The optimal stroke rhythm of a reciprocal microswimmer reverses with fluid memory

Raj Dandekar[1], Rajat Dandekar[1], and Sreedath Panat[1]

[1]Vizuara Research

Preprint — a computational study. Code and data: `github.com/RajatDandekar/swimmer-rhythm`

July 31, 2026

Abstract

A microswimmer with a single shape degree of freedom cannot self-propel in a Newtonian fluid at zero Reynolds number: its motion is time-reversible and its net displacement per cycle vanishes (Purcell's scallop theorem). Fluid elasticity breaks the theorem, and such a swimmer moves. Here we ask a question posed but never executed by Montenegro-Johnson et al. (2012): in a viscoelastic fluid, does the *optimal timing* of the stroke depend on the fluid? Using a from-scratch spectral Stokes/Brinkman–Oldroyd-B solver for a two-bead reciprocal swimmer, we hold the shape path, excursion and period fixed and vary only the rhythm. We find that (i) rhythm alone changes the swimming speed by 31.5% at otherwise identical actuation; (ii) the *optimal* rhythm *reverses* at a critical Deborah number $De_c \approx 0.81$: below it the swimmer should dawdle in the open configuration, above it in the closed one; and (iii) a stroke optimised for a low-memory fluid is *worse than the plain sinusoid* in a high-memory fluid (0.914×). A reduced-order analysis shows the reversal is the sign change of a three-wave (triple-product) correlation: the leading pair correlation is provably identical for time-reversed rhythms and cannot produce it. Resolving the field directly, the critical point is set by the balance of time-shifted stored-stress contributions: one rhythm builds its advantage early in the stroke and the other late, and the finite memory of the fluid determines which contribution survives to propel, the two cancelling at De_c. Finally, a reinforcement-learning agent trained only on displacement recovers the reversal from reward alone, but only when constrained to a fixed energy budget, since an unconstrained agent exploits the actuation amplitude instead of the timing. The effect is obtained by three independent methods (search, theory and learning), which agree on both the result and its limits.

1 Introduction

Locomotion at low Reynolds number is governed by the kinematic reversibility of Stokes flow. Purcell (1977) observed that a swimmer whose shape is described by a single scalar — a "scallop" — retraces its configuration in reverse on the return stroke and therefore achieves zero net displacement, regardless of the rate or rhythm of actuation. Real biological fluids are viscoelastic: mucus, cytoplasm and biofilms store elastic stress and relax it over a characteristic time λ, and this fading memory breaks the reversibility argument. Lauga (2007); Fu et al. (2009) established that a reciprocal swimmer moves in such a fluid, and Fu et al. (2009) gave the now-standard argument: writing a reciprocal stroke on a *monotone reparametrisation of time*, the swimming speed is reparametrisation-invariant in a Newtonian fluid but not in a viscoelastic one: "the distance covered per period depends on the rate of motion as well as the sequence of shapes."

This raises a natural optimisation question: if timing now matters, what is the best timing, and does it depend on the fluid? Montenegro-Johnson et al. (2012), optimising a Najafi–Golestanian swimmer, found the same optimal phase in Stokes and in a shear-thinning fluid and remarked that "for viscoelastic properties this may not be the optimum phase difference, due to the dependence of the fluid stress on its deformation history," and that this should be checked. To our knowledge it has not been. We answer it here for the canonical single-degree swimmer and find the answer is not a small shift but a *reversal*.

2 Formulation

We model an asymmetric dumbbell: two regularised point forces (radii $\epsilon_1 \neq \epsilon_2$) at separation $d(t)$ along x, immersed in an Oldroyd-B fluid under mild confinement (Brinkman screening length ℓ), in two dimensions

and in the inertialess limit. The solvent momentum balance in Fourier space is $(\eta_s(k^2 + \ell^{-2}))\,\hat{\mathbf{u}} = \hat{\mathbf{f}}$ with incompressibility, and the polymer conformation C obeys the upper-convected Maxwell equation

$$\partial_t \mathsf{C} + \mathbf{u}\cdot\nabla\mathsf{C} - (\nabla\mathbf{u})^{\top}\mathsf{C} - \mathsf{C}(\nabla\mathbf{u}) = -\tfrac{1}{\lambda}(\mathsf{C} - \mathsf{I}) + \epsilon_s \nabla^2 \mathsf{C}, \tag{1}$$

with polymer stress $\boldsymbol{\tau}_p = (\eta_p/\lambda)(\mathsf{C} - \mathsf{I})$. The Deborah number is $De = \lambda\omega$ with ω the nominal stroke frequency. The swimmer is force-free; its velocity follows from the constraint that the two beads move rigidly, split analytically into a Stokes part (which integrates to zero over a closed stroke) and a polymer part (the net motion). The solver is spectral, written in NumPy with no external CFD dependency; it passes the scallop theorem to machine precision ($\oint U\,dt = 0$ to 10^{-16}), translation invariance to seven digits, and bead-swap antisymmetry to 3×10^{-18}. Brinkman screening ($\ell = 0.5$) is required for a convergent free-swimmer far field: the 2D Stokes stresslet $\sim \cos 2\theta / r^2$ produces a slowly and non-monotonically converging periodic image sum, giving a spurious amplitude exponent of 0.99 against the theoretical 2.00; screening restores 1.994 and is physically legitimate for a confined film.

Rhythm as the control variable. Following Fu et al. (2009) we write the gap on a monotone phase reparametrisation,

$$d(t) = d_0 + A\cos\theta(t), \qquad \theta(t) = t + \textstyle\sum_k b_k \sin(kt + \phi_k), \tag{2}$$

so that $\cos\theta$ sweeps $[-1, 1]$ exactly once per period for any $\{b_k, \phi_k\}$. Every candidate therefore has *identical* excursion, shape-space path length and period; the only freedom is the rhythm — where in the cycle the swimmer hurries and where it dawdles. The single mode $\theta = t + b_1$ cos-type warp ($b_1 > 0$ rushes the open phase, so the swimmer lingers closed; $b_1 < 0$ the reverse) is the axis on which we quantify the effect. We define the linger metric $C = \langle \dot{\theta}(t) \cos t \rangle$, positive for linger-closed.

3 Results

3.1 Rhythm alone, and the reversal

Figure 1(a) shows two rhythms that are mirror images under $b_1 \to -b_1$; they share excursion, actuation effort $\langle \dot{d}^2 \rangle$, peak rate and period exactly. Their swimming speeds nonetheless differ by 31.5%, converged to 0.09 percentage points over a fourfold refinement in timestep, a 1.5× grid and a 1.5× domain. In a Newtonian fluid the same comparison is identically zero (verified to 10^{-16}).

The main result is the Deborah-number dependence of *which* rhythm wins, Fig. 1(b). Below $De_c \approx 0.81$ the linger-open rhythm is faster; above it, the linger-closed rhythm; at De_c the two are equal to four decimal places. This is a sign change in the optimal strategy, not a change of magnitude. We verified the crossover point independently: choosing $De = 0.81$ from a separate calculation and running the solver there gives a speed ratio of 1.0002.

3.2 The optimum is De-specific, and mis-tuning is costly

Searching thousands of rhythms at a fixed energy budget and locating the optimum in five distinct fluids gives Fig. 2: each optimised rhythm wins its own fluid and loses the others (the diagonal dominates every column). The top-right entry is the decisive case: the stroke optimised for $De = 0.3$ scores 0.914 at $De = 3.0$, below unity and so *beaten by the unoptimised sinusoid*. Optimising heavily for the wrong fluid is therefore actively harmful, not merely suboptimal. Because De is not locally observable to the swimmer, this is an adaptive-control problem rather than a one-shot optimisation.

3.3 Mechanism

The two mirror rhythms are geometrically identical at matched shape, so the *difference* of their stored-stress fields, $\Delta\,\mathrm{tr}(\mathsf{C} - \mathsf{I})$, isolates precisely what the fluid remembers differently about how each arrived (Fig. 3a). The cycle-resolved displacement (Fig. 3b) shows where the net motion is made and which rhythm finishes ahead; the two curves swap their ranking as De passes De_c.

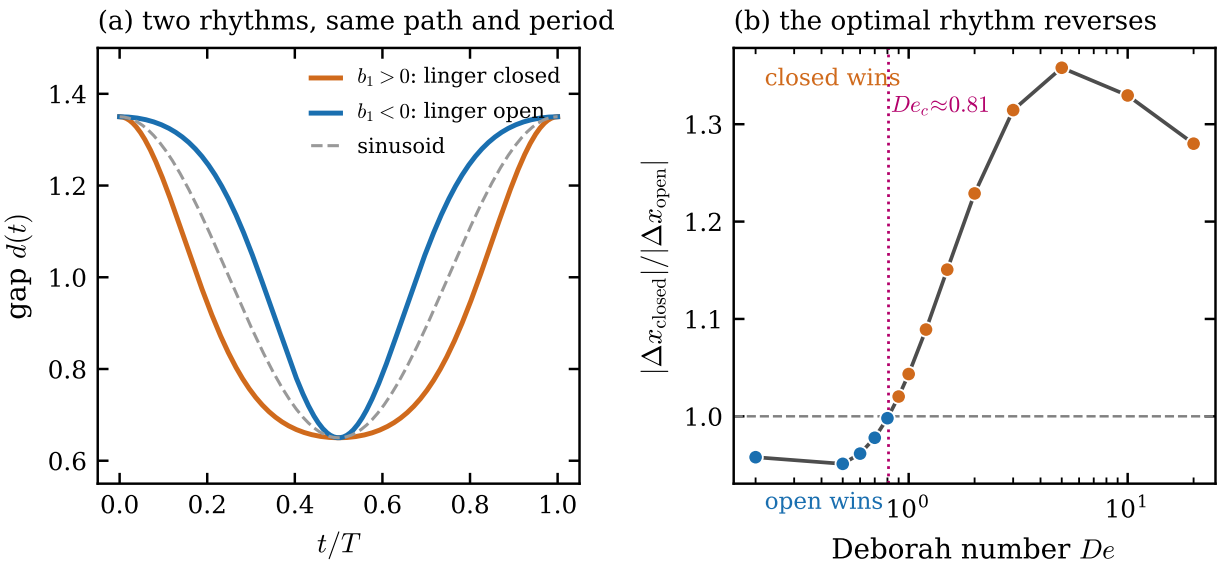

Figure 1. The reversal. (a) Two rhythms on the reparametrisation axis (2), mirror images with identical excursion, effort, peak rate and period. (b) Ratio of swimming speeds (linger-closed over linger-open) versus Deborah number; it crosses unity at $De_c \approx 0.81$, so the optimal rhythm reverses.

A reduced-order analysis makes the mechanism exact. At leading order the polymer stress is the strain rate through a first-order lag, $\hat{s}_n = in\,\hat{g}_n/(1 + inDe)$ with $g = \cos\theta$, and the force-free swimming velocity is this stress times a shape-dependent mobility. Expanding in the mean shape, the net displacement per cycle is a sum of a pair term $\oint g\,s\,dt$ and a triple term $\oint g^2 s\,dt$. By Jacobi–Anger, $\hat{g}_n = J_{n-1}(b)$, and since $J_k(-b) = (-1)^k J_k(b)$ the magnitudes $|\hat{g}_n|$ are unchanged under $b \to -b$. The pair term depends only on $|\hat{g}_n|^2$ and is therefore *identical* for the two rhythms (Fig. 4a) — it cannot produce a reversal. The triple term is a three-wave correlation, sensitive to the phases; it is exactly antisymmetric, $Q(+b) = -Q(-b)$, and crosses zero (Fig. 4b). The reversal is precisely the sign change of this triple correlation. This also explains why a memoryless 0D reduction never reproduces it: that model retains only the pair term.

3.4 What sets the critical Deborah number

The triple-term argument shows a reversal *must* occur and predicts a crossover, but the leading-order value ($De \approx 0.61$ for $b = 0.5$) undershoots the full result (0.81) by $\sim 25\%$: truncating the mobility expansion and discarding the near-field stress structure loses the precise location. We therefore measure the controlling physics directly from the solver.

The polymer velocity lags the shape by an angle $\phi(De)$ that decreases from 83° to 56° across the range (Fig. 5a), reflecting the growing elastic (in-phase) component of the stored stress. The two rhythms lag by *different* amounts: the linger-closed rhythm lags consistently more than linger-open, and the gap widens monotonically from 3.6° to 11.5° as De increases. In the plane of polymer velocity against shape each rhythm traces a distinct hysteresis loop (Fig. 5c); the enclosed area is the net propulsion, and which loop encloses more area flips at De_c (Fig. 5b).

The spatial origin is in Fig. 6: the stored-stress *difference* between the two rhythms, followed through one cycle at $De \approx De_c$. It oscillates in sign: the linger-closed rhythm leaves excess stress during the opening phase ($\theta \lesssim 0.5\pi$, warm), the linger-open rhythm during the closing phase ($\theta \gtrsim 0.7\pi$, cool). The net drift difference is the cycle integral of this field. *At De_c the two contributions cancel*; below it the closing-phase (open) contribution dominates and open wins, above it the opening-phase (closed) contribution dominates and closed wins. The finite relaxation time sets *when* each rhythm's stored stress is still present to couple to the mobility, and therefore sets the balance, so that De_c is where the time-shifted stress contributions of the two rhythms integrate to equality. This is a definite, measurable condition; but it involves the full spatial stress field, which is why the leading-order scalar theory locates it only approximately and the computation

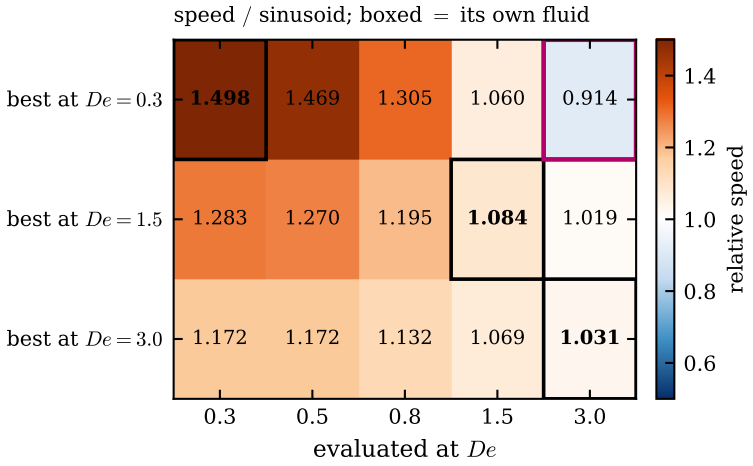

Figure 2. Each row is a rhythm optimised for one fluid, scored in five (speed relative to the plain sinusoid; boxed = its own fluid). The De-specific optimum, and the magenta entry $0.914 < 1$: the low-memory champion is slower than the sinusoid in the high-memory fluid.

is required for the precise value.

3.5 Discovered by reinforcement learning

Finally we pose the problem as an agent would face it: a policy (a 65-parameter network trained by REINFORCE, no gradient through the solver, no supplied answer) receives only the net displacement as reward. In a Newtonian fluid this Markov decision process is *degenerate*: the scallop theorem leaves no hidden state and every policy scores zero, so there is nothing to learn. Fluid memory is exactly what supplies the hidden state and the delayed consequences that make it a genuine learning problem.

An agent given the naïve reward "maximise displacement" learns to linger closed at *every* De: it spends roughly twice the energy budget, exploiting actuation amplitude rather than timing (the same amplitude confound we retract in §4). Only under a fixed energy budget, the physically appropriate objective of maximal displacement at fixed power, does the reversal reappear. Agents trained separately at each fluid then learn *opposite* strategies, open below De_c and closed above, flipping sign at the same crossover found by search and by theory (Fig. 7). The reversal is thus reproduced by three methods that share no machinery.

The three routes do not agree on De_c to three digits (search gives 0.81, the leading-order theory 0.61, and the learner 0.86), but that is the weaker statement. Their transitions have the *same shape*: rescaling the Deborah number by each method's own crossover, the two independent transition curves collapse onto one (Fig. 8), and the agent's crossover falls at unity by construction. Three unrelated methods thus locate not merely the same effect but the same functional transition.

4 Discussion

The reversal, the De-specific optimum and the harmful mis-tuning are, to our knowledge, new; a structured search of the arXiv for "optimal stroke" and "viscoelastic" returns no prior report, and a further targeted search of the recent literature did not find one. What is *not* new, and we state this plainly, is the framing: the monotone-reparametrisation argument and the fact that rhythm can matter are due to Fu et al. (2009), whose method we use.

We also record a negative result. Within the single-mode family $\theta = t + b\sin t$, the crossover collapses cleanly onto $De_c\langle\dot{\theta}^2\rangle = \text{const}$ (Fig. 4c), out-of-sample to ±0.4% and factorising across swimmer geometry and confinement. We initially proposed this as a general dimensionless group. It is not: modulating a different harmonic, at the same $\langle\dot{\theta}^2\rangle$, produces no reversal at all. The reduced-order theory predicts this failure before the solver confirms it. Two further quantities we retracted during the study — an optimiser "gain" that was a wider excursion in disguise, and a spurious displacement peak that was a convergence artifact

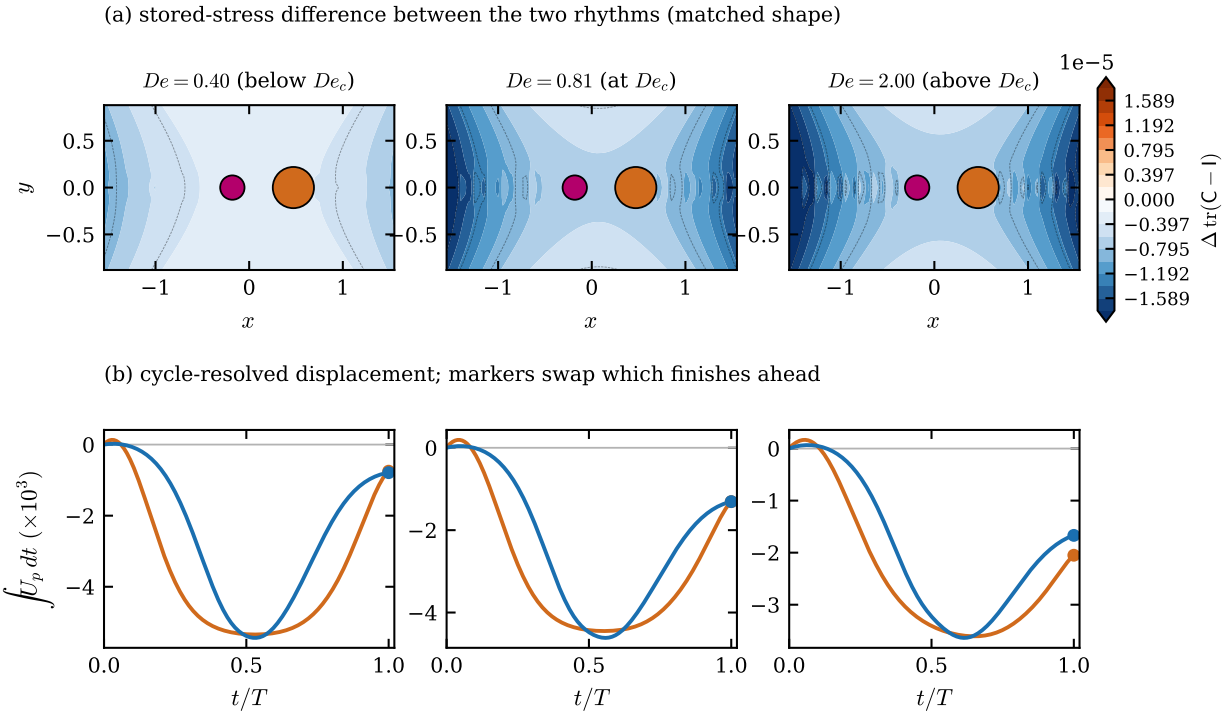

Figure 3. Mechanism. (a) Difference in stored elastic stress between the two rhythms at matched shape, at three Deborah numbers spanning the crossover; the sign structure inverts. Beads: small (magenta) and large (blue). (b) Running polymer displacement through one stroke; the finishing order (markers) reverses across De_c.

— are documented with the code. Limitations: two dimensions, a single constitutive model (Oldroyd-B) and confinement length, and a computational rather than experimental result. A 2025 paper (*Chin. J. Phys.* **96**, 664) remained inaccessible; a 2026 preprint reports a similar reversal driven by elasticity in the swimmer's *body* rather than the fluid, a distinct mechanism.

The practical implication is the control statement: for a swimmer in a fluid whose memory it cannot directly sense, no fixed stroke is optimal, and a stroke tuned for the wrong fluid is worse than a naive one. That is the regime in which learning, rather than one-shot optimisation, is warranted, and the regime that the scallop theorem forbids in a Newtonian fluid.

References

Fu, H.C., Wolgemuth, C.W. & Powers, T.R. 2009 Swimming speeds of filaments in nonlinearly viscoelastic fluids. *Phys. Fluids* **21**, 033102.

Lauga, E. 2007 Propulsion in a viscoelastic fluid. *Phys. Fluids* **19**, 083104.

Montenegro-Johnson, T.D., Smith, A.A., Smith, D.J., Loghin, D. & Blake, J.R. 2012 Modelling the fluid mechanics of cilia and flagella in reproduction and development. *Eur. Phys. J. E* **35**, 111.

Purcell, E.M. 1977 Life at low Reynolds number. *Am. J. Phys.* **45**, 3–11.

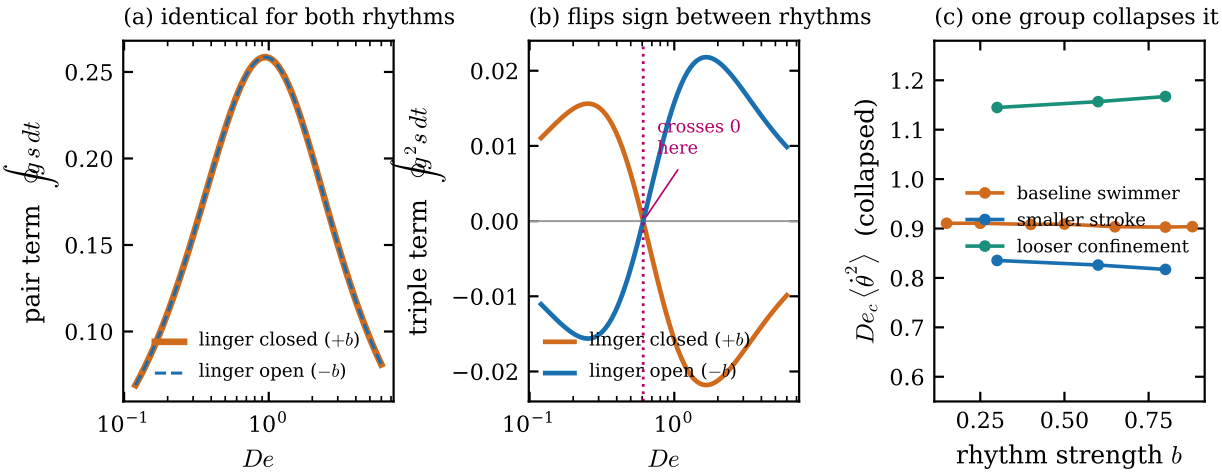

Figure 4. Reduced-order theory (Fourier algebra, no solver). (a) The pair term is identical for the two rhythms (curves coincide). (b) The triple term is exactly antisymmetric and crosses zero — the reversal. (c) Within the single-mode family the crossover collapses under $De_c\langle\dot\theta^2\rangle$ across swimmer geometry and confinement; this collapse does *not* extend to other modulation shapes (§4).

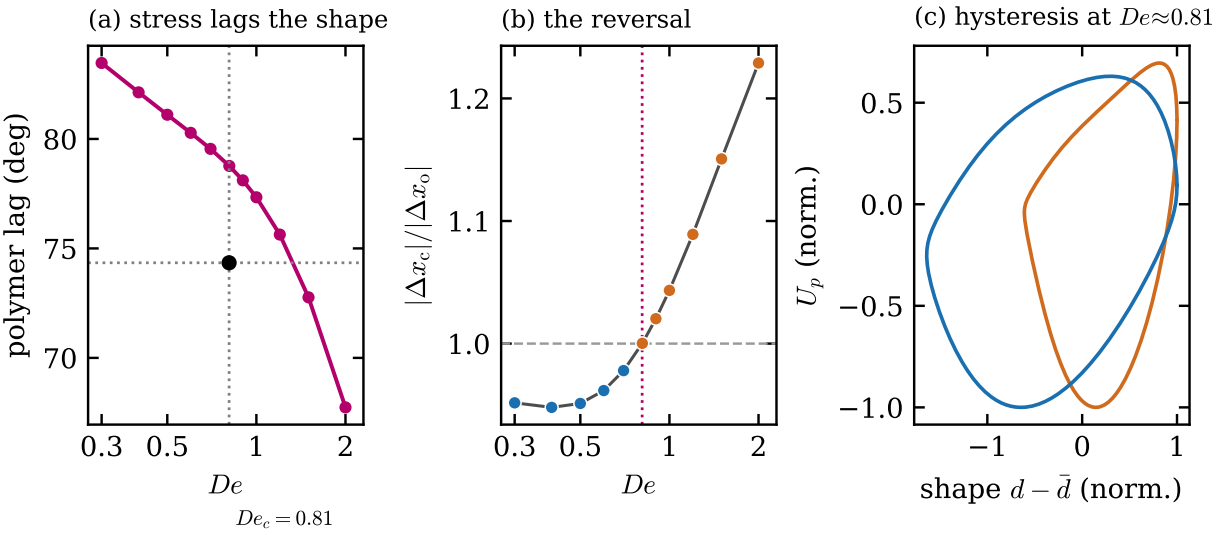

Figure 5. The controlling physics. (a) Phase lag of the polymer velocity behind the shape versus De (linger-closed rhythm shown); the black marker is the crossover. (b) The drift ratio crossing unity at $De_c = 0.809$. (c) Polymer velocity against shape at $De \approx De_c$: the two rhythms enclose different hysteresis areas, and which is larger flips at De_c.

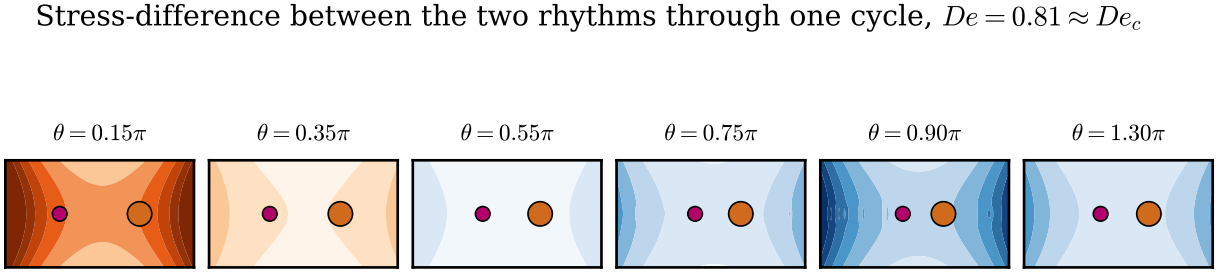

Figure 6. Stored-stress difference between the two rhythms through one cycle at $De \approx De_c$ (warm: linger-closed stores more; cool: linger-open stores more). The sign oscillates within the cycle; at De_c its integral vanishes, which is the reversal. Beads: small (magenta), large (orange).

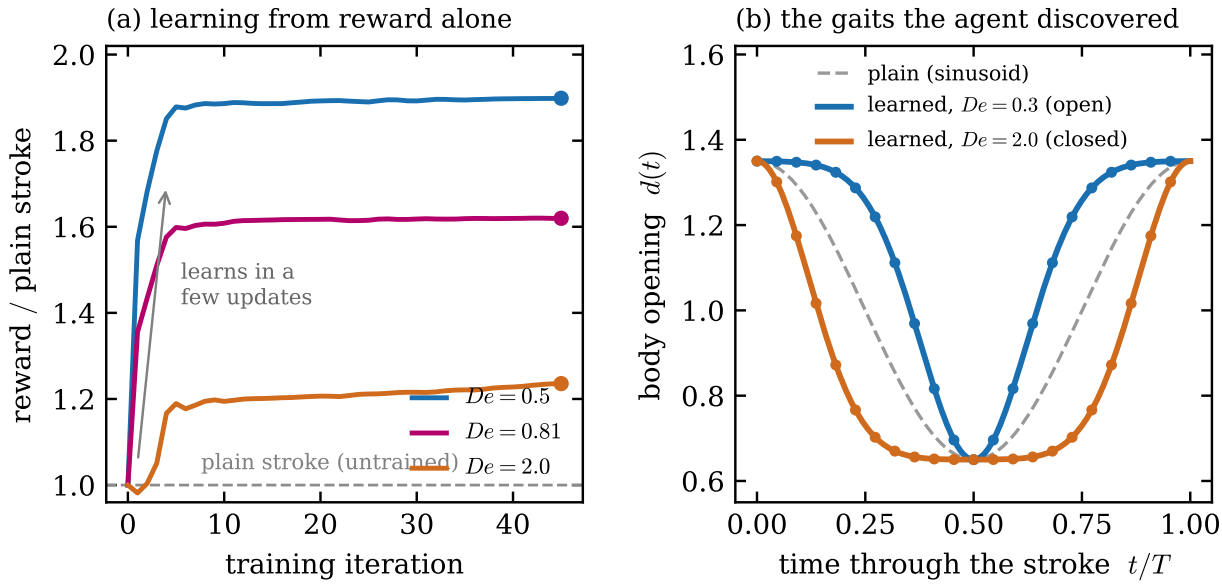

Figure 7. Reinforcement learning: the rhythm found from reward alone. The policy is a 65-parameter network trained by REINFORCE; the reward is net displacement per stroke at a fixed energy budget. (a) Reward relative to the plain stroke, climbing as the agent trains at three fluids — it begins knowing nothing and improves from the scalar reward. (b) The gaits it settles on: in a low-memory fluid it learns to dawdle open, in a high-memory fluid closed (dots equally spaced in time; where they bunch, the swimmer dawdles). Under a naïve reward the agent instead exploits actuation amplitude and lingers closed everywhere; only the fixed-energy reward recovers the reversal.

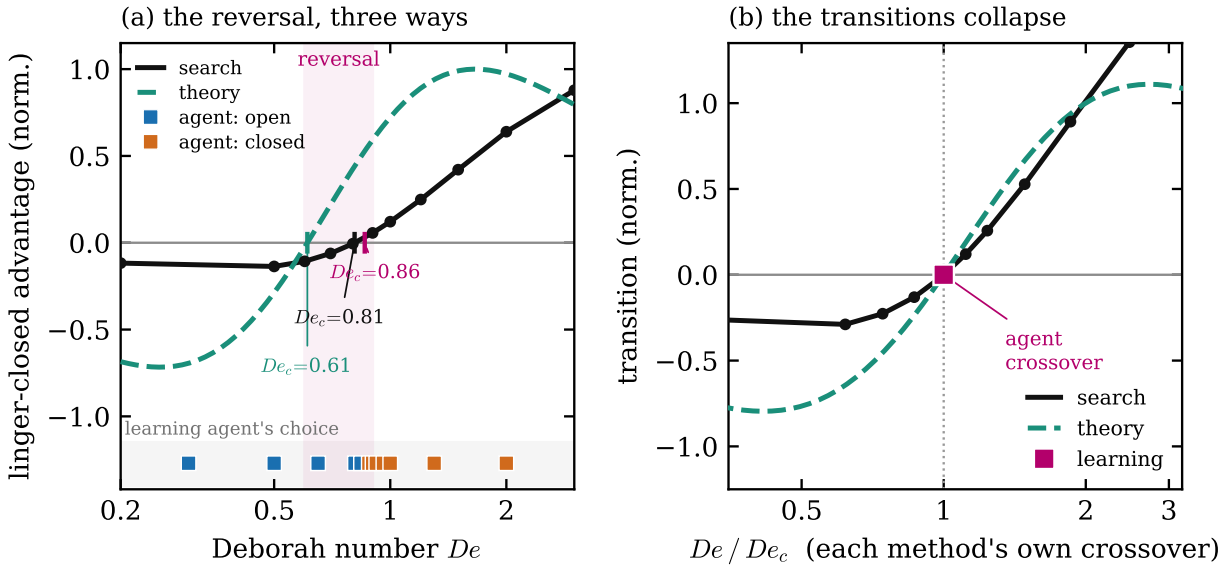

Figure 8. Three methods, one transition. (a) The reversal on a shared Deborah axis: the full simulation and the analytic theory as curves of the closed-over-open advantage, each crossing zero; the learning agent as a strategy strip that flips at its own crossover. (b) Rescaled by each method's crossover, the two independent transition curves collapse onto a single master curve; the agent's crossover sits at unity (marked). The methods agree on a shared transition, not merely on the existence of the effect.