Estimates for Dynamics of Coupled Flagella

Bati Sengul[*][†]

January 18, 2011

Abstract

It is theorised that the dynamics of coupled flagella behave much like a system of weakly coupled noisy oscillators. This enables the biologists to regard the phase difference of the flagella as an SDE and in this case the unknown parameters of this SDE give important biological insight, see for example [PTD$^+$09]. However the techniques used by Polin et. al. do not work in the case when the number of unknowns is large.

Various methods in the literature can be combined to give a powerful and robust method of statistical analysis for SDEs of a specific type. Using this, the paper aims to outline methods of statistical analysis that could be used in analysing the dynamics of coupled flagella.

Keywords: *Coupled Flagella, SDE Parameter Fitting, Kramers Exit Problem, Nonlinear Autoregressive Model*

1 Introduction

Phase dynamics of eukaryotic flagella is responsible for a variety of phenomena in biology ranging from the embryonic left-right asymmetry to the enhancement of nutrient uptake.

Mathematically, in order to obtain the dynamics of the flagella, it is of interest to first make observations on a pair of flagella that are coupled. In [PTD$^+$09] Polin et. al. make observations on flagella dynamics in *Chlamydomonas reinhardtii*, a species of unicellular green algae. The two flagella are thought to be weakly coupled, through a combination of their roots and the fluid in which they beat. The dynamics of these biological objects are similar to a systems of weakly coupled phase oscillators which are well understood, see for example [PRKH02]. If ϕ_1 and ϕ_2 are the phases of the oscillators, then

$$\frac{d}{dt}\phi_1 = \omega_1 + U(\phi_1, \phi_2)$$
$$\frac{d}{dt}\phi_2 = \omega_2 + U(\phi_2, \phi_1)$$

[*] Department of Pure Mathematics and Mathematical Statistics, Cambridge

[†] This work was submitted as a part of the CCA course.

where ω_1, ω_2 are constants describing the initial frequencies and U is a 2π-periodic function in both variables which describes the effects of one particle on the other. We consider the system of phase differences $\phi_1 - \phi_2$ and perturb the system by Gaussian noise, i.e.

$$\dot{X}_t = Q(X_t) + \xi_t$$

where $Q \in \mathcal{C}^\infty$ and has a Fourier expansion and $\{\xi_t\}_{t\geq 0}$ is centred Gaussian white noise. Alternatively we can write this as a stochastic differential equation (SDE):

$$dX_t = Q(X_t)dt + \sigma dB_t \tag{1}$$

where B is a standard Brownian motion on $\mathbb{R}$ and $\sigma > 0$ is the diffusivity constant. One can look at this as modelling the over-damped motion in the potential field $-\int_0^x Q$ while being subject to Brownian noise.

The subject of interest is the Fourier exponents of the function Q and σ which are unknown, thus requiring a method of fitting to data collected during experimentation. The layout of the paper is as follows; in Section 2 the statistical inference is described, Section 3 describes the steady states of the motion which is relevant both physically and statistically, the work of Polin et. al. and other physical aspects are discussed in Section 4 and Section 5 gives some results on simulated paths.

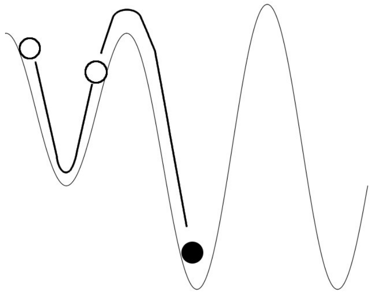


Figure 1: *The visualisation of X as it rolls along a potential well, while getting bombarded with small perturbations.*

Acknowledgements

I am in large debt to Marco Polin for his support and guidance during this project and his continued enthusiasm in my work. I would also like to thank Kostas Papafitsoros and Kolyan Ray for their help with my various questions.

2 Time Series Analysis

Suppose we are given a time series $\{x_{t_n}\}_{n=0}^{N}$ of a continuous processes $X = (X_t : t \geq 0)$, observed at regular intervals δ (without loss of generality we assume that $t_0 = 0$). We would like to use a maximum likelihood estimation on the parameters of Q, for which the likelihood function takes the form

$$\mathbb{P}(X_{t_0} = x_{t_0}, \ldots, X_{t_N} = x_{t_N}). \tag{2}$$

Unfortunately, in order to compute (2) analytically, one must compute the laws of the integrals $\int_0^t Q(X_t)dt$ in terms of the parameters of Q.This in general cannot be done which gives us a reverse problem, that is to estimate an SDE using discrete time steps given that we know it's drift and diffusivity. This problem will be analysed in 2.1 and afterwards we will shed some light on the estimation processes for the discrete estimate in 2.2.

2.1 Discrete Estimations to SDEs

This author is familiar with two estimation methods for SDE which are described below in their full generality. This general disposition is not necessary but certainly very interesting and illuminating.

2.1.1 The Euler Scheme

Consider a general SDE of the form

$$dX_t = c(X_t)dt + \sigma(X_t)dB_t$$

with $X_0 = x$, where σ and c are Lipschitz and are assumed to be known. The Euler scheme estimates the SDE on equidistant time intervals $t_n := n\delta$ by

$$\tilde{X}_{t_n} - \tilde{X}_{t_{n-1}} = c(\tilde{X}_{t_{n-1}})\delta + \sigma(\tilde{X}_{t_{n-1}})(B_{t_n} - B_{t_{n-1}}) \tag{3}$$

and the process is interpolated between the discrete points.

The rate of convergence of this is well known (c.f. [All07]).

Theorem 2.1.1. *The process $\tilde{X}$ converges weakly to X and moreover there exists a constant C depending on t_T such that for each $n \leq T$*

$$\mathbb{E}[|\tilde{X}_{t_n} - X_{t_n}|^2] \leq C\delta.$$

2.1.2 Milstein's Higher Order Method

Th SDE can be approximated better by looking at higher derivatives. In the case of the Euler method, we ignored the higher derivative terms in the Taylor series. Milstein's method works by applying Itô's formula to the lower order terms and obtains an approximation $\tilde{X} = (\tilde{X}_t : t \geq 0)$ as

$$\begin{aligned}\tilde{X}_{t_n} - \tilde{X}_{t_{n-1}} &= c(\tilde{X}_{t_{n-1}})\delta + \sigma(\tilde{X}_{t_{n-1}})(B_{t_n} - B_{t_{n-1}}) \\ &+ \frac{1}{2}\sigma(\tilde{X}_{t_{n-1}})\sigma'(\tilde{X}_{t_{n-1}})((B_{t_n} - B_{t_{n-1}})^2 - \delta).\end{aligned}$$

This approximation converges much faster than the Euler scheme, at order δ (c.f. [KP77]), i.e.

$$\mathbb{E}[|\tilde{X}_{t_n} - X_{t_n}|^2] \leq C\delta^2.$$

2.2 Parameter Estimation

Now let us look at the problem outlined in the introduction i.e. the system described by

$$dX_t = Q(X_t)dt + \sigma dB_t$$

with $Q(x) = \sum_{k\geq 0} a_k \sin(kx) + b_k \cos(kx)$ where a_k, b_k and $\sigma > 0$ are parameters to be estimated. We assume that the function Q is given by a finite sum, hence making it Lipschitz and also treat σ as a nuisance parameter, that is, a parameter which is unknown and assumed to be fixed. Notice that the two methods described above give the same model as σ is constant.

The problem now reduces to estimate the following non-linear AR model[1]

$$x_n = f(x_{n-1}|\theta) + \epsilon_n \tag{4}$$

where $f(x) = x + Q(x)\delta$, ϵ_n are i.i.d. $N(0, \sigma^2\delta)$ and θ is the vector of parameters a_k, b_k. Two methods could be employed here to estimate θ, one is to use a least squares estimate and the second is to use a maximum likelihood estimation. As the error terms are Gaussian, the log-likelihood function is

$$\log \mathcal{L}(\theta) = (2\pi\delta)^{-T/2} - \frac{\sum_{n\leq T}(x_n - f(x_{n-1}|\theta))^2}{2\delta\sigma^2}$$

which is maximised when the square residuals are minimised, thus estimates using least squares and likelihood coincide to give the same result. Also note that the value of σ does not effect the minima of the log-likelihood function, giving justification to the assumption that the parameter σ is a nuisance parameter.

For the estimates on σ we can again use maximum likelihood. Once an estimate $\hat{\theta}$ for θ is obtained, we can assume that $\bar{x}_{n+1} - f(\bar{x}_n|\hat{\theta})$, where $\bar{x}$ are observations on x, are i.i.d. $N(0, \sigma^2\delta)$ samples. Then the unbiased likelihood estimate is given by

$$\hat{\sigma}^2 = \frac{1}{\delta(T-1)} \sum_{n\leq T} (\bar{x}_{n+1} - f(\bar{x}_n|\hat{\theta}))^2.$$

[1] We adopt the convention $t_n = n$ here to avoid heavy notation.

2.2.1 Likelihood Ratio Test

Suppose we have fitted two parameters sets θ_K and θ_M with lengths $K \leq M$. Then we can use a ratio test to determine which model we should pick. Fix a significance level p, we set a hypothesis test of the form $H_0 : \theta = \theta_K$ and $H_1 : \theta = \theta_M$, with the same known diffusivity σ, then the likelihood ratio is given by

$$\Gamma(x) := \frac{\mathcal{L}(\theta_M)}{\mathcal{L}(\theta_K)} = \exp\left(\frac{1}{2\sigma^2\delta}\sum_{n\leq N}[(x_{n+1} - f(x_n|\theta_K))^2 - (x_{n+1} - f(x_n|\theta_M))^2]\right).$$

Let $\bar{x}$ denote the observed quantities, then we reject H_0 when $\Gamma(\bar{x}) \geq c$, where c is a critical value determined by $\mathbb{P}(\Gamma(x) \geq c|H_0) = p$, which we assume is greater than one.[2]

This simple idea is often difficult to work with due to the complicated nature of the probabilities under the non-linear system, that is to say that the distribution of $\Gamma(x)$ is almost impossible to derive analytically. The solution is then given by the asymptotic properties of the likelihood ratio. It is well known (c.f. [GH80]) that asymptotically $\Gamma(x)$ conditioned on H_0 is distributed χ^2_d, a χ^2 random variable with $d = M - N$ degrees of freedom. So the constant c may be calculated as $\mathbb{P}(\chi^2_d \geq c) = p$.

Notice now that $c \mapsto \mathbb{P}(\chi^2_d \geq c)$ is decreasing, hence as we reject the null hypothesis when $\Gamma(\bar{x}) \geq c$, we can reject the null hypothesis if

$$\mathbb{P}(\chi^2_d \geq \Gamma(\bar{x})) \leq p. \tag{5}$$

2.3 Comments

Notice that as in the likelihood estimate we use $x_n - f(x_{n-1}|\theta)$, we are only looking at the one step estimate. In essence our discrete processes is given by

$$\tilde{X_n} = X_{n-1} + Q(X_{n-1})\delta + (B_n - B_{n-1})$$

thus making the constant in the convergence absolute (that is it does not depend on the time interval on which we work in).

This approach to the problem of parameter estimation can be very ineffective if the data set is small, i.e. N is small. The likelihood estimations are well known to deliver closest approximations asymptotically,[3] however they may encounter some problems with a small sample. So in cases of small samples one may wish to adopt an other statistical tools such as the Wald statistic or the Lagrange multiplier.

There are however considerable advantages to using likelihood statistics. It is efficient, and under certain circumstances it is also strongly consistent, that is the likelihood estimate converges a.s. to the true value of the parameters, see for example [Fry80].

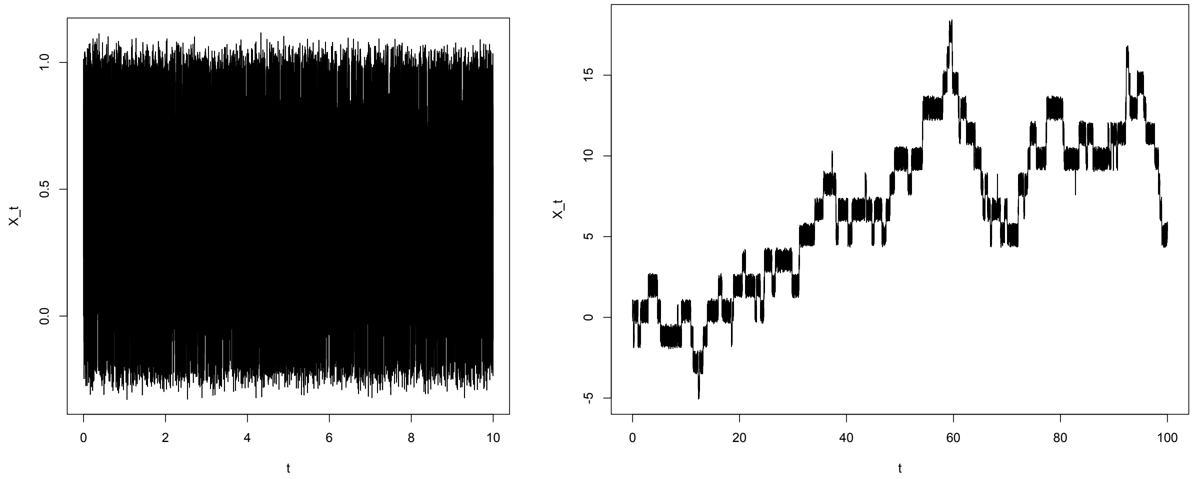
[2] Notice that this requirement then forces the hypothesis test to be the most powerful test in the sense of the Neyman-Pearson lemma.

[3] In other words the estimator is efficient, i.e. has the lowest mean square error asymptotically.

3 Stable Steady States

3.1 Kramers Exit Problem

Stable steady states (here after referred to as steady states) is the point in which the motion of the particle becomes confined around a point due to the strong influence of the drift. If we think of X_t as a ball rolling in a field while being perturbed by Gaussian noise (see Figure 1), we notice that at some "wells", the noise may take a long while to get the ball to escape. Figure 2 illustrates this problem. On a small scale, the process looks like noise, while actually the process is jumping between the steady states.



(a) X at a close distance

(b) X at a larger distance

Figure 2: *The first path shows the process close up, where one may think that this is just the noise. The second picture is zoomed out, showing the steady states where the process hangs around a point due to the strong drift.*

Certainly from a mathematical point of view, the escape time are finite however practically if our observations were taken only during a steady state, the true values of the parameters may not be extracted. The important factor in this is the expected time of escape.

Denote by V the potential field on which our ball is rolling, that is

$$V(x) := -\int_0^x Q(u)du$$

Let $A := \{x \in \mathbb{R} : V'(x) = 0, V''(x) < 0\}$ be the set of local maxima, $B := \{x \in \mathbb{R} : V'(x) = 0, V''(x) > 0\}$ be the set of local minima and $T := \inf\{t \geq 0 : X_t \in A\}$ the first escape time.

The problem is then if $X_0 = x_0$ for some $x_0 \in B$, how does T behave? In particular we would be interested in $\mathbb{E}_{x_0}[T]$.

This problem was first considered by Kramers in [Kra40] where he considers the motion of a particle in a potential well, perturbed by Brownian noise. We follow the more recent account given by [MS82]. Let $T_z := inf\{t \geq 0 : X_t = z\}$, $u := \inf\{y \in A : y > x_0\}$ and $v := \sup\{y \in A : y < x_0\}$ denote the escape points to the right and left of the starting location respectively. Then $g(x) := \mathbb{E}_x[T_u]$ solves Dynkin's problem:

$$\mathcal{L}g = -1$$

for $x < u$, where $\mathcal{L} := \frac{1}{2}\sigma^2\Delta + Q(x)\nabla$ is the generator of the SDE.

Now the boundary conditions are given by $g(u) = 0$ and $g(-\infty) = \infty$, hence the solution is given by

$$\mathbb{E}_x[T_u] = Ce^{(V(u)-V(x))\frac{2}{\sigma^2}}. \tag{6}$$

3.2 Exiting to the Side

Now let us examine the probability of exit on the left or right of the potential well. Define

$$s(x) := \int_0^x e^{-2V(l)/\sigma^2} dl$$

then by an application of Itô's formula $M_t := s(X_t)$ is a continuous local martingale. We may apply optional stopping to M_t^T as this is a bounded martingale, which gives

$$\mathbb{E}[M_T] = M_0 = M_{T_u}\mathbb{P}(T_u < T_v) + M_{T_v}\mathbb{P}(T_v \leq T_u).$$

The case when $s(v) = s(u)$, i.e. the well is symmetric, then the probabilities are both a half, otherwise we have that

$$P_+ := \mathbb{P}(T_u < T_v) = \frac{s(v) - s(x_0)}{s(v) - s(u)}$$

which is the probability of escaping to the right of the well. Similarly we have

$$P_- := \mathbb{P}(T_u < T_v) = \frac{s(x_0) - s(u)}{s(v) - s(u)}$$

as the escape probability to the left of the well.

3.3 Application to Data Fitting

While the methods of Section 2.2 work in theory, in practice there may be problems such as an estimate of θ that is obviously wrong, but happens to minimise the residuals better than the true value of θ. These are not too uncommon while doing simulations where one knows what the true values of parameters, and can be prevented by constricting parameters to a certain range. However in practice, we do not know the true values so an indication of the rough range of the parameter values are immensely helpful.

The idea is that looking at a set of data we can see and calculate the average time between the "slips", i.e. when the rolling ball exits the well. From (6) we see that log of this is roughly the height of the wells that appear in V, and the maximal height of the well in V is greater than $\max_{k\leq N}\{\frac{a_k}{k}, \frac{b_k}{k}\}$. Hence we can deduce that we would expect to see the maximal value of θ to be near $N\sigma^2 \log \tau_{slip}$ where τ_{slip} is the expected time of slipping.

4 Physical Interpretations

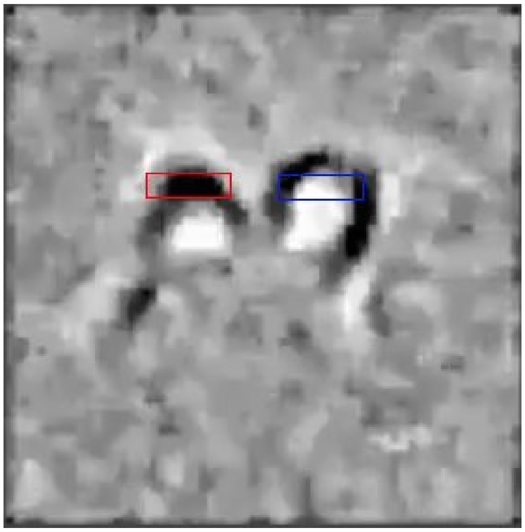


Figure 3: *The measurements taken during experiments by Polin et. al. The red and blue boxes are used to find the first entrance time of each flagellum.*

Following the mathematical exposition we can see that using SDE model

$$dX_t = Q(X_t)dt + dB_t$$

if indeed the physical phenomena follows this SDE then we would expect to see stationary noise followed after an exponential time by slips. This is indeed what one can observe from the data collected by Polin et. al.

The physical phenomena of *Chlamydomonas reinhardtii* flagella phase synchronization was measured by looking at the first entrance time of each flagellum into a given box (see Figure 3). Then it is hypothesised in the paper that the difference in the phases $\Delta := \phi_1 - \phi_2$, where ϕ_1 and ϕ_2 are the measurements taken on each flagellum, is given by the simple model

$$d\Delta_t = (a + b\sin(\Delta_t))dt + \sigma dB_t$$

where a, b, σ are constants. They use the fact that the autocorrelation decay of Δ is estimated by $Ce^{\frac{-t}{\tau}}$, with C, τ as constants, to estimate the parameters. Specifically, the three parameters (C, τ, P_+, P_-) result in estimations of the form $\sigma^2 = \frac{\pi C}{\tau}, a = \pi\sigma^2 \log(P_+/P_-)$ and $b = 2\pi(a^2 + (2\pi\tau)^{-2})^{-1/2}$.

The exposition shown here approaches the problem in other direction, once we have estimated parameters of the model, one can then compute the quartet (C, τ, P_+, P_-). While the two methods of parameter estimation coincide for this case (and indeed any case with less than four parameters), the method used by Polin et. al. fails to work for cases when the number of unknown parameters are four or more.

5 Simulations

5.1 Methodology

In the results that follow we have taken a small time interval (in the range of 10^{-6}) to approximate the SDE as closely as possible. From this we have drawn the data at the sample rate given. The parameters for the simulations were taken from a uniform random variable on $[-5, 5]$, the variance was taken to be the identity and we have omitted the constant drift factor. The code used in the simulations is given in the appendix.

The fitting was done using the *nls* function built in R. Some simulations were done under less than ideal situations to highlight the possible weaknesses in the fitting method.

5.2 Results

Points	Sampling Interval	Parameters	Error (avg., per cent)
1000	0.1	1	$8 \cdot 10^{-6}$
1000	0.1	2	4.616517
1000	0.1	4	14.12664
1000	0.001	3	5.793519
10000	0.001	3	0.79469
10000	0.1	2	0.03375826
10	1	2	514.1449
1000	1	3	86.78013

5.3 Conclusion

As we can see from the simulation results, the theoretical aspects of the effects become apparent. As the discrete approximation relies on the Lipschitz constant of Q, the number of parameters (as well as their values) directly affect the approximation. This can be seen in the bottom of the table where the sampling intervals were high.

Other aspects such as the the sampling interval or the number of points, which is needed for the likelihood estimation to give close results, is also seen in the table.

6 Epilogue

From the analysis on the fitting methods there are three main factors influencing the errors in the data fitting, which can be minimised as follows:

- the observations should be taken frequently
- the quantity of observations should be high.

The author hopes that the methods described here could be used in analysis of coupled flagella even when the coupling function Q is not simple. The fit could then be checked by computing P_+ and P_- and comparing this to the estimates obtained from the data.

There is also an extension of this theory to when the diffusivity is not constant, i.e. to consider an SDE of the form

$$dX_t = Q(X_t)dt + W(X_t)dB_t$$

with some regularity conditions on W. One can use either method of discrete estimation for this and arrive on a likelihood estimate to fit the parameters of W. This is however a lot more computationally expensive, as well as the possibility of impossibility to solve numerically. The likelihood ratio in this case does not work, leaving the hypothesis testing to much more *adhoc* methods.

Appendix

A Simulation Code

The R code used for the simulations is given below.

```
# F is the function f(x) described in section 2
# x is the input vector, A is a vector where the first half is a_k and the
    second half is b_k and del is the sampling interval
F = function(x,A,del) {
        a = A[1:(length(A)/2)]
        b = A[(length(A)/2+1):length(A)]
        ret = numeric(length(x))
        for(i in (1:length(x))){
                for (j in (1:length(a))) {
                                ret[i] = ret[i]+a[j]*sin(j*x[i])+b[j]*cos(j*x[
                                    i])
                        }
                        ret[i] = ret[i]*del+x[i]
                }
                return(ret)
        }
```

```
# Simulates an SDE of the form dX_t = Q(X_t)dt + \sigma dB_t
# N is the number of points, del is the sampling interval, A is a vector where
     the first half is a_k and second half b_k and sig is the diffusitivity of
    the Brownian motion
Sim = function(N,del,A,sig){
        x = numeric(N)
        B = rnorm(N-1,0,del*sig^2)
        x[1] = 0
        for(i in 1:(N-1)){
                        x[i+1] = F(x[i],A,del)+B[i]
                }
                return(x)
        }
```

References

[All07] E. Allen, *Modeling with Itô stochastic differential equations*, Springer Verlag, 2007.

[Fry80] R. Frydman, *A proof of the consistency of maximum likelihood estimators of nonlinear regression models with autocorrelated errors*, Econometrica: Journal of the Econometric Society **48** (1980), no. 4, 853–860.

[GH80] A.R. Gallant and A. Holly, *Statistical Inference in an Implicit, Nonlinear, Simultaneous Equation Mode in the Context of Maximum Likelihood Estimation*, Econometrica: Journal of the Econometric Society (1980), 697–720.

[KP77] P.E. Kloeden and RA Pearson, *The numerical solution of stochastic differential equations*, The ANZIAM Journal **20** (1977), no. 01, 8–12.

[Kra40] HA Kramers, *Brownian motion in a field of force and the diffusion model of chemical reactions*, Physica **7** (1940), no. 4, 284–304.

[MS82] B. Matkowsky and Z. Schuss, *Kramers' diffusion problem and diffusion across characteristic boundaries*, Theory and Applications of Singular Perturbations (W. Eckhaus and E. de Jager, eds.), Lecture Notes in Mathematics, vol. 942, Springer Berlin / Heidelberg, 1982, 10.1007/BFb0094756, pp. 318–345.

[PRKH02] A. Pikovsky, M. Rosenblum, J. Kurths, and R.C. Hilborn, *Synchronization: A universal concept in nonlinear science*, American Journal of Physics **70** (2002), 655.

[PTD$^+$09] M. Polin, I. Tuval, K. Drescher, JP Gollub, and R.E. Goldstein, *Chlamydomonas Swims with Two" Gears" in a Eukaryotic Version of Run-and-Tumble Locomotion*, Science **325** (2009), no. 5939, 487.