
8) Murdie, G., Hassell, M.P. (1973) Food distribution, searching success and predator prey models, In: *The Mathematical Theory of the Dynamics of Biological Populations*, (M.S.Bartlett & R.W. Hiorns, Eds.), New York: Academic Press, 87-101.
9) Romos-Fernández, G., Mateos, J. L., Miramontes, O., Cocho, G., Larralde, H., Ayala-Orozco, B. (2004) Lévy Walk Pattans in the Foraging Movements of Spider Monkeys (Ateles Goeoffroyi), *Behavioral Ecology and Sociobiology* **55**, 223-230.
10) Sakamoto, Y., Ishiguro, M. and Kitagawa, G.: *Akaike Information Criterion Statistics* (Kluwer Academic Publishers, Tokyo, 1986).
11) Takahashi, H. (2003) Ehrenfest Model with Large Jumps in Finance, *Physica D* **189**, 61-69.
12) Tamura, T., Chiang, A.S., Ito, N., Liu, H.P., Horiuchi, J., Tully, T., Saitoe, M., (2003) Aging Specifically Impairs amnesiac-Dependent Memory in Drosophila *Neuron* **40**, 1003-1011.
13) van Swinderen B., Nitz, D.A., Greenspan, R.J. (2004) Uncoupling of Brain Activity from Movement Defines Arousal States in Drosophila, *Current Biology* **14**, 81-87.
14) Viswanathan,G. M., Buldyrev, S. V., Havlin, S., de Luz, M. G. E., Raposo, E. P., Stanley, H. E. (1999) Optimizing the Success of Random Searches, *Nature* **401**, 911-914.

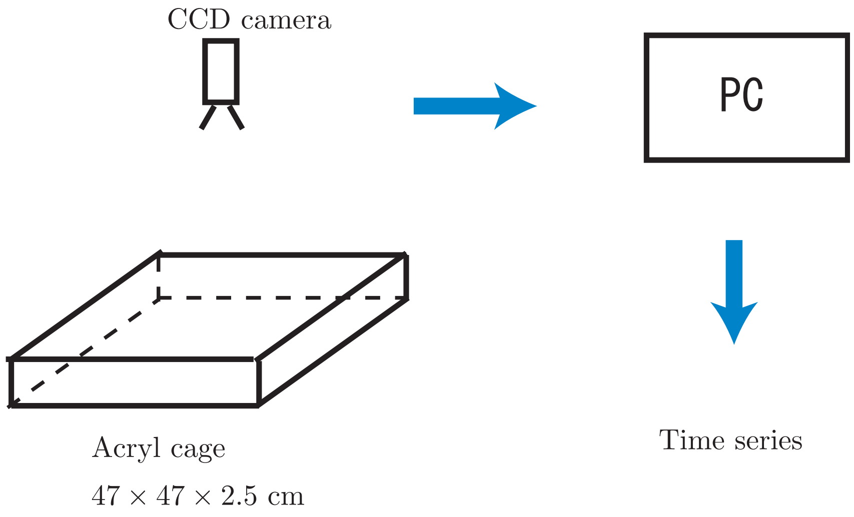

Fig. 1. Experimental system

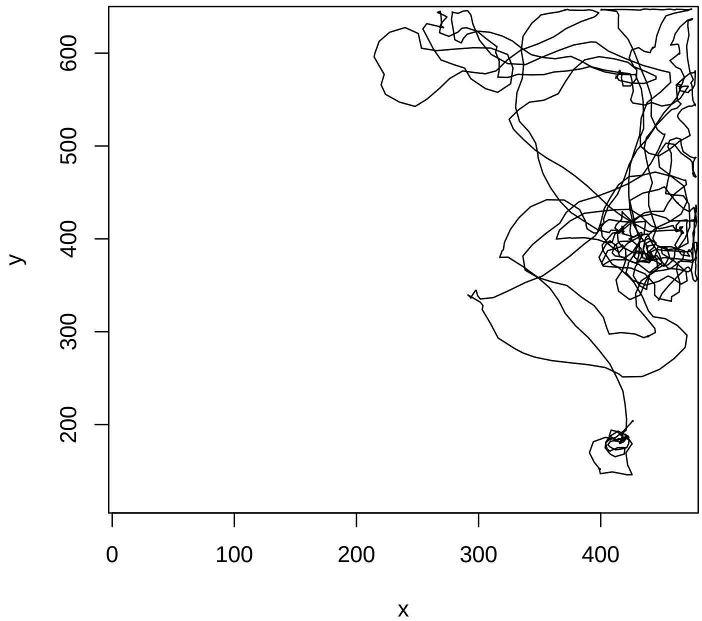

Fig. 2. Trajectory of a fly for about 7 minutes with sugar solution's droplets.

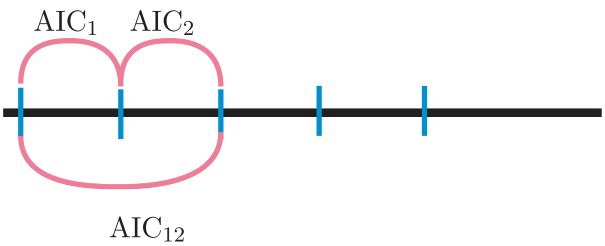

Fig. 3. Local stationary AR model

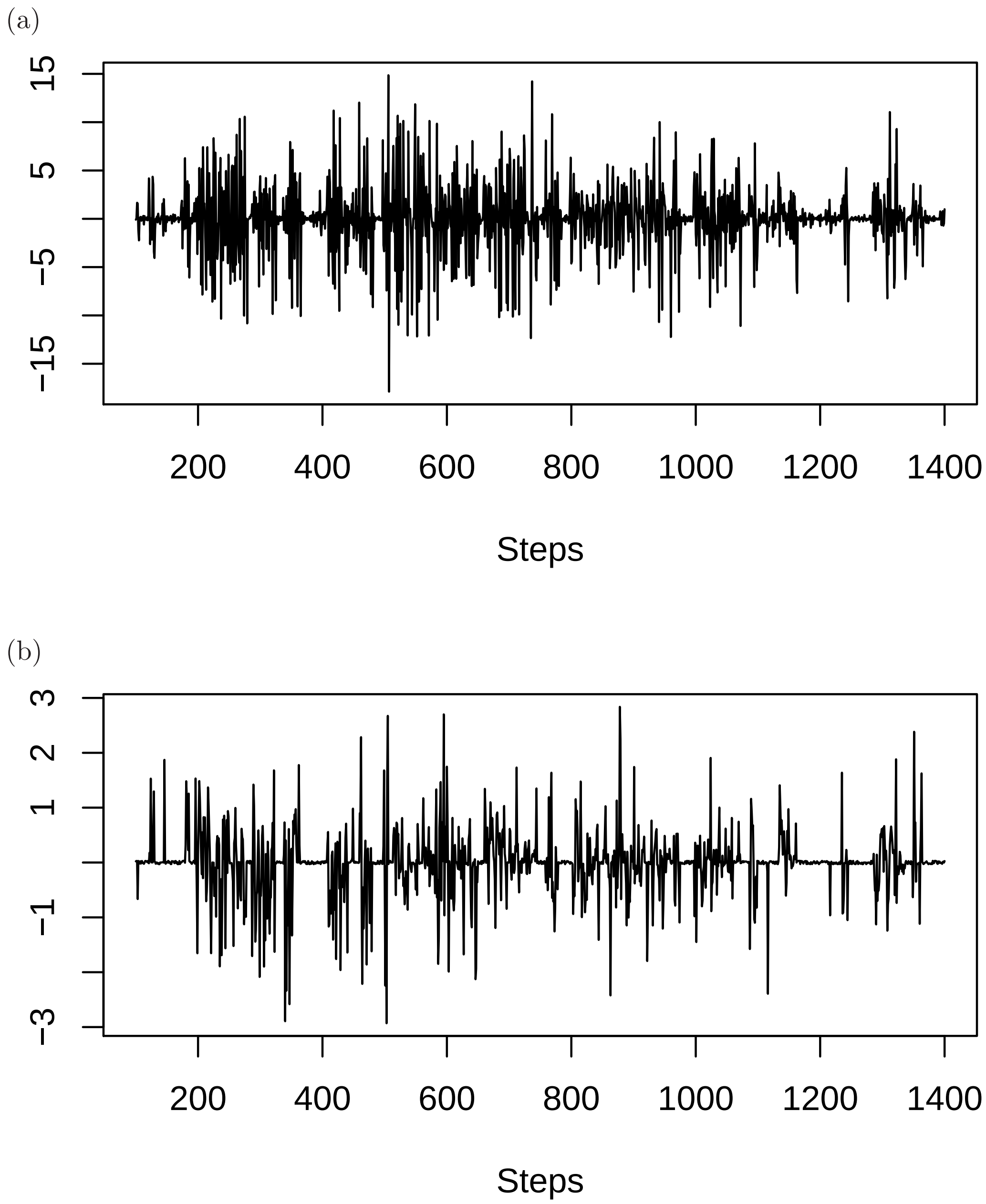

Fig. 4. Element of the velocity's difference (a) and the angular difference (b).

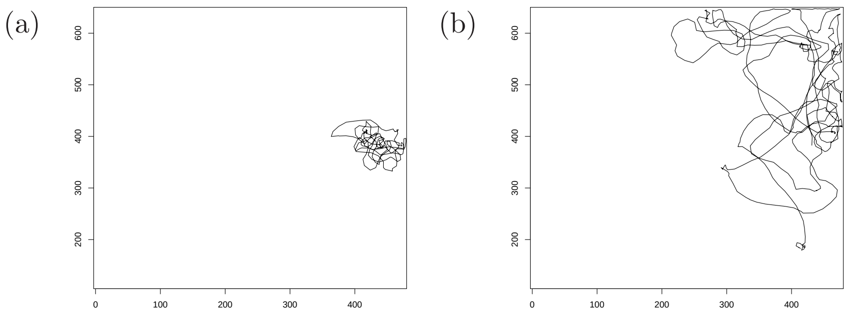

Fig. 5. Trajectory of a fly for one minute after feeding (a), and for about three minutes after the one minute (b). (a) is the interval from 171 to 470 steps and (b) is from 471 to 1370.

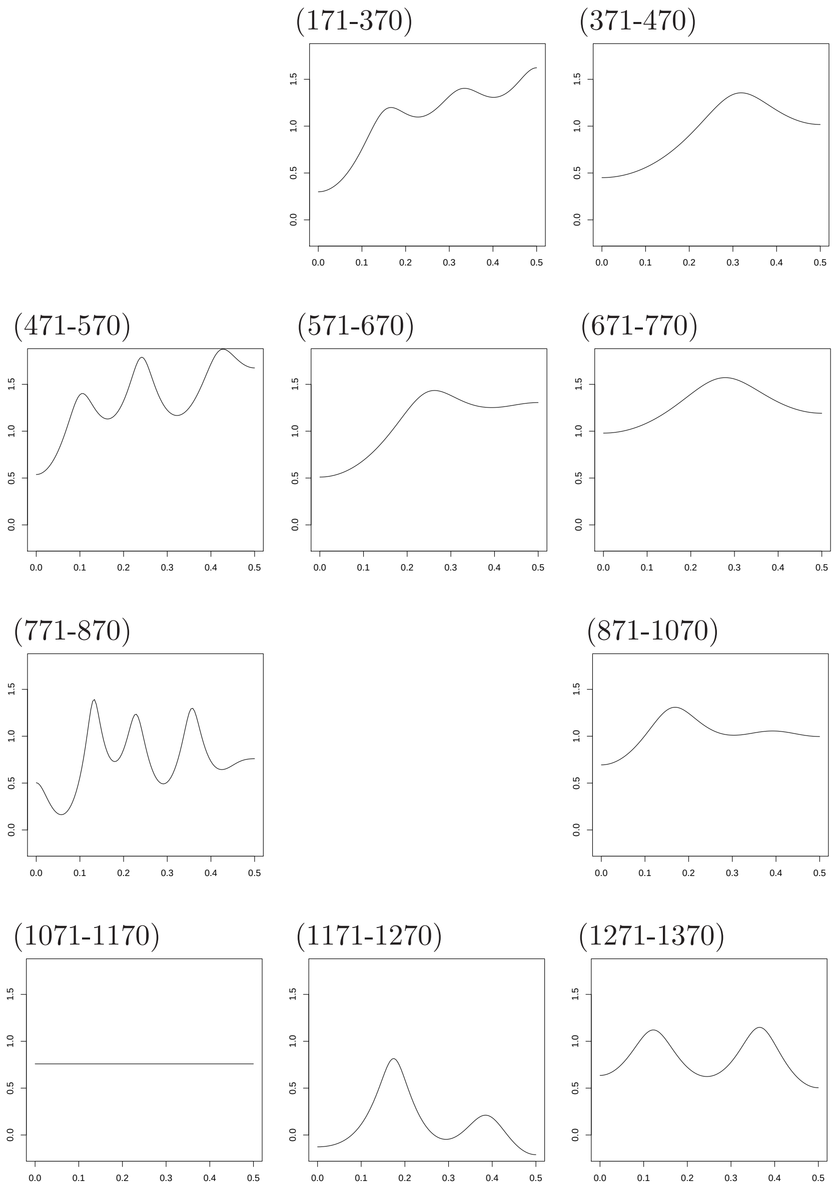

Fig. 6. Rational spectrum for the velocity's difference. The corresponding intervals are described at the top of each figure.

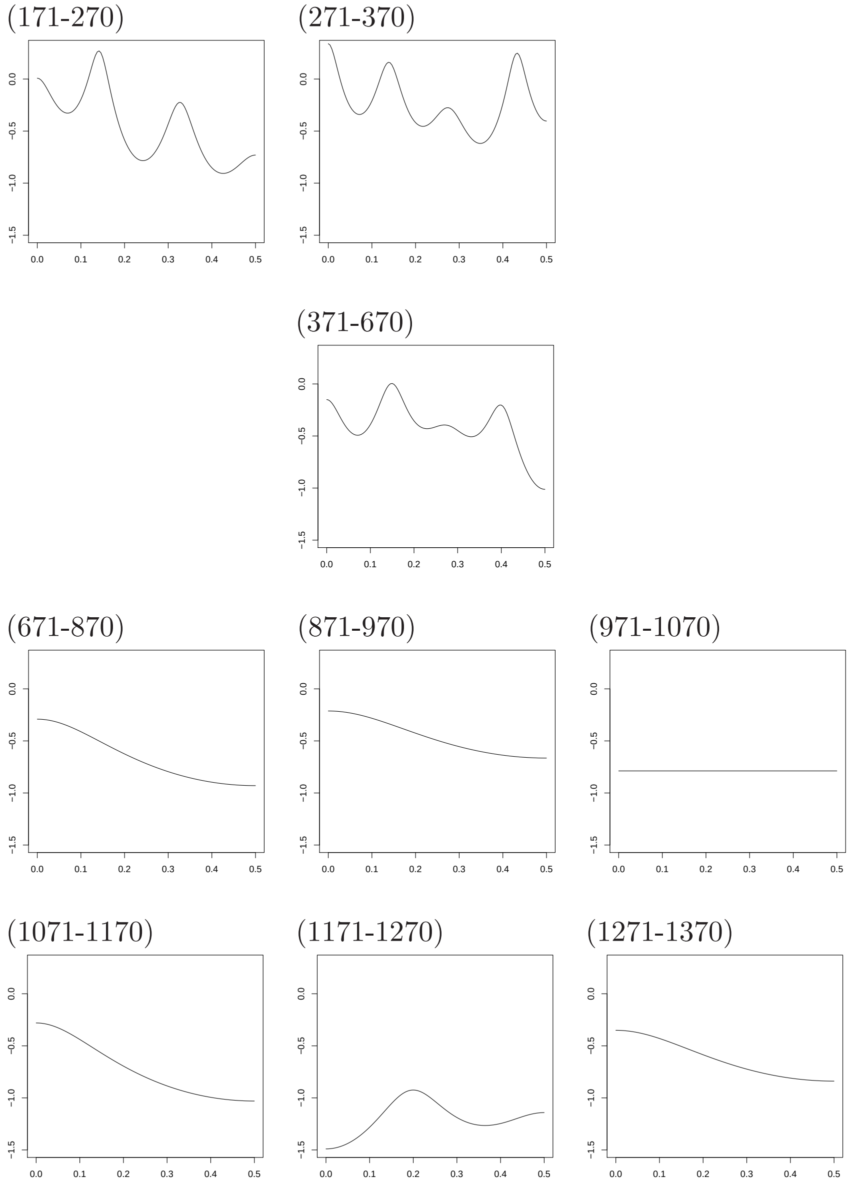

Fig. 7. Rational spectrum for the angular difference. The corresponding intervals are described at the top of each figure.

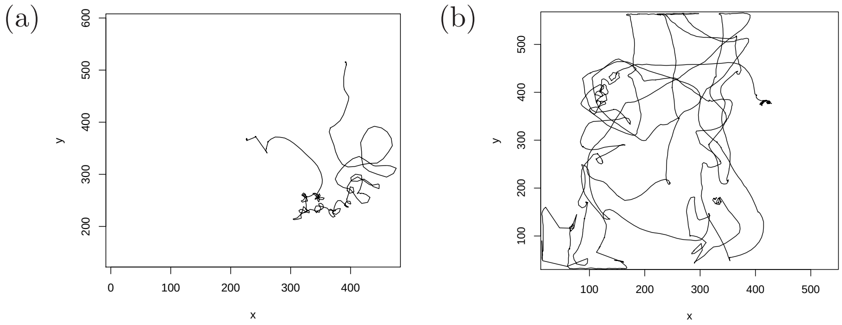

Fig. 8. Trajectory of a fly for about 14 minutes with sugar solution's droplets (a), and for about 27 minutes with no-sugar solution's droplets (b).

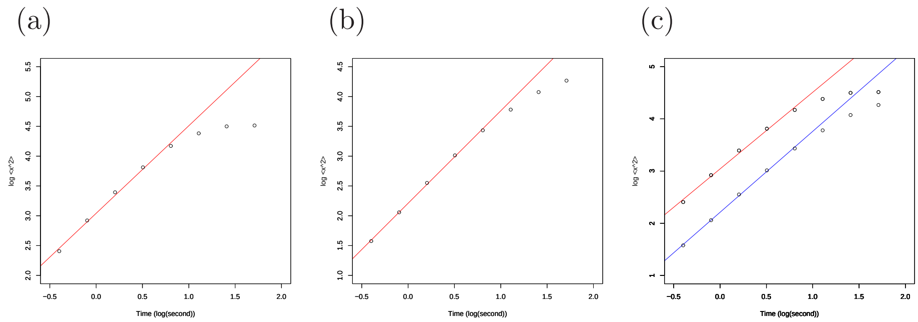

Fig. 9. Log-log plots for the eq. (2) with the duration 0.4, 0.8, 1.6, 3.2, 6.4, 12.8, 25.6, 51.2 seconds and the line represents $\alpha = 1.47$ in Fig. (a) and $\alpha = 1.55$ in Fig. (b). Figure (a) is coresponding the case of with sugar solution: Fig. 2, and Fig (b) is with no sugar solution case: Fig. 8 (b)

(a)

steps	AR order	AR coefficient	σ^2
171 - 370	5	-0.521, -0.322, -0.287, -0.207, -0.173	12.563
371 - 470	2	-0.420, -0.335	8.699
471 - 570	7	-0.425, -0.176, -0.296, -0.121, -0.064, -0.318, -0.142	22.312
571 - 670	3	-0.416, -0.366, -0.168	12.329
671 - 770	2	-0.154, -0.271	19.372
771 - 870	9	-0.149, -0.243, -0.190, -0.220, 0.015, 0.056, -0.054, 0.255, 0.210	5.569
871 - 1070	4	-0.049, -0.157, -0.169, -0.105	10.831
1071 - 1170	0	0	5.737
1171 - 1270	5	0.112, -0.230, -0.195, -0.188, 0.151	1.361
1271 - 1370	4	0.044, -0.043, 0.057, -0.299	6.654

(b)

steps	AR order	AR coefficient	σ^2
171 - 270	6	0.402, -0.069, 0.033, -0.194, -0.043, 0.285	0.350
271 - 370	7	0.102, 0.184, -0.130, -0.003, 0.092, -0.045, 0.284	0.581
371 - 670	7	0.204 -0.1424, 0.060, -0.182, 0.193, -0.049, 0.174	0.390
671 - 870	1	0.352	0.215
871 - 970	1	0.255	0.341
971 - 1070	0	0	0.163
1071 - 1170	1	0.407	0.185
1171 - 1270	3	-0.056, -0.176, -0.176	0.064
1271 - 1370	1	0.274	0.235

Table I. Results of the local RA model for the velocity's difference (a), and for the angular difference (b).