J. Fluid Mech. (2008), *vol.* 601, *pp.* 25–61. © 2008 Cambridge University Press
doi:10.1017/S0022112008000384 Printed in the United Kingdom

Shape optimization of a sheet swimming over a thin liquid layer

JON WILKENING[1] AND A. E. HOSOI[2]

[1]Department of Mathematics and Lawrence Berkeley National Laboratory
University of California, Berkeley, CA 94721, USA

[2]Hatsopoulos Microfluids Laboratory, Department of Mechanical Engineering,
Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

(Received 30 January 2006 and in revised form 21 December 2007)

Motivated by the propulsion mechanisms adopted by gastropods, annelids and other invertebrates, we consider shape optimization of a flexible sheet that moves by propagating deformation waves along its body. The self-propelled sheet is separated from a rigid substrate by a thin layer of viscous Newtonian fluid. We use a lubrication approximation to model the dynamics and derive the relevant Euler–Lagrange equations to simultaneously optimize swimming speed, efficiency and fluid loss. We find that as the parameters controlling these quantities approach critical values, the optimal solutions become singular in a self-similar fashion and sometimes leave the realm of validity of the lubrication model. We explore these singular limits by computing higher-order corrections to the zeroth order theory and find that wave profiles that develop cusp-like singularities are appropriately penalized, yielding non-singular optimal solutions. These corrections are themselves validated by comparison with finite element solutions of the full Stokes equations, and, to the extent possible, using recent rigorous *a priori* error bounds.

1. Introduction

Peristaltic locomotion is ubiquitous among invertebrates; it is used by flatworms, limpets, annelids (e.g. earthworms), gastropods and even sea anemones, which employ peristalsis for digging (Vlès 1907; Trueman 1975; Denny 1980). In each of these organisms, muscles drive a wave of deformation along the body, much like a waving sheet, which in turn generates fluid stresses that propel the animal forward. This mechanism is adopted by a wide variety of species, leading researchers to hypothesize that "this is mechanically the most effective solution to the problem of rapid movement over rigid substrates by soft-bodied animals" (Trueman 1975, p. 24). While this may be an overstatement, the striking prevalence of this locomotive strategy provides ample motivation to study the mechanics of a sheet swimming (or crawling) over a thin layer of fluid.

In addition, there have been numerous recent studies involving mechanical swimmers and crawlers that mimic the waving motions observed in biological systems, including self-propelled liquid-crystal elastomers (LCEs) (Camacho-Lopez *et al.* 2004) and RoboSnail, a mechanical crawler that mimics the locomotive strategies employed by selected gastropods (Chan, Balmforth & Hosoi 2005). In the first case, when a small LCE disk at an air/water interface is triggered with a laser, the elastomer will 'swim into the dark' by propagating a saddle-shaped wave along its flexible body.

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

In the second example, RoboSnail propels itself by mechanically actuating its foot. In both cases, a flexible body swims or crawls over a thin layer of fluid near a rigid boundary using a geometrical configuration similar to that shown in figure 1. It has been well-documented that, in general, the presence of a rigid boundary can dramatically affect the behaviour of low-Reynolds-number swimmers; e.g. bacteria swimming near a rigid boundary tend to swim in circles (rather that in a random walk) due to hydrodynamic interactions with the wall (DiLuzio *et al.* 2005; Lauga *et al.* 2006). It has even been suggested that the interaction of rigid boundaries with low-Reynolds-number propulsion systems may be relevant to the early stages of biofilm formation and pathogenic infection (DiLuzio *et al.* 2005).

Swimming at low Reynolds numbers – in the absence of rigid boundaries and substrates – has long been a topic of interest, particularly in the context of locomotive strategies of micro-organisms such as spermatozoa and nematodes. One of the earliest studies of this subject was G. I. Taylor's work on the motion of a two-dimensional swimming flapping sheet (Taylor 1951), in which he analysed the swimming speed of such an organism in the limit of low-amplitude flapping oscillations. Shortly after Taylor's pioneering investigation, Hancock (1953) proposed an alternative approach, calculating swimming speeds by modelling the swimmer's tail with a distribution of Stokeslets and doublets. As Hancock's approach is more general than Taylor's, he was able to investigate a variety of geometries including finite and infinite tails as well as several steady shape disturbances including helical motions, transverse waves and longitudinal waves. These theoretical predictions were then successfully compared with experimentally measured swimming velocities of nematodes. Numerous subsequent analyses are discussed in Lighthill (1976), Childress (1981), and Pedley & Kessler (1992).

While the mechanics of free low-Reynolds-number swimmers has been analysed extensively, far fewer studies exist in which the optimization of such systems has been investigated. Shapere & Wilczek (1989) study the efficiency of swimming motions due to small deformations of spherical and cylindrical bodies at low Reynolds number. They define a natural notion of efficiency and show how to find optimally efficient swimming strokes composed of propagating waves, symmetric about the axis of propulsion and limited to a finite range of spherical harmonics or Fourier modes. Their work is intended to model ciliated organisms that swim by synchronously waving a layer of short, densely packed cilia. Becker, Koehler & Stone (2003) study optimal swimming strokes for Purcell's-three link swimmer using slender-body dynamics in the inertialess limit and compare its efficiency (0.77%) to the optimal undulating rod (7.4%) and rotating helix (8.1%). The simplicity of Purcell's swimmer may prove useful in the design of micro-robots for minimally invasive therapeutic treatments, and is a useful model to explore fundamental aspects of self-propulsion at low Reynolds numbers. Avron, Gat & Kenneth (2004) study a two-dimensional microbot that swims by changing the shape of its boundary. To rank different swimming strokes, they introduce a swimming drag coefficient, which is equivalent to Shapere & Wilczek's efficiency for small-amplitude perturbations of a cylinder, but is also applicable to swimmers whose shape changes substantially during the swimming stroke. They go on to find the optimal swimmer within a certain (four-dimensional) space of swimming strokes using conformal mapping techniques. Felderhof (2006) derives expressions for swimming velocity and mean dissipation in various asymptotic regimes for a chain of beads (one large bead followed by several smaller beads) swimming in a Stokesian fluid. Felderhof does not discuss optimization, but his formulae could easily be optimized. Lobaton & Bayen (2007) model a bacterial flagellum as a helical

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

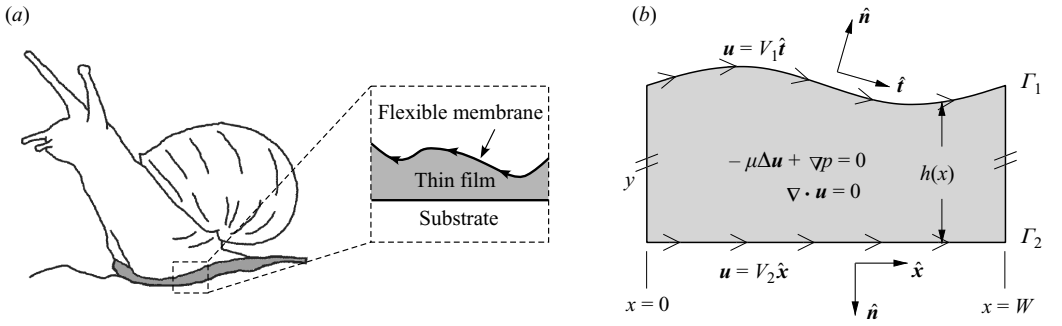


FIGURE 1. (*a*) Schematic of gastropod locomotion showing a thin viscous fluid sandwiched between a periodically deforming flexible foot and a rigid substrate. (b) Idealized model depicting steady state swimming of a sheet via wave propagation. This system is conveniently studied in the reference frame in which the wave profile Γ_1 remains stationary. In this frame, material points on the sheet move tangent to Γ_1 with constant speed V_1 and the substrate moves with constant speed V_2.

structure attached to a spheroid. They use the method of regularized Stokeslets to discretize the system into particles that are connected by spring forces, and derive an optimal elasticity distribution to maximize forward speed when the base of the flagellum is rotated. They use an adjoint technique (commonly used in PDE constrained optimization) to compute the gradient of the objective function with respect to perturbation of the spring coefficients.

In the current paper we will consider optimal geometries for a variant of Taylor's swimming sheet near a rigid boundary. The quantities we will optimize are swimming speed, mechanical efficiency, and fluid loss (which is relevant for gastropod locomotion). To our knowledge, this is the first study of optimal swimming in which the solution is sought from the full infinite-dimensional space of curves (without limiting ourselves to a few Fourier modes or discretizing the curve into beads or rigid segments). In other words, we optimize first and discretize second rather than the other way around. In § 2 we define the mathematical model, calculate the swimming velocity and power dissipation in the lubrication limit, and briefly discuss the forces at work in this type of propulsion. In § 3, we formulate several optimization problems and derive the relevant Euler–Lagrange equations, which turn out to be integro-differential equations for the shape of the optimal wave profile. In § 4, we present two numerical methods to solve the optimization equations and use them to study the solutions in various regimes of the parameter space in which the optimal solutions become singular in a self-similar fashion and sometimes leave the realm of validity of the lubrication model. We explore these singular limits by computing higher-order corrections to the zeroth-order theory and find that wave profiles that develop cusp-like singularities are appropriately penalized, yielding non-singular optimal solutions. These corrections are validated in Appendix C by comparing them to solutions of the full Stokes equations using a blend of finite element and rigorous analysis. We end with a discussion of our results in § 5, which includes a one page summary of our main findings.

2. Mathematical model

Our crawling 'organism' will be modelled as a flexible sheet swimming over a viscous Newtonian fluid near a rigid substrate as shown in figure 1. (While this model

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

is not ideal for real snails, it is directly relevant to the design of mechanical crawlers in which the mucus simulant may be chosen to be Newtonian (Chan *et al.* 2005)). In the simplest case, we can assume the foot shape is periodic and completely determined by the muscular input; thus we assume the organism can select a deformation that is most effective for a given environment. The wave profile and the fluid velocity are assumed to be periodic in the x-direction with period W. At steady state, there exists a reference frame (the wave frame) in which the shape of the foot – as described by the wave envelope $\Gamma_1 = \{(x, h(x))\}$ – remains stationary in time. In this frame, the substrate moves with constant speed V_2 in the x-direction while the membrane moves tangent to Γ_1 with constant speed V_1 (see figure 1*b*); thus, we consider the case where the sheet remains inextensible as it swims. Although the foot of a real gastropod can stretch in the tangential direction, other swimmers such as water snakes have flexible backbones which allow changes in curvature but not length. In addition, our mechanical crawlers such as RoboSnail are designed with an inextensible foot. Our goal is to find optimal shapes that maximize speed and efficiency while minimizing arclength and fluid production costs. Since the optimal shapes turn out to be smooth, single-valued functions, it does not appear necessary to pose the problem more generally to allow the wave profile to overturn.

2.1. *Swimming speed and power dissipation*

Two quantities that will be used repeatedly in our analysis are the swimming speed and power required to swim, which we derive here in their general form for Stokes flow. Consider the Stokes equations

$$\nabla p = \mu \nabla^2 \boldsymbol{u} \qquad \nabla \cdot \boldsymbol{u} = 0, \tag{2.1}$$

where p denotes pressure, μ is the dynamic viscosity, and $\boldsymbol{u}$ is the velocity of the fluid. These equations must be solved subject to the boundary conditions $\boldsymbol{u} = V_1 \hat{\boldsymbol{t}}$ on the upper boundary, Γ_1, and $\boldsymbol{u} = V_2 \hat{\boldsymbol{x}}$ on the lower boundary, Γ_2, where $\hat{\boldsymbol{t}}$ and $\hat{\boldsymbol{x}}$ are unit vectors tangent to Γ_1 and Γ_2, respectively, oriented as shown in figure 1. We define

$$P(V_1, V_2) = \int_{\Gamma_1} \hat{\boldsymbol{t}} \cdot \boldsymbol{\sigma} \hat{\boldsymbol{n}} \, \mathrm{d}s, \qquad F(V_1, V_2) = -\int_{\Gamma_1} \hat{\boldsymbol{x}} \cdot \boldsymbol{\sigma} \hat{\boldsymbol{n}} \, \mathrm{d}s = \int_{\Gamma_2} \hat{\boldsymbol{x}} \cdot \boldsymbol{\sigma} \hat{\boldsymbol{n}} \, \mathrm{d}x, \tag{2.2}$$

where $\boldsymbol{\sigma} = -p\boldsymbol{I} + \mu[\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^T]$ is the stress tensor and $\hat{\boldsymbol{n}}$ denotes the outward unit normal to Γ_1 and Γ_2. Physically, $V_1 P$ (rather than P itself) is the power required to maintain the steady motion of the sheet, F is the x-component of the net force exerted by the fluid on the sheet (which is the same as that exerted by the substrate on the fluid), and $V_2 F$ is the power required to maintain the steady motion of the substrate. We emphasize that there is a solution of the Stokes equations with these boundary conditions for any choice of V_1 and V_2, but most choices require non-zero *external* forces $\mp F(V_1, V_2)$ to be applied to the system from above and below. To model a self-propelled swimmer (with no external forces), we require V_1 and V_2 to be related so that $F(V_1, V_2) = 0$. Since the Stokes equations are linear, $F(V_1, V_2) = V_1 F(1, 0) + V_2 F(0, 1)$, and we obtain the relationship

$$V_2 = -\frac{F_1}{F_2} V_1, \tag{2.3}$$

where we have defined $F_1 = F(1, 0)$ and $F_2 = F(0, 1)$ to be the external forces required to drive two auxiliary problems with V_1 and V_2 alternately set to 1 or 0. Note that V_1 is directly controlled by the swimmer: $(W/L)V_1$ is the speed at which the sheet propagates waves to the left (relative to its material points) in order to swim to the

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

right. From (2.3), we find that the swimming speed in the laboratory frame is given by

$$V = \frac{W}{L}V_1 - V_2 = \left(\frac{W}{L} + \frac{F_1}{F_2}\right)V_1, \tag{2.4}$$

which relates the self-propelled swimmer's speed V to the input wave speed V_1.

It is useful to derive expressions for P and F in terms of the velocity field $\boldsymbol{u} = (u, v)$. Details of this calculation are given in Appendix A. In brief, we substitute the Newtonian stress tensor into (2.2), enforce continuity, and apply the no-slip boundary conditions to obtain

$$P(V_1, V_2) = \int_{\Gamma_1} \mu u u_y \left(1 + h_x^2\right)^{3/2} \mathrm{d}x, \tag{2.5}$$

$$F(V_1, V_2) = -\int_{\Gamma_1} \left[p h_x + \mu u_y \left(1 + h_x^2\right)\right] \mathrm{d}x = -\int_{\Gamma_2} \mu u u_y \, \mathrm{d}x, \tag{2.6}$$

where subscripts denote partial derivatives. These formulae will be used to compute higher-order corrections to the swimming speed and power dissipation obtained from lubrication theory.

2.2. *The lubrication approximation*

We now briefly review the work of Chan *et al.* (2005) in which the swimming speed of the sheet is computed in the lubrication limit (Oron, Davis & Bankoff 1997). The validity and limitations, of this approximation are discussed further in Appendix C.

To study this small-amplitude limit (where $h \ll W$), we non-dimensionalize the problem, choosing different scales for the x- and y-coordinates to keep the rescaled aspect ratio of the domain of order unity. Specifically, we choose a characteristic height H (e.g. the maximum film thickness) and velocity U (e.g. the wave speed) and set

$$\hat{x} = Wx, \quad \hat{y} = Hy, \quad \hat{h}(\hat{x}) = Hh(x), \quad \hat{V}_i = UV_i,$$
$$\hat{\boldsymbol{u}} = (\hat{u}, \hat{v}) = U(u, Hv/W), \quad \hat{p} = \frac{\mu U W}{H^2} p,$$

where, henceforth, the carat distinguishes a physical variable from its dimensionless counterpart. Dropping terms of order $\varepsilon = H/W$, the momentum equations (2.1) become $p_x = u_{yy}$ and $p_y = 0$. To first order in ε, the boundary conditions on u are $u(x, h(x)) = V_1$, $u(x, 0) = V_2$; thus, the fluid velocity is given by

$$u(x, y) = \frac{y}{2}(y - h)p_x + \frac{y}{h}V_1 + \left(1 - \frac{y}{h}\right)V_2. \tag{2.7}$$

Integrating from 0 to h and solving for p_x, we obtain

$$p_x = \frac{6}{h^2}(V_1 + V_2) - \frac{12}{h^3}Q, \tag{2.8}$$

where $Q = \int_0^h u \, \mathrm{d}y$ is the (constant) volume flux through any cross-section of the fluid. Since p is periodic, $\int_0^1 p_x \, \mathrm{d}x = 0$ and we find that

$$Q = \frac{V_1 + V_2}{2}\frac{I_2}{I_3}, \qquad \text{where} \qquad I_k \equiv \int_0^1 h(x)^{-k} \, \mathrm{d}x. \tag{2.9}$$

Finally, the force and power functions defined in the previous section can be computed in the lubrication limit. The terms involving $\hat{h}_{\hat{x}}$ in (2.5) and (2.6) vanish as they are

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

all higher order. Rescaling

$$\hat{P} = \mu U \frac{W}{H} P, \quad \hat{F} = \mu U \frac{W}{H} F,$$

substituting u from above and dropping higher-order terms, we obtain

$$P = \left(4I_1 - 3\frac{I_2^2}{I_3}\right) V_1 + \left(2I_1 - 3\frac{I_2^2}{I_3}\right) V_2, \tag{2.10}$$

$$F = \left(2I_1 - 3\frac{I_2^2}{I_3}\right) V_1 + \left(4I_1 - 3\frac{I_2^2}{I_3}\right) V_2. \tag{2.11}$$

It follows from (2.3) and (2.4) that

$$V_2 = -\frac{2-3\zeta}{4-3\zeta} V_1, \qquad V = \frac{6(1-\zeta)}{4-3\zeta} V_1, \tag{2.12}$$

where the shape parameter, ζ, is defined as

$$\zeta \equiv \frac{I_2^2}{I_1 I_3}. \tag{2.13}$$

Finally, we use (2.10) and (2.12) to obtain the power that must be generated by the sheet in order to swim at steady state:

$$\text{power} = V_1 P = 2 I_1 V_1 V. \tag{2.14}$$

A natural choice for the characteristic speed U is the wave speed, $\hat{V}_1$ – which is directly controlled by the organism — in which case $V_1 = 1$. Note that we can put bounds on the achievable swimming velocities using the Cauchy–Schwarz inequality:

$$I_2^2 = \left(\int h^{-1/2} h^{-3/2}\, \mathrm{d}x\right)^2 \leqslant \int h^{-1}\, \mathrm{d}x \int h^{-3}\, \mathrm{d}x = I_1 I_3. \tag{2.15}$$

Hence $0 < \zeta \leqslant 1$ and, from (2.12), the dimensionless velocity satisfies $0 \leqslant V < 3/2$. We assume from now on that $h(x)$ is not a constant function, which implies $\zeta < 1$ and $V > 0$.

2.3. *Speedy shapes*

The observation by Chan *et al.* (2005) that the organism can swim faster than the wave speed, i.e. that $V > V_1$ is possible, is surprising. This result can be explained as follows. First, the swimming speed is determined by a force balance calculation. For both the full Stokes equations and the lubrication approximation, this balance of forces can be reduced to solving two externally driven auxiliary problems (with V_1 and V_2 alternately set to 1 and 0) resulting in equation (2.4), namely $V = (W/L + F_1/F_2)V_1$. In the small-amplitude limit, $W/L \approx 1$ and the signs of F_1 and F_2 determine whether the sheet swims faster or slower than the wave speed.

Now, $F_2 = F(0, 1)$ is positive because it can be interpreted as a power dissipation– in this auxiliary problem the wall is externally driven at unit speed $V_2 = 1$ with the top sheet held fixed, and $F_2 = V_2 F_2$ is the power required to maintain this motion. On the other hand, $F_1 = F(1, 0)$ is the force exerted by the wall on the fluid when the sheet is driven at unit speed and the wall is held fixed. If Γ_1 is flat enough, the solution of this auxiliary problem will resemble a shear flow with $u_y > 0$ everywhere; the net force $F_1 = -\int_{\Gamma_2} u_y\, \mathrm{d}x$ will then be negative. But if Γ_1 passes close to Γ_2 and creates a narrow constriction, there will be a large recirculation region where

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

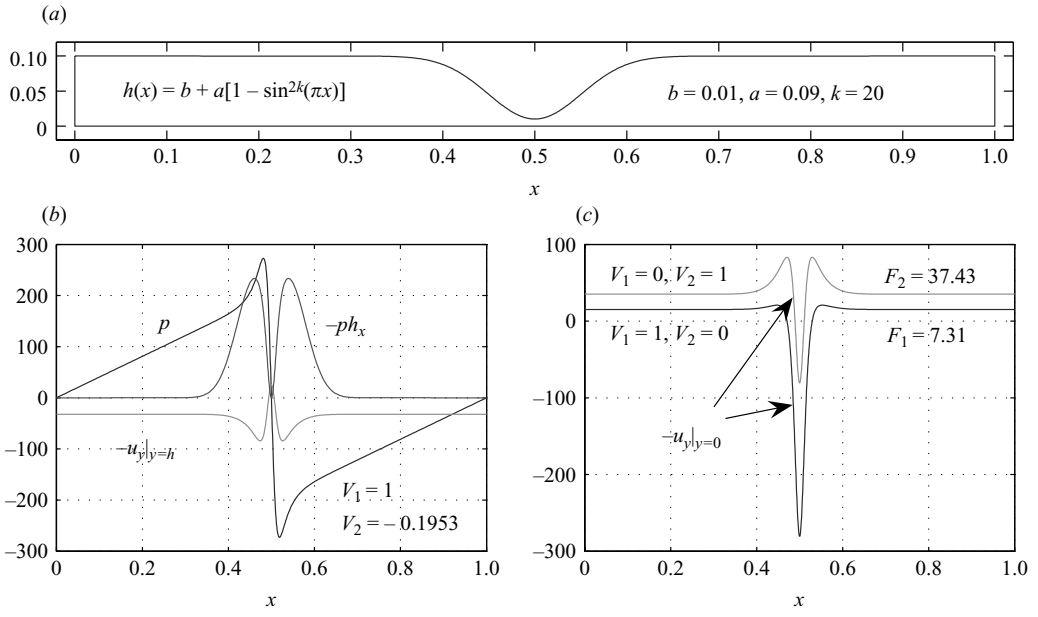


FIGURE 2. (*a*) A sample wave profile with a narrow constriction. (*b*) Pressure and drag forces: a large pressure drop across the fluid gap pushes the self-propelled sheet forward against the viscous shear forces acting on the sheet. (*c*) Forces in auxiliary problems: the equilibrium condition $V_2 = -(F_1/F_2)V_1$ is found by computing the shear forces F_1, F_2 of the bottom wall on the fluid for two non-equilibrium auxiliary problems.

the fluid flow near Γ_2 is actually right to left (even though the sheet moves left to right). In this region, $u_y < 0$ along Γ_2 and the contribution to F_1 is positive – only in the narrow gap is $u_y > 0$ and the contribution negative. This is illustrated in figure 2, where the positive contribution to F_1 in the 'recirculation region' is stronger than the negative contribution in the 'gap region', leading to a swimming speed of $V/V_1 = 1 + F_1/F_2 = 1.1954 > 1$.

The fact that narrow constrictions lead to faster swimming is not surprising. Consider the fluid forces acting on the sheet in the x-direction from the the combined velocity field (rather than on the wall for each auxiliary problem separately). Upstream of the constriction there is an increase in pressure as fluid is squeezed through the narrow gap. Conversely, the pressure decreases on the downstream side as the channel expands. These pressures act normal to the sheet resulting in a net propulsive force on the organism. At steady state, these propulsive forces are balanced by drag. As shown in figure 2(*b*), the net contribution of pressure $(-ph_x)$ is positive and is exactly balanced by the shear forces $(-u_y|_{y=h})$. Maximizing the efficacy of this propulsive strategy requires: (i) a localized narrow constriction to increase the pressure drop and (ii) a (relatively) large film thickness between these constrictions to minimize viscous drag. As we shall see in the following analyses, both of these features are present in all of the optimized wave forms.

There are many families of curves for which the swimming speed V (with $V_1 = 1$) approaches its maximal value of 3/2. We observe that $V = 6(1-\zeta)/(4-3\zeta)$ decreases monotonically from 3/2 to zero as $\zeta = I_2^2/I_1 I_3$ varies from zero to one. A necessary (but not sufficient) condition for ζ to approach zero is that $h_{\min}/h_{\max}$ approaches zero, where $h_{\min}$ and $h_{\max}$ are the minimum and maximum values attained by $h(x)$.

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

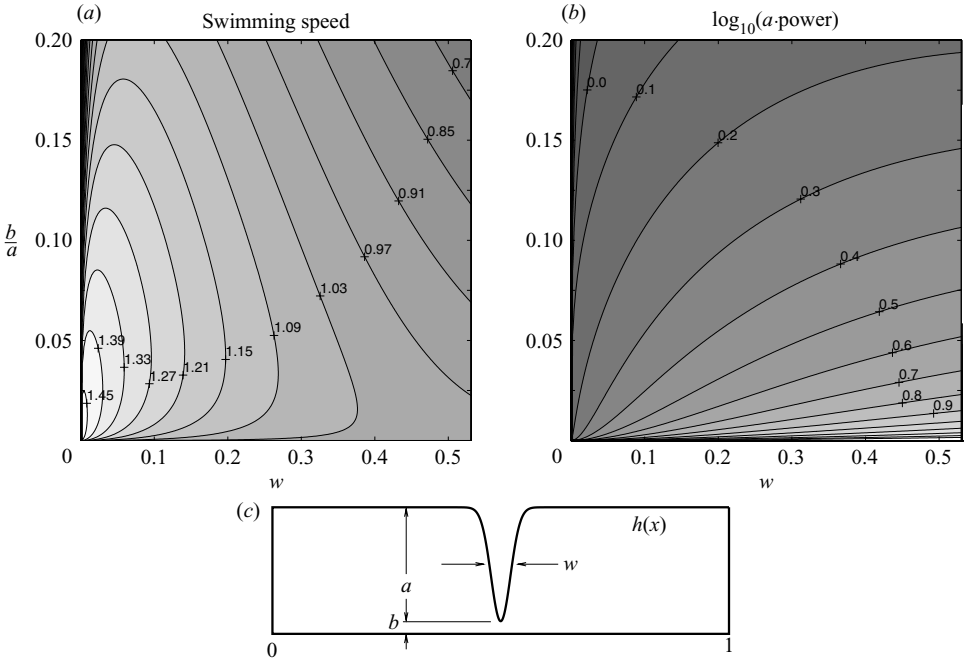


FIGURE 3. Contour plots of (a) swimming speed V and (b) power P as functions of gap thickness b and ripple width w for the wave profile $h(x) = b + a[1 - (\sin^2(\pi x))^{4\ln 2/\pi^2 w^2}]$ shown in (c). If $w > 0$ is held fixed, $\zeta \to 2/3$, $V \to 1$ and $P \to \infty$ as $b \to 0$. If $b > 0$ is held fixed, $V \to 0$ and $P \to 0$ as $w \to 0$. If the ratio b/w is held fixed, $V \to 3/2$, $I_2 \to \infty$, $I_3 \to \infty$, and P and I_1 remain bounded as $b, w \to 0$.

To see this, note that $I_2 = \int_0^1 h^{-2}\,\mathrm{d}x \geqslant h_{\max}^{-2}$ and $I_1 I_3 \leqslant h_{\min}^{-4}$; hence, $\zeta \geqslant h_{\min}^4/h_{\max}^4$. If $h_{\min}/h_{\max}$ does approach zero, then I_1, I_2 and I_3 may diverge and the limiting value of ζ – in particular, whether ζ approaches zero – depends on the relative rates at which they do so.

To illustrate this, consider the family of curves shown in figure 3. For this particular family,

$$h(x) = b + a\left[1 - (\sin^2(\pi x))^{4\ln 2/(\pi^2 w^2)}\right], \tag{2.16}$$

the only way to approach a swimming speed of 3/2 is to simultaneously decrease the ratio $h_{\min}/h_{\max} = b/(a+b)$ and the spike width w to zero (so that the maximum curvature $\kappa = h''(1/2) = (8a\ln 2)w^{-2}$ diverges). If the ratio b/w is held fixed in this process (so that the origin in both plots is approached along a straight line), then I_2 and I_3 both diverge while I_1 remains bounded and $\zeta = I_2^2/I_1 I_3$ approaches zero. Interestingly, there are other families of curves for which the maximum curvature remains bounded as $V \to 3/2$. For example, the one-parameter family $h(x) = \sqrt{\cos^6 \pi x + b^2 \cos^2 \pi x + b^4}$ yields $\zeta \to 0$ and $\kappa \to \pi^2$ as $b \to 0$. This shows that one cannot characterize the swimming behaviour of the sheet solely in terms of its local behaviour near its minima. In §4, we will see that the optimal curves $h(x)$ approach touchdown in a self-similar fashion as a parameter involving the constraints approaches a critical value.

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

3. Optimization: the Euler–Lagrange equations

We now define several natural optimization problems in the context of swimming sheets that will then be solved numerically in § 4. First, for a sheet of fixed length swimming over a fixed volume of fluid, we wish to

(1) maximize the swimming speed,
(2) maximize the efficiency,
(3) maximize the efficiency for a given swimming speed,
(3a) same as (3), but without the fluid volume constraint.

The third problem may also be posed as that of minimizing arclength subject to swimming speed, efficiency and fluid volume constraints — in other words, among all curves that achieve these desired goals, find the smoothest one. In the spirit of gastropod locomotion, where the animal must produce the fluid in the thin film and has limited resources for doing so, we will then replace the fluid volume constraint by a constraint on the 'fluid left behind' per unit distance travelled, namely

(4) minimize arclength for a given swimming speed, efficiency and fluid loss.

In this way, holding fluid loss constant, all the curves in the two-parameter family (with speed and efficiency as parameters) have the same 'cost' in terms of the energy associated with mucus production, which is important for efficiency comparisons to be meaningful.

The four problems considered here are intended to strike a balance between mathematical appeal and biological relevance. In order to solve them, we face obstacles that would arise for any choice of objective function or constraints: non-existence of solutions, multiple local extrema, complicated parameter spaces, and optimal solutions that leave the realm of validity of the (lubrication) model. Although Problem 4 takes into account the cost of mucus production, it remains idealized in our neglect of viscoelastic effects, finite length effects, three-dimensional effects, tangential stretching of the sheet, and muscular limitations (e.g. in terms of maximum sustainable pressure and shear forces). Nevertheless, we will start by investigating these illustrative model problems with the understanding that the methods we develop could be extended to include other objectives and restrictions.

3.1. *Problem 1 (maximize speed)*

In the first problem under consideration, we wish to find the wave profile $h(x)$ that maximizes the swimming speed V subject to the constraints

$$\ell^2[h] = \ell_0^2, \qquad A[h] = A_0, \tag{3.1}$$

where ℓ_0 and A_0 are fixed constants while

$$\ell^2[h] \equiv \frac{1}{2}\int_0^1 h_x^2 \,\mathrm{d}x, \qquad A[h] \equiv \int_0^1 h \,\mathrm{d}x \tag{3.2}$$

are functionals of the curve shape $h(x)$ representing arclength and area (i.e. fluid volume), respectively. Since V is a monotonically decreasing function of ζ, maximizing V is equivalent to minimizing ζ. Although the lubrication approximation includes only zeroth-order terms in an expansion in powers of $\varepsilon = H/W$, the arclength $\hat{L} = \int_0^W (1 + \hat{h}_{\hat{x}}^2)^{1/2}\,\mathrm{d}\hat{x}$ must be treated at second order to obtain zeroth-order results for swimming speed and power. Therefore, we expand $\hat{L} \approx W(1+\varepsilon^2\ell^2)$ and drop higher-order terms. A scaling argument (presented below) provides further justification for doing this.

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

For a general functional $f[h]$, the variational derivative $\delta f/\delta h$ is defined via

$$Df[h]g = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f[h+tg] = \int_0^1 \frac{\delta f}{\delta h}(x)g(x)\,\mathrm{d}x. \tag{3.3}$$

In particular,

$$\frac{\delta I_k}{\delta h} = -kh^{-k-1}, \quad \frac{\delta \zeta}{\delta h} = \left(\frac{1}{I_1h^2} - \frac{4}{I_2h^3} + \frac{3}{I_3h^4}\right)\zeta, \quad \frac{\delta \ell^2}{\delta h} = -h_{xx}, \quad \frac{\delta A}{\delta h} = 1. \tag{3.4}$$

Assuming an optimal solution exists which is at least C^2, it must satisfy the equation

$$\frac{\delta \zeta}{\delta h} + \lambda_1 \frac{\delta \ell^2}{\delta h} + \lambda_2 \frac{\delta A}{\delta h} = 0$$

for some Lagrange multipliers λ_1, λ_2. For this problem, it is possible to compute λ_1 and λ_2 analytically. To this end, consider the new objective function

$$G[h] = \zeta[\alpha h + \beta] + (\alpha - 1)^2/2 + \beta^2/2, \tag{3.5}$$

where

$$\alpha = \ell_0 \left(\frac{1}{2}\int_0^1 h_x^2\,\mathrm{d}x\right)^{-1/2}, \qquad \beta = A_0 - \alpha \int_0^1 h\,\mathrm{d}x \tag{3.6}$$

are chosen so that $\alpha h + \beta$ satisfies the constraints. Any global minimizer of G will be a constrained minimizer of ζ and satisfy $\alpha = 1$, $\beta = 0$. One can readily check that

$$\frac{\delta G}{\delta h} = \alpha \frac{\delta \zeta}{\delta h} + \left(\int_0^1 \frac{\delta \zeta}{\delta h} h\,\mathrm{d}x\right)\frac{\delta \alpha}{\delta h} + \left(\int_0^1 \frac{\delta \zeta}{\delta h}\,\mathrm{d}x\right)\frac{\delta \beta}{\delta h} + (\alpha - 1)\frac{\delta \alpha}{\delta h} + \beta \frac{\delta \beta}{\delta h}, \tag{3.7}$$

where $\delta\zeta/\delta h$ is evaluated at $\alpha h + \beta$ in each of the integrals and

$$\frac{\delta \alpha}{\delta h} = \frac{\alpha^3}{2\ell_0^2} h_{xx}, \qquad \frac{\delta \beta}{\delta h} = -\left(\int_0^1 h\,\mathrm{d}x\right)\frac{\delta \alpha}{\delta h} - \alpha. \tag{3.8}$$

Setting $\alpha = 1$, $\beta = 0$, $\delta G/\delta h = 0$ we obtain

$$\frac{\delta \zeta}{\delta h} - \left(\int_0^1 \frac{\delta \zeta}{\delta h}\,\mathrm{d}x\right)\left(\frac{A_0}{2\ell_0^2} h_{xx} + 1\right) = 0,$$

which leads to the integro-differential equation

$$h_{xx} = \left(\frac{2\ell_0^2}{A_0}\right)\left(\frac{h^{-2}/I_1 - 4h^{-3}/I_2 + 3h^{-4}/I_3}{I_2/I_1 - 4I_3/I_2 + 3I_4/I_3} - 1\right). \tag{3.9}$$

Note that solutions of this equation will be symmetric about their extrema due to the absence of a term involving h_x. To see this, imagine we are given a solution $h(x)$. Then the initial value problem emanating from an extremum (x^*, h_0) with fixed parameters I_k has a unique solution with initial condition $h(x^*) = h_0$, $h'(x^*) = 0$. But the function $h(2x^* - x)$ also solves this initial value problem, so $h(2x^* - x) = h(x)$ for all x.

Although there are two parameters in the problem statement, ℓ_0 and A_0, we can simplify the problem by showing that the optimal solution depends only on the ratio A_0/ℓ_0. Indeed, if we scale ℓ_0 and A_0 by the same factor θ, the optimal solution $h(x)$ merely scales by θ as well, and the optimal speed remains the same. This is clear from the Euler–Lagrange equations (3.9) and the formula $\zeta = I_2^2/I_1I_3$, where we observe that if h is replaced by θh then I_k will become $\theta^{-k}I_k$. This is to be expected from the formulation in the lubrication limit since the analysis is intended to

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

hold independently of the small parameter H/W, and the physical problem remains unchanged if we replace h by θh and H by $\theta^{-1}H$. Therefore, there is only one essential parameter and we seek a family of optimal solutions depending on the ratio A_0/ℓ_0.

When the objective function involves only the swimming speed without regard to power consumption, the sheet will attempt to pass near the substrate to create a narrow fluid gap (see § 2.3 above). The problem becomes singular if the area and arclength constraints allow the sheet to reach the wall – there will then be infinitely many solutions with speeds arbitrarily close to 3/2 but none that equals it. (Recall that we are not interested in the development of dry patches; hence, we only consider wave profiles satisfying $h(x) > 0$ everywhere.) To determine the allowable values of ℓ_0 and A_0 (the values that lead to a well-posed optimization problem), consider the non-allowable curve satisfying $h(0) = h(1) = 0$ that maximizes area subject to the constraint $\ell^2[h] = \ell_0^2$. A calculation shows that the solution is $h_{\text{crit}}(x) = \sqrt{6}\ell_0(x - x^2)$, which has (dimensionless) area $A_{\text{crit}} = \ell_0/\sqrt{6}$. If $A_0 > A_{\text{crit}}$, any periodic function $h(x)$ that satisfies the constraints (3.1) is strictly positive. In § 4.2 we will study the asymptotics of the optimal solution as A_0 approaches this critical value. These solutions are self-similar in a neighbourhood of the point closest to touchdown and approach $h_{\text{crit}}(x)$ outside this neighbourhood, where the complicated term involving the I_k on the right-hand side in (3.9) becomes negligible. In this limit, swimming velocities approach the optimal speed $V = 3/2$ while the power required to sustain these speeds grows without bound.

Note again that our system is limited in that we explicitly prescribe the kinematics of the sheet rather than the internal forces. Naturally, in real crawlers, these kinematics are dictated by the dynamics, which include the influences of geometry, material properties, muscular limitations, etc. In particular, for a real organism, we would not expect to observe sharp, cusp-like features as there is an energetic cost associated with bending that we do not account for here. One way to correct for this would be to include a penalty on regions of large curvature in the objective function – e.g. one could add a term that includes the square of the local curvature multiplied by a small parameter. However, in order to gain insight into the simplest system first, we begin by tackling the problem as defined above without this penalty, i.e. we rely only on arclength as a regularizer to penalize discontinuities and infinite slopes in the wave profile $h(x)$.

3.2. *Problem 2 (maximize efficiency)*

In this second problem, we wish to find the wave profile $h(x)$ that maximizes the ratio of speed to power (i.e. the "efficiency") subject to the constraints (3.1). From (2.14), we see that $V/P = 1/(2I_1)$; thus, to maximize this ratio, we must minimize I_1. Taking the same approach as before, we set

$$G[h] = I_1[\alpha h + \beta] + (\alpha - 1)^2/2 + \beta^2/2 \tag{3.10}$$

and compute

$$\frac{\delta G}{\delta h} = \alpha \frac{\delta I_1}{\delta h} + \left(\int_0^1 \frac{\delta I_1}{\delta h} h \, \mathrm{d}x \right) \frac{\delta \alpha}{\delta h} + \left(\int_0^1 \frac{\delta I_1}{\delta h} \, \mathrm{d}x \right) \frac{\delta \beta}{\delta h} + (\alpha - 1)\frac{\delta \alpha}{\delta h} + \beta \frac{\delta \beta}{\delta h}, \tag{3.11}$$

where $\delta I_1/\delta h$ is evaluated at $\alpha h + \beta$ in each of the integrals. Setting $\alpha = 1$, $\beta = 0$, $\delta G/\delta h = 0$ we obtain

$$\frac{\delta I_1}{\delta h} - I_1 \frac{\delta \alpha}{\delta h} + I_2 \left(A_0 \frac{\delta \alpha}{\delta h} + 1 \right) = 0,$$

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

i.e.

$$h_{xx} = 2\ell_0^2 \frac{h^{-2} - I_2}{I_2 A_0 - I_1}. \tag{3.12}$$

Interestingly, we will see that solutions of this equation maximize I_1 rather than minimize it. In other words, there are well-behaved functions h which are optimally *in*efficient, but one cannot achieve the theoretical optimal (dimensionless) efficiency of $V/P = A_0/2$. Moreover, as solutions approach this optimal efficiency, the swimming speed goes to zero; therefore, it is more useful to reformulate the problem as follows.

3.3. *Problem 3 (maximize efficiency for a given swimming speed)*

Given an attainable speed V_0, we wish to find the wave profile $h(x)$ that minimizes the power required to swim at that speed, subject to the constraints (3.1). Because the speed constraint is non-trivial, we are unable to obtain analytic expressions for the Lagrange multipliers, and instead leave the Euler–Lagrange equation

$$\frac{\delta \ell^2}{\delta h} - \lambda_1 \frac{\delta I_1}{\delta h} + \frac{\lambda_2}{\zeta}\frac{\delta \zeta}{\delta h} + \lambda_3 \frac{\delta A}{\delta h} = 0$$

in the form

$$h_{xx} = \frac{\lambda_1}{h^2} + \lambda_2 \left(\frac{1}{I_1 h^2} - \frac{4}{I_2 h^3} + \frac{3}{I_3 h^4} \right) + \lambda_3. \tag{3.13}$$

As in Problems 1 and 2 above, solutions of Problem 3 depend only on the (dimensionless) ratio A_0/ℓ_0 in the sense that scaling A_0 and ℓ_0 by the same constant θ causes the optimal wave profile $h(x)$ (and the optimal efficiency) to scale by θ as well, but does not affect the swimming speed. Therefore, although there are three constraints, only two parameters change the solution in an essential way, and the third can be absorbed into the vertical length scale H of the physical problem.

In § 4.2, we will study this two-parameter family from various perspectives. First, holding ℓ_0 constant, we will drop the constraint on $A[h]$ and obtain a one-parameter family of curves ("Problem 3a") depending only on V_0. We will then hold ℓ_0 fixed and vary A_0 and V_0 over their attainable values to study the connection between Problems 1, 2 and 3a. For choices of A_0 and V_0 near the boundary of attainable values (found by solving Problem 1 for $V_{\max}$ in terms of A_0), there are two solutions of (3.13), one maximizing the efficiency and the other minimizing it. For choices of A_0 and V_0 far inside the attainable region, other local extrema also exist, but they can be interpreted as higher-frequency versions of these two primary families of extremal wave profiles.

The arclength and fluid volume constraints are appealing for their simplicity, but require re-interpretation for biological relevance. The role of arclength is to regularize the problem while the role of fluid volume is to set the vertical length scale: by imposing the constraint $A[h] = A_0$, we remove the indeterminacy of scaling h by θ and H by θ^{-1} without changing the physical problem being modelled. To explore this point of view, we will exchange the roles of efficiency and arclength, treating the former as a constraint and the latter as the objective function. The quantities of physical interest (speed, efficiency and fluid volume) are now imposed as constraints and the optimization problem is to find the smoothest wave profile that can achieve these goals. This has the added benefit that it eliminates the difficulty that (3.13) has multiple solutions when optimizing efficiency: these solutions are mapped to unique locations in the new parameter space with arclength as the objective function.

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

3.4. *Problem 4 (minimize arclength given swimming speed, efficiency and fluid loss)*

As a final variant of Problem 3, we replace the fluid volume constraint with a constraint on the fluid loss per unit distance travelled. The distance travelled in the laboratory frame in time $T = W/\hat{V}_1$ is $\hat{V}T$ while the volume of fluid that has passed under a vertical cross-section attached to a material point on the sheet during this cycle is $\hat{A} - \hat{Q}T$:

$$\hat{Z} = \frac{\hat{A} - \hat{Q}T}{\hat{V}T} = \frac{HWA - (HUQ)(W/U)}{(UV)(W/U)} = \frac{A - Q}{V}H = ZH. \tag{3.14}$$

Thus, we replace the constraint $A[h] = A_0$ in Problem 3 above with

$$Z[h] = Z_0, \qquad Z[h] \equiv \frac{A[h] - Q[h]}{V[h]}, \qquad Q[h] \equiv \left(\frac{1}{4 - 3\zeta[h]}\right)\frac{I_2[h]}{I_3[h]}, \tag{3.15}$$

where Z_0 is given. We freeze $V[h] = V_0$ and $\zeta[h] = \zeta_0 \equiv (6 - 4V_0)/(6 - 3V_0)$ when computing $\delta Z/\delta h$ in the Euler–Lagrange equation

$$\frac{\delta \ell^2}{\delta h} - \lambda_1 \frac{\delta I_1}{\delta h} + \frac{\lambda_2}{\zeta}\frac{\delta \zeta}{\delta h} + \lambda_3 V_0 \frac{\delta Z}{\delta h} = 0$$

to obtain

$$h_{xx} = \frac{\lambda_1}{h^2} + \lambda_2\left(\frac{1}{I_1 h^2} - \frac{4}{I_2 h^3} + \frac{3}{I_3 h^4}\right) + \lambda_3\left[1 + \frac{1}{4 - 3\zeta_0}\left(\frac{2}{I_2 h^3} - \frac{3}{I_3 h^4}\right)\frac{I_2}{I_3}\right]. \tag{3.16}$$

As shown in (4.15) below, $Z[h] > \frac{2}{3}A[h]$ for any non-constant, positive wave profile $h(x)$, so fluid loss is a positive quantity. Moreover, fluid loss (like fluid volume) scales linearly with the wave profile, i.e. $Z[\theta h] = \theta Z[h]$ for any positive number θ and wave profile h.

The fluid loss constraint in this problem serves the same purpose as the area constraint in Problem 3, namely, to pin down a degree of freedom in the dimensionless problem so that the vertical length scale of the physical problem is entirely controlled by H. This constraint normalizes the solutions in the two-parameter family (with speed and efficiency as parameters) to have the same cost (for fixed H) in terms of the energy associated with mucus production. Once the optimal solutions have been found with Z held fixed, the vertical length scale H of the physical problem may be chosen to balance the cost due to fluid production with the mechanical energy per unit distance travelled (the efficiency) required to swim: the former varies linearly with H while the latter varies inversely with H.

4. Numerical methods and results

In this section we describe two numerical methods for solving the optimization problems posed in the previous section and analyse the results. In the first approach, we represent the wave profile $h(x)$ using periodic cubic splines and use a limited-memory BFGS method (Nocedal & Wright 1999) to descend to the optimal shape. In the second approach, we solve the Euler–Lagrange equations using a quadratically convergent multi-shooting Runge–Kutta algorithm that uses the Levenberg–Marquardt method to vary the parameters of the ODE until the constraints are satisfied. The first approach is fast and robust unless the solution is very singular. It also gives useful information in cases where the optimization problem has no solution, and can be generalized to study optimization in the non-lubrication

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

regime. In this regime, one cannot write down Euler–Lagrange equations for the solution, but can often solve an adjoint PDE to find the gradient of the objective function with respect to perturbations of the curve; see e.g. Pironeau (1984). The second approach requires a good starting guess, but is otherwise fast and highly accurate even when the solution is nearly singular.

We use the spline approach to study Problems 1 and 2 of the previous section and show that solutions of (3.12) minimize efficiency rather than maximize it. We use the ODE approach to study the two-parameter families of Problems 3 and 4 and to unify the one-parameter families of Problems 1, 2, and 3a. We also explore various singular limits as the constraint parameters approach critical values and the optimal curves develop singularities. By monitoring the second- and fourth-order corrections to the lubrication theory, we conclude that the lubrication approximation remains uniformly valid for Problem 1 even in the critical limit $A_0 \to \ell_0/\sqrt{6}$, but breaks down for Problem 3a in the critical limit $V_0 \to 3/2$ due to the formation of a cusp in the optimal shape. We also find that for Problems 3 and 4, optimization within lubrication theory drives us out of its realm of applicability; however, if higher-order corrections are taken into account, cusp-like singularities are appropriately penalized and the optimal solutions for finite ε are non-singular.

4.1. *Direct minimization using periodic cubic splines*

4.1.1. *Problem 1 (maximize speed)*

The basic strategy in this section is to represent the wave profile $h(x)$ using periodic cubic splines and use a gradient-based minimization algorithm to modify the degrees of freedom in the spline space to minimize the functional $G[h]$ in (3.5). We used the L-BFGS algorithm (Nocedal & Wright 1999), which is a limited-memory, quasi-Newton line search method that builds an approximate Hessian incrementally from the history of gradients $g(q)$ it has evaluated. We start with relatively few knots in the spline space (typically 32 or 64) and find the optimal shape in this space. We then repeatedly double the number of knots using the previous solution as the starting guess until the variational derivative $\delta G/\delta h$ is sufficiently small; (see Appendix B.1 for details).

Figure 4 shows the optimal solution for $\ell_0 = 0.32$, $A_0 = 1.2\ell_0/\sqrt{6}$. This choice of ℓ_0 was arbitrary – a different choice would change the labelling of the y-axis, but not the shape of the optimal curve $h(x)$, which depends only on A_0/ℓ_0. Although the variational derivative $\delta G/\delta h$ is not identically zero, it is orthogonal to the spline space and any perturbation of the spline degrees of freedom will lead to a larger value of $G[h]$. As we refine the mesh, the true solution becomes better approximated by spline functions and $\delta G/\delta h$ converges to zero. The relative error in the computed swimming speed (using $V = 1.03017842643553$ obtained from the ODE method of §4.2 as a benchmark) and the number of iterations the BFGS algorithm required to converge are given in table 1. The sixth-order convergence rate is due to our mesh refinement strategy, which involves quadratic clustering of spline knots near the origin. The entire computation took approximately 2 s on a 2.4 GHz desktop machine. When A_0 is closer to $A_{\text{crit}} = \ell_0/\sqrt{6}$ and the solution is more sharply kinked at the origin, the number of iterations required by the BFGS algorithm can increase into the tens of thousands, causing the computation to take minutes to hours. By contrast, the ODE method of §4.2 continues to require only a few seconds to run, and maintains accuracy for A_0 much closer to A_{crit}. On the other hand, the spline method generalizes to the non-lubrication regime.

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

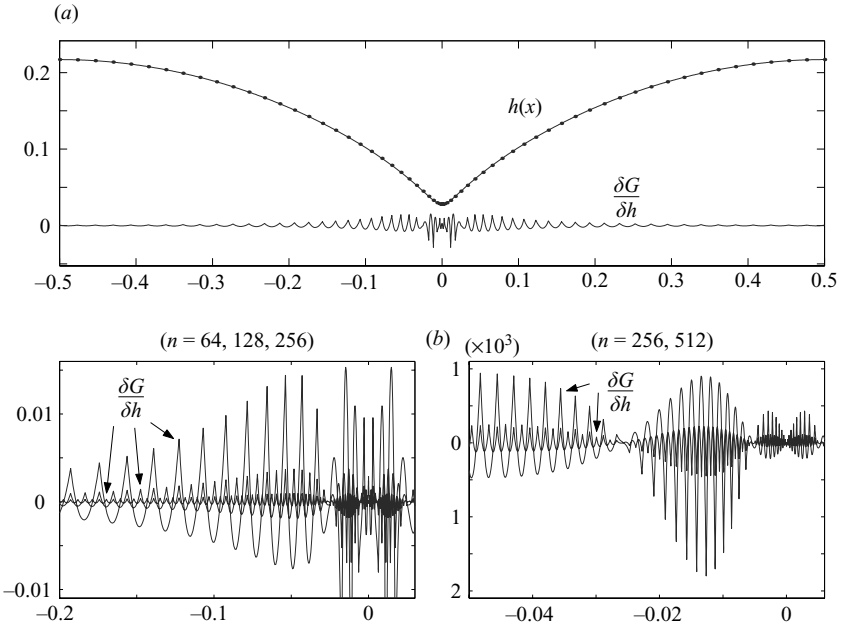


FIGURE 4. (*a*) Fastest wave profile and its variational derivative: solution obtained by direct minimization using periodic splines with 64 knots ($\ell_0 = 0.32$, $A_0 = 1.2\ell_0/\sqrt{6}$). (*b*) The variational derivative of the optimal solution in the spline space converges to zero as the grid is refined. This variational derivative is computed exactly (up to roundoff error), and is orthogonal to the spline space.

n	Relative error	Function evaluations
64	4.2×10^{-7}	306
128	5.0×10^{-9}	178
256	7.3×10^{-11}	82
512	1.3×10^{-12}	50

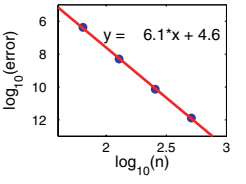


TABLE 1. The number of iterations that the BFGS algorithm required to converge to the optimal solution in the spline space.

4.1.2. *Problem 2 (maximize efficiency)*

We now try to optimize efficiency, defined as the ratio of swimming speed to mechanical power required to swim at that speed. As described in § 3, $V/P = 1/(2I_1)$; hence, we wish to minimize I_1 in order to maximize efficiency. We seek the global minimizer of the functional

$$G[h] = I_1[\alpha h + \beta] + (\alpha - 1)^2/2 + \beta^2/2, \tag{4.1}$$

where $\alpha = \ell_0(\frac{1}{2}\int_0^1 h_x^2 \,\mathrm{d}x)^{-1/2}$ and $\beta = A_0 - \alpha \int_0^1 h \,\mathrm{d}x$ are used to enforce the area and arclength constraints; (see Appendix B.1 for details). Using the formula (3.11) for $\delta G/\delta h$ and proceeding as in Problem 1, we discover that the optimal discrete solution (in the spline space) is nearly equal to the constant function $h(x) = A_0$, with the

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

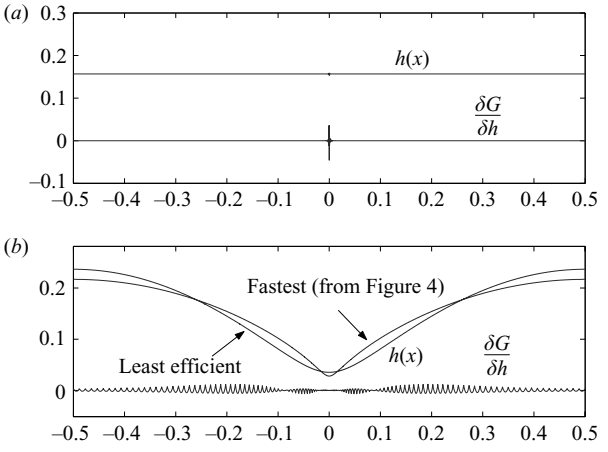


FIGURE 5. (*a*) The most efficient wave profile ($n = 128$, $\ell_0 = 0.32$, $A_0 = 1.2\ell_0/\sqrt{6}$) in the spline space is as close to a constant as possible without violating the constraints. There is no solution in the mesh refinement limit. (*b*) The least efficient wave profile is flatter near its minimum than the fastest profile for the same ℓ_0 and A_0. The relationship between these two curves will be explored in § 4.2.

exception of a few high-frequency oscillations to lengthen the curve and satisfy the arclength constraint (see figure 5). This can be understood as follows. By Jensen's inequality,

$$I_1 = \int_0^1 h(x)^{-1}\,\mathrm{d}x \geqslant \left(\int_0^1 h(x)\,\mathrm{d}x\right)^{-1} = A_0^{-1}. \tag{4.2}$$

It can be shown that any minimizing sequence h_k (such that $I_1[h_k] \to A_0^{-1}$) that satisfies the constraints (3.1) converges uniformly to the constant function $h(x) = A_0^{-1}$. Since we assume $\ell_0 > 0$, this limiting function violates the arclength constraint and there is no optimal solution. Moreover, in order to approach the optimal efficiency $V/P = A_0/2$, the swimming speed V must approach zero; this is because a function h which is uniformly close to the constant function A_0 will satisfy $I_k \approx A_0^{-k}$, hence $\zeta = I_2^2/(I_1 I_3) \approx 1$ and $V = 6(1-\zeta)/(4-3\zeta) \approx 0$. In other words, a sheet with no deflection does not swim. Interestingly, if we put a minus sign in front of I_1 in (4.1), we find that as long as $A_0 > \ell_0/\sqrt{6}$, there is a well-defined "least efficient" swimmer (see figure 5*b*). Solutions of the Euler–Lagrange equation (3.12) actually solve this problem rather than the problem we intended to solve. We remark that Jensen's inequality also implies that the denominator of (3.12) is strictly positive for any non-constant positive function h; however, if h comes close to maximizing efficiency by approaching the constant function $h(x) = A_0$, the denominator tends to zero.

4.2. *Solution of the Euler–Lagrange equations via ODE methods*

4.2.1. *Problem 1 (maximize speed)*

In this section we use the Levenberg–Marquardt method (Nocedal & Wright 1999) to vary the parameters I_k and the initial condition $h(0) = b$ in the Euler-Lagrange equation

$$h_{xx} = \left(\frac{2\ell_0^2}{A_0}\right)\left(\frac{h^{-2}/I_1 - 4h^{-3}/I_2 + 3h^{-4}/I_3}{I_2/I_1 - 4I_3/I_2 + 3I_4/I_3} - 1\right) \tag{4.3}$$

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

until the solution is periodic, the constraints are satisfied, and the integrals of the solution match the parameters. More precisely, given the vector

$$q = (b, I_1, I_2, I_3, I_4), \tag{4.4}$$

we define the function $h_q(x)$ as the solution of the ordinary differential equation (4.3) with initial conditions $h_q(0) = b$, $h_q'(0) = 0$. We then use Minpack (which employs the Levenberg–Marquardt method) to solve the nonlinear system of equations $r(q) = 0$, where

$$r_0 = h_q'(1/2), \qquad r_{\{k=1,2,3\}} = \frac{2}{I_k}\int_0^{1/2} h_q^{-k}\,\mathrm{d}x - 1, \qquad r_4 = \frac{2}{A_0}\int_0^{1/2} h_q\,\mathrm{d}x - 1. \tag{4.5}$$

In order to use the Levenberg–Marquardt method, we also need to compute the Jacobian $J = \nabla_q r$, which we do by solving variational equations for the derivatives $\partial h/\partial q_i$ (see Appendix B.2 for details). Since (4.3) is invariant under translation $x \to x + a$ and reflection $x \to -x$, its solutions are symmetric about their extrema and all critical points are extrema; therefore, requiring that $h'(0) = 0$ and $h'(1/2) = 0$ is equivalent to enforcing periodicity via $h(1) = h(0)$ and $h'(1) = h'(0) = 0$. The remaining two conditions $\frac{1}{2}\int_0^1 h_x^2\,\mathrm{d}x = \ell_0^2$ and $\int_0^1 h^{-4}\,\mathrm{d}x = I_4$ are satisfied automatically, which can be seen by integrating (4.3) and its product with h from 0 to 1 and using periodicity.

We use this strategy to study the behaviour of the swimming sheet as the area constraint A_0 decreases to the critical value $A_{\mathrm{crit}} = \ell_0/\sqrt{6}$ at which the sheet can touch the wall. We ran 201 trials with the constraints $\ell_0 = 0.32$, $A_0 = (1 + \delta)A_{\mathrm{crit}}$, where δ ranged from $\delta_0 = 0.2$ to $\delta_{200} = 9.945 \times 10^{-8}$ via the recursion $\delta_{k+1} = 0.93\delta_k$. Since the Levenberg–Marquardt method is quadratically convergent when used to solve nonlinear equations (as opposed to overconstrained least-squares problems), very few iterations (typically 8–30) are required to converge to the solution as long as a good initial guess is known in advance. On the first and second iteration, we used the method of § 4.1 to compute a starting guess for $q = (b, I_{\{1,2,3,4\}})$. After that, we used logarithmic extrapolation to obtain the starting guess q, e.g. $b_{k+1} = \exp[2\log(b_k) - \log(b_{k-1})]$.

The results of this study are summarized in figure 6. Figure 6(*a*) shows that the optimal curves $h_k(x)$ do indeed approach the parabola $h_{\mathrm{crit}}(x) = \sqrt{6}\ell_0(|x| - x^2)$ as $\delta_k \to 0$ and $(A_0)_k \to A_{\mathrm{crit}}$. Figure 6(*c*) shows how the power, $P = 2I_1V$, and shape parameter, ζ, depend on δ_k: the power required to swim eventually grows linearly with ζ^{-1}, and diverges as the swimming speed approaches $V = 3/2$. Recall that the swimming speed V is related to ζ via $V = 6(1 - \zeta)/(4 - 3\zeta)$, so for small ζ we have $V \approx \frac{3}{2} - \frac{3}{8}\zeta$. We also find that solutions are self-similar near touchdown when rescaled by

$$y_k(\xi) = \frac{1}{\delta_k \ln(1/\delta_k)} h_k(\delta_k \ln(1/\delta_k)\,\xi). \tag{4.6}$$

Rigorous validation of this scaling law is difficult as it requires a delicate matched asymptotic analysis between the inner expansion for $h(x)$ near $x = 0$ (described via $y(\xi)$) and the outer expansion $h(x) \approx \sqrt{6}\ell_0(|x| - x^2)$ over the rest of the interval. We arrived at this scaling law by trial and error; our first guess was that $y_k(\xi) = \delta_k^{-1}h_k(\delta_k\xi)$ might work, but this ansatz yields functions y_k that drift upward (i.e. translate vertically without changing their shape) as $\delta_k \to 0$, which led us to expect a logarithm to be involved.

This family of wave profiles exhibits the interesting property that as $\delta \to 0$, a singularity in curvature develops in the region where the gap thickness approaches

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

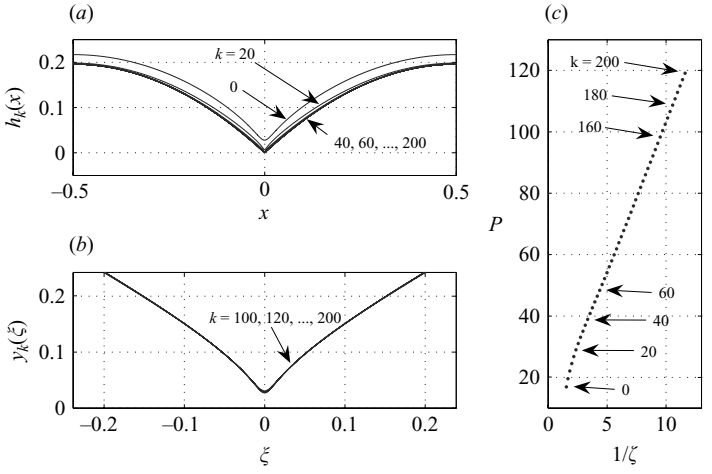


FIGURE 6. (*a*) Wave profiles that optimize speed: the optimal wave profiles $h_k(x)$ approach the periodic parabola $\sqrt{6}\ell_0(|x| - x^2)$ as $\delta_k \to 0$. (*b*) Self-similarity near touchdown: the curves appear to be self-similar when scaled using (4.6); the curves shown lie in the range $1.4 \times 10^{-4} \geqslant \delta_k \geqslant 1.0 \times 10^{-7}$. (*c*) Power and speed: as $\delta_k \to 0$, the swimming speed $V \approx \frac{3}{2} - \frac{3}{8}\zeta$ approaches $\frac{3}{2}$ while the power required to swim at that speed grows without bound.

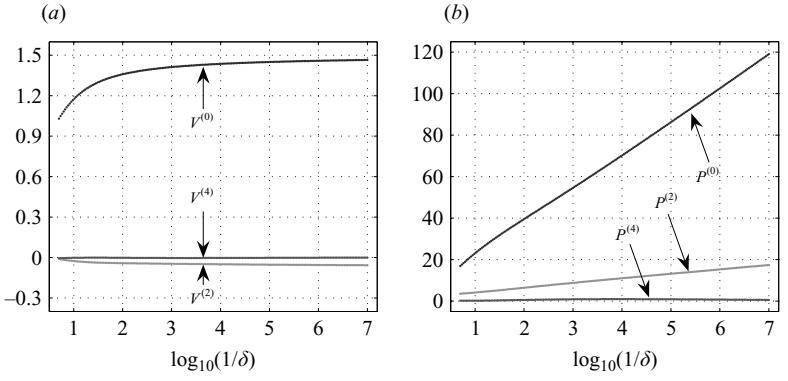


FIGURE 7. Plots of the expansion coefficients for (*a*) speed $V = V^{(0)} + \varepsilon^2 V^{(2)} + \varepsilon^4 V^{(4)} + \cdots$ and (*b*) power $P = P^{(0)} + \varepsilon^2 P^{(2)} + \varepsilon^4 P^{(4)} + \cdots$, where $\varepsilon = H/W$. Since the higher-order coefficients remain small in comparison to $V^{(0)}$ and $P^{(0)}$, the lubrication approximation appears to remain valid as the constrained area $A_0 = (1 + \delta)A_{\text{crit}}$ approaches $A_{\text{crit}} = \ell_0/\sqrt{6}$.

zero. Singularities in curvature tend to invalidate the lubrication approximation while narrow constrictions are favourable. To determine which tendency prevails, we consider higher-order terms in the expansion of the solution of the Stokes equations in powers of the aspect ratio $\varepsilon = H/W$ (see Appendix C). In figure 7, we plot the coefficients of the expansions of the speed and power in powers of ε up to fourth order,

$$V = V^{(0)} + \varepsilon^2 V^{(2)} + \varepsilon^4 V^{(4)} + \cdots, \qquad P = P^{(0)} + \varepsilon^2 P^{(2)} + \varepsilon^4 P^{(4)} + \cdots, \tag{4.7}$$

and find that the coefficients of the higher-order terms are bounded by the zeroth-order terms as $\delta \to 0$. This is strong evidence that the second- and fourth-order corrections can be made arbitrarily small in comparison to the zeroth-order terms

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

(uniformly in δ) by a single choice of $\varepsilon = H/W$. We note that although $P^{(2)}$ diverges as $\delta \to 0$, it is bounded by $P^{(0)}$, which also diverges; hence, $\varepsilon^2 P^{(2)} \ll P^{(0)}$ for small ε. In summary, owing to the proximity of the substrate, the lubrication approximation remains valid as $\delta \to 0$ in spite of the formation of a singularity in curvature. See Appendix C for further discussion of the validity of lubrication expansions such as (4.7) on nearly singular domains.

4.2.2. *Problems 3 and 4 (maximize efficiency for a given swimming speed)*

We now turn to the problem of optimizing efficiency when arclength, speed and fluid volume (or fluid loss) are constrained. The Euler–Lagrange equations (3.13) and (3.16) may be studied simultaneously via

$$h_{xx} = \frac{\lambda_1}{h^2} + \lambda_2\left(\frac{1}{I_1 h^2} - \frac{4}{I_2 h^3} + \frac{3}{I_3 h^4}\right) + \lambda_3\left[1 + \frac{\gamma}{4-3\zeta_0}\left(\frac{2}{I_2 h^3} - \frac{3}{I_3 h^4}\right)\frac{I_2}{I_3}\right], \quad (4.8)$$

where $\gamma = 0$ for Problem 3 and $\gamma = 1$ for Problem 4. We wish to vary

$$q = (b, I_1, I_2, I_3, \lambda_1, \lambda_2, \lambda_3) \quad (4.9)$$

until the constraints $r(q) = 0$ are satisfied, where

$$\left.\begin{aligned} r_0 = h'(1/2), \quad r_{k=1,2,3} &= \frac{2}{I_k}\int_0^{1/2} h^{-k}\,\mathrm{d}x - 1, \\ r_4 = \frac{1}{\zeta_0}\left(\frac{I_2^2}{I_1 I_3}\right) - 1, \quad r_5 = \frac{2}{A_0}\int_0^{1/2} h\,\mathrm{d}x - 1, \quad r_6 &= \frac{2}{\ell_0^2}\int_0^{1/2} h_x^2\,\mathrm{d}x - 1. \end{aligned}\right\} \quad (4.10)$$

If we instead choose to minimize arclength subject to an efficiency constraint, we replace the formula for r_6 by $2I_1 V_0/P_0 - 1$, where V_0/P_0 is the prescribed efficiency. We can solve Problems 1 and 2 by dropping the speed constraint and driving the appropriate Lagrange multiplier to zero, i.e. by setting $r_4 = \lambda_1$ or $r_4 = \lambda_2$, respectively: in the case of Problem 2, we obtain the least efficient swimmer of § 4.1.2. We can solve Problem 3a by setting $r_5 = \lambda_3$. We can solve Problem 4 by setting $\gamma = 1$ in (4.8) and

$$r_5 = \frac{1}{V_0}\left[2\int_0^{1/2} h\,\mathrm{d}x - \frac{1}{4-3\zeta_0}\frac{I_2}{I_3}\right] - Z_0, \qquad r_6 = 2I_1\frac{V_0}{P_0} - 1. \quad (4.11)$$

Thus, all the optimization problems described in § 3 are closely related.

Before studying Problems 3 and 4 in full generality, let us consider the one-parameter family (Problem 3a) of solutions obtained by dropping the fluid volume constraint. Given $0 < V_0 < 3/2$, we wish to

$$\text{maximize} \quad \frac{V[h]}{P[h]} \quad \text{subject to the constraints :} \quad \ell^2[h] = \ell_0^2, \qquad V[h] = V_0, \quad (4.12)$$

where ℓ_0 is arbitrarily chosen to be 0.32. The optimal solution scales linearly with ℓ_0, so this choice merely pins down the vertical length scale of the dimensionless problem. This problem is equivalent to minimizing $I_1[h]$ subject to $\ell^2[h] = \ell_0^2$ and $\zeta[h] = \zeta_0$, where $\zeta_0 = (6-4V_0)/(6-3V_0)$. We ran 101 trials with $(\zeta_0)_k = \delta_k$ ranging from $\delta_0 = 0.5$ to $\delta_{100} = 0.001$ via the recursion $\delta_{k+1} = 0.94\delta_k$. The results of this study are summarized in figure 8. Through trial and error, we discovered that the solutions are self-similar near the origin in the sense that the re-scaled functions

$$y_k(\xi) = \frac{1}{\delta_k^2} h_k\left(\delta_k^3 \xi\right) \quad (4.13)$$

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

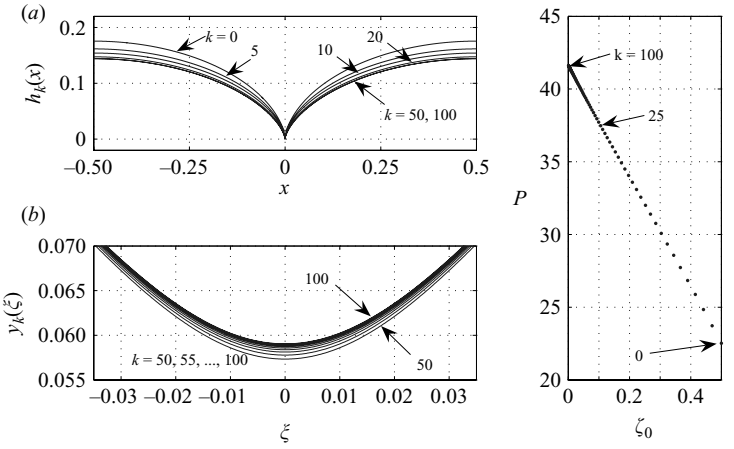


FIGURE 8. (*a*) Wave profiles that optimize efficiency: the optimal wave profiles $h_k(x)$ of Problem 3a approach a limiting curve with a cusp at the origin. (*b*) Self-similarity near touchdown: the curves y_k appear to be self-similar as $\delta_k \to 0$ when scaled as in (4.13). (*c*) Power and speed: unlike the area constrained problem, the power $P = 2I_1V$ remains bounded as the speed $V \approx \frac{3}{2} - \frac{3}{8}(\zeta_0)_k$ approaches $\frac{3}{2}$.

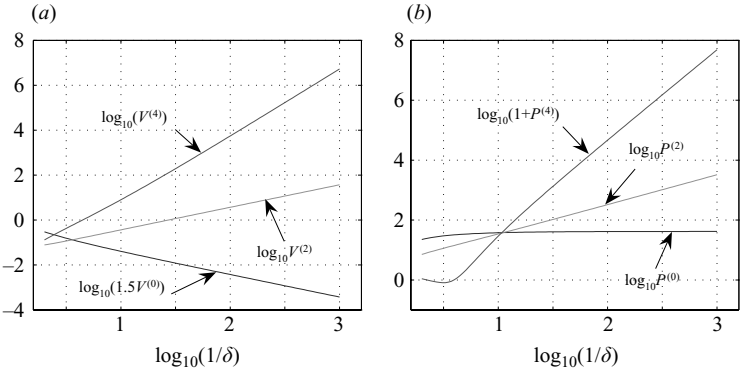


FIGURE 9. Expansion coefficients for (*a*) speed and (*b*) power. In contrast to Problem 1, the higher-order coefficients grow without bound as $\delta \to 0$ while the zeroth-order coefficients remain bounded; hence the lubrication approximation cannot be used to study the critical limit $\delta \to 0$.

converge to a common curve as $\delta_k \to 0$. This suggests that the cusp in the limiting curve in figure 8(*a*) is asymptotically of the form $\alpha|x|^{2/3}$, which has a finite first inverse moment I_1 in the power formula $P = 2I_1V$. By contrast, I_1 is infinite for the limiting curve $\sqrt{6}\ell_0(|x| - x^2)$ of Problem 1. Unfortunately, this cusp also causes the maximum slope to blow up and the curvature to diverge faster than the gap thickness approaches zero (so that $h(0)h_{xx}(0)$ blows up); as explained in Appendix C, we expect as a result that as $\delta \to 0$ with $\varepsilon > 0$ held fixed, the geometry will leave the lubrication regime. This is confirmed in figure 9, where we see that the higher-order expansion coefficients $V^{(2)}$, $V^{(4)}$, $P^{(2)}$, and $P^{(4)}$ grow without bound relative to $V^{(0)}$ and $P^{(0)}$ as $\delta \to 0$. Before addressing this important issue, we wish to include fluid volume (or fluid loss) in the problem and shuffle the objective function and constraints as discussed previously in § 3.

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

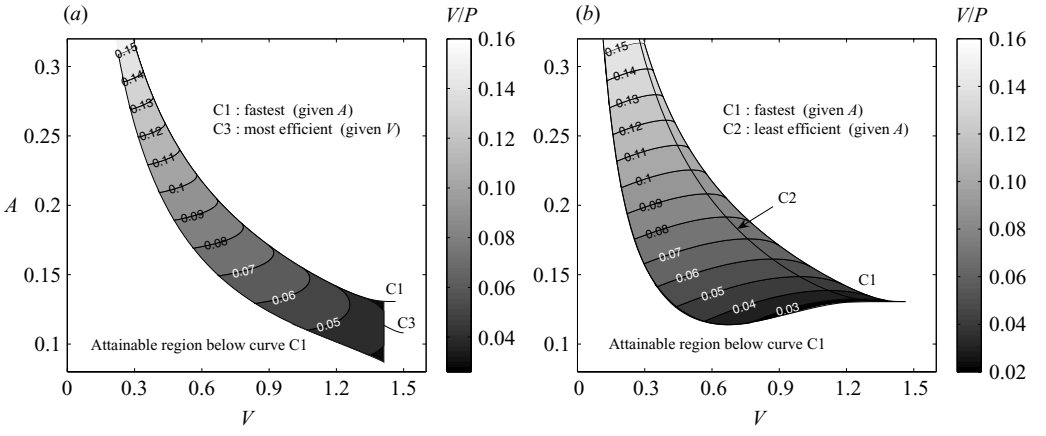


FIGURE 10. Contour plots of maximum (*a*) and minimum (*b*) efficiency vs. speed and fluid volume for Problem 3 with $\ell = 0.32$ held fixed. The boundary C1 of the attainable region is the solution of Problem 1, while curves C2 and C3 correspond to Problems 2 and 3a, respectively.

Figures 10–12 show the solutions of Problem 3 from various complementary perspectives. In figure 10, we optimize efficiency, varying V_0 and A_0 with $\ell_0 = 0.32$ held fixed. The attainable region is determined by solving Problem 1 for the maximum speed $V_{\max}$ as a function of fluid volume A_0. Inside this region, solutions are found using a homotopy method in which the constraints V_0 and A_0 are varied slightly and the parameters q_j in (4.9) of the previous solution are used as an initial guess in the Levenberg–Marquardt method. Interior boundaries of the region shown were empirically chosen to avoid parameters V_0 and A_0 for which the problem becomes numerically ill-conditioned and our Levenberg–Marquardt method fails.

For each choice of V_0 and A_0 there are two solutions of (4.8) and (4.10): one maximizing the efficiency and the other minimizing it. Curve C2 contains the minimum-efficiency solutions when the speed constraint is dropped (the "least efficient" variant of Problem 2). Curve C3 contains the maximum-efficiency solutions when the fluid volume constraint is dropped (Problem 3a). Note that the curves C1 and C3 maximizing speed and efficiency are close to each other for $V \leqslant 1$, but differ substantially for V close to 3/2. For any integer $n \geqslant 2$, we can replace A_0 by A_0/n and $h(x)$ by $h(nx)/n$ to obtain a "higher-frequency" solution of the Euler–Lagrange equation with efficiency multiplied by $1/n$. We have not found any local extrema other than the two main families shown in figure 10 and their higher frequency variants.

In figure 11, we reverse the roles of fluid volume and efficiency, treating the former as the objective function and the latter as a constraint. The solid and dotted contour lines correspond to maximizing or minimizing fluid volume, respectively. This time the boundary C3 of the attainable region is found by solving Problem 3a, and the curves C1 or C2 maximize fluid volume when the efficiency or speed constraint is dropped, respectively. In each of these cases, although it is biologically preferable to minimize fluid volume, the solution of the Euler–Lagrange equations with these constraints turns out to maximize it, as happened in Problem 2 when optimizing efficiency. The points labelled B1, B3, E1 and E3 lie on curves C1 or C3 with $V = 1.2$ or $V = 0.5$, respectively. The corresponding wave profiles are shown. The points labelled M and m correspond to the two solutions of the Euler-Lagrange equations with $V = 0.7$, $V/P = 0.083$ and $\ell = 0.32$; A is maximized by M and minimized by m.

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

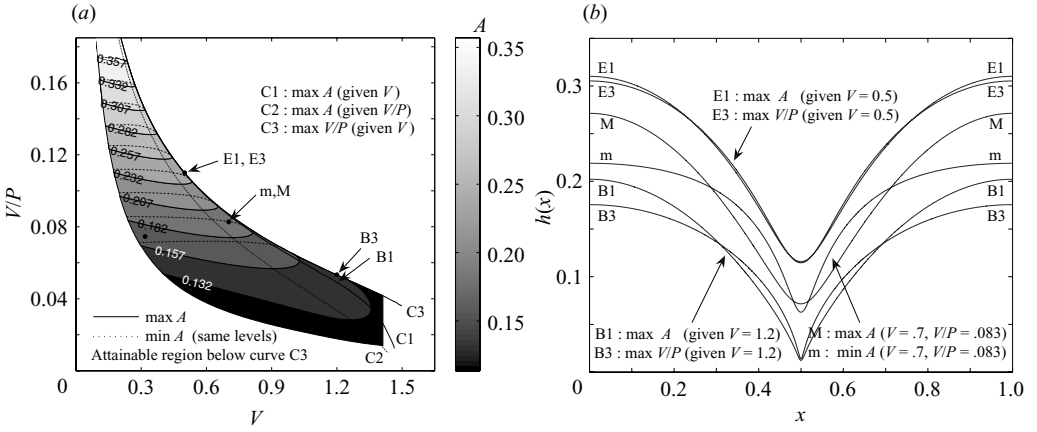


FIGURE 11. (*a*) Maximum and minimum fluid volume A versus speed V and efficiency V/P with $\ell = 0.32$ held fixed. (*b*) Optimal curves $h(x)$ corresponding to the labelled points.

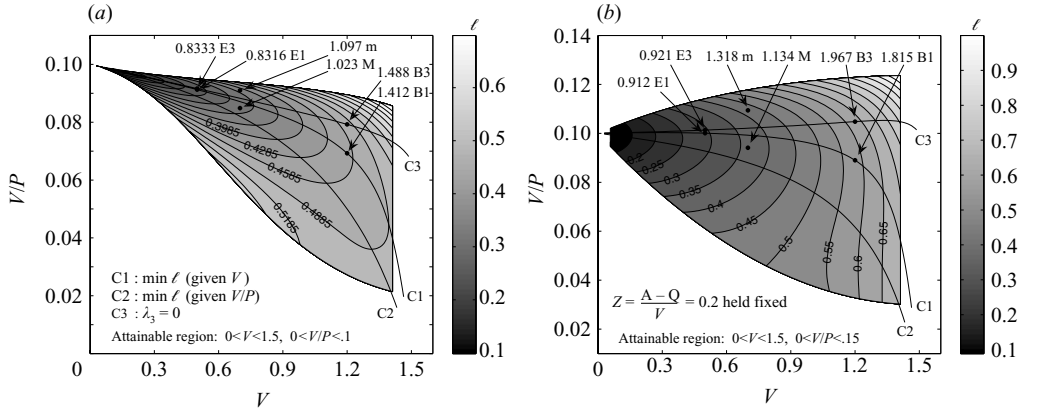


FIGURE 12. (*a*) Minimum arclength ℓ versus speed V and efficiency V/P. The curves $h(x)$ minimizing ℓ with $A = 0.2$ held fixed are re-scaled versions of the curves of figures 10 and 11. Re-scaling changes the efficiency of each curve but not the speed, and 'unfolds' the contour plot in figure 11 across C3. (*b*) A different re-scaling of the curves to normalize their cost in terms of 'fluid lost per distance travelled', namely $Z = 0.2$.

In figure 12(*a*), we minimize arclength, varying speed and efficiency with $A_0 = 0.2$ held fixed. The fluid volume constraint should now be thought of as pinning down a degree of freedom of the dimensionless problem so that the vertical length scale of the physical problem is controlled entirely by H. Minimizing arclength then gives the smoothest solution that achieves the desired goal (of speed and efficiency). The curves C1, C2 and C3 are the zero level sets of the Lagrange multipliers λ_1, λ_2 and λ_3, respectively. C1 and C2 cross the contour lines of ℓ at vertical and horizontal tangents, but C3 does not have an obvious geometrical interpretation when A_0 is held fixed.

The points on the contour plots of figures 10–12 correspond to re-scaled versions of the same wave profiles $h(x)$. For example, the curves $h(x)$ labeled m and M in figure 11 must be multiplied by 1.023 and 1.097, respectively, to satisfy the fluid volume constraint $A_0 = 0.2$. This re-scaling affects the efficiency of each curve but not the speed, and "unfolds" the contour plot in figure 11 across C3, making it single

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

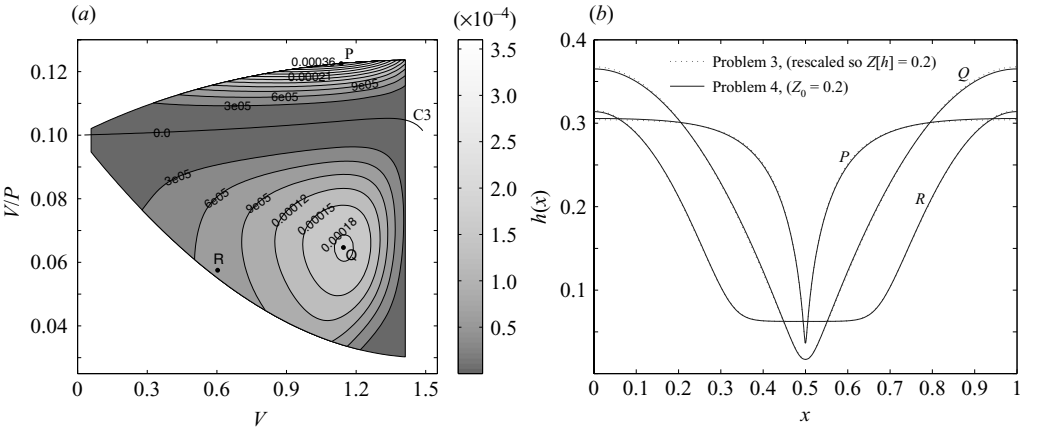


FIGURE 13. Comparison of Problems 3 and 4. (*a*) Deviation from minimum arclength when solutions of Problem 3 are re-scaled for constant fluid loss. We subtract the minimum arclength solution found by solving Problem 4 with $Z_0 = 0.2$ from the value shown in figure 12(*b*). Problems 3 and 4 have identical solutions along C3, where $\lambda_3 = 0$ in both problems. (*b*) The re-scaled wave profiles for Problem 3 are nearly identical to the minimum arclength solutions. $P: V = 1.136$, $V/P = 0.1225$; $Q: V = 1.146$, $V/P = 0.0647$; $R: V = 0.603$, $V/P = 0.0575$.

valued. For example, m and M now lie on opposite sides of C3 whereas before they fell on top of each other.

In figure 12b, we re-scale the optimal wave profiles of Problem 3 so that the fluid loss per distance travelled is constant ($Z = 0.2$). These solutions are not optimized for fluid loss; instead, the fluid volume constraint A_0 is chosen to vary with speed and efficiency in such a way that the fluid loss Z turns out to be a constant for all curves in the two-parameter family. The curves C1, C2 and C3 remain the zero level sets of the Lagrange multipliers, but they no longer cross the contours of ℓ in a geometrically significant manner (since A_0 now varies with V_0 and V_0/P_0). Although these solutions were not optimized for fluid loss, they are almost identical to the optimal solutions of Problem 4 (see figure 13).

Optimal swimming should involve maximizing efficiency and swimming speed while minimizing fluid loss. We have already argued that one way to treat fluid loss is to constrain it to a fixed value in the dimensionless problem so that it becomes controlled by the vertical length scale H of the physical problem. Once the optimal solutions have been found with Z held fixed, H may be chosen to balance the energetic cost of fluid loss with that of mechanical energy consumption: the former varies linearly with H while the latter varies inversely with H. Thus, we wish to maximize both V and V/P in figure 12(*b*). By (2.15) and (4.2), we have the upper bounds

$$V < \frac{3}{2}, \qquad \frac{V}{P} = \frac{1}{2I_1} < \frac{A}{2} < \frac{3}{4}Z, \tag{4.14}$$

where we have used Jensen's inequality

$$I_2 = \int_0^1 h^{-2}\,\mathrm{d}x > \left(\int_0^1 h^{-1}\,\mathrm{d}x\right)^2 > A^{-1}I_1$$

in

$$Z = \frac{A-Q}{V} = \frac{(4-3\zeta)A - I_2/I_3}{6(1-\zeta)} = \frac{2}{3}A + \frac{\zeta}{6}\left(\frac{A - I_1/I_2}{1-\zeta}\right) > \frac{2}{3}A. \tag{4.15}$$

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

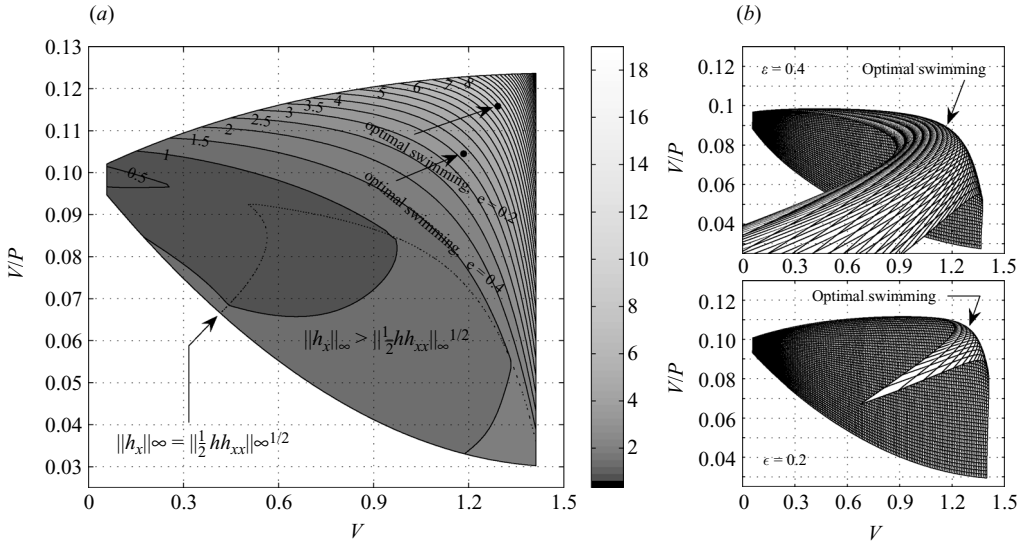


FIGURE 14. (*a*) $\max(||h_x||_\infty, ||\frac{1}{2}hh_{xx}||_\infty^{1/2})$ versus speed and efficiency for solutions of Problem 3, re-scaled for constant fluid loss ($Z[h] = 0.2$). These quantities dictate how small $\varepsilon = H/W$ must be for the lubrication approximation to remain valid. For smaller values of ε, $(V, V/P)$ can approach the theoretical maximum $(1.5, 0.15)$ more closely without leaving the realm of lubrication theory. (*b*) Corrections to lubrication theory: if corrections up to fourth order in ε are included, the attainable region 'folds over' itself, penalizing wave profiles that are too cusp-like.

We have assumed here that $h(x)$ is a non-constant, positive function. The bounds (4.14) can be simultaneously approached arbitrarily closely, i.e. we can find a sequence of wave profiles such that $(V, V/P) \to (\frac{3}{2}, \frac{3}{4}Z_0)$, but not without leaving the realm of validity of lubrication theory.

We now face a dilemma: optimization within lubrication theory drives us out of its realm of applicability. To optimize speed and efficiency using lubrication theory, we want to find wave profiles for which I_1 and $I_2^2/(I_1 I_3)$ are both small – but this leads to cusp-like wave profiles $h(x)$ with large slopes and gap-weighted curvatures:

$$\varepsilon_0^{-1} = \max\left(\|h_x\|_\infty, \sqrt{\|\tfrac{1}{2}hh_{xx}\|_\infty}\right) \gg 1, \qquad \|f\|_\infty \equiv \max_{0\leqslant x\leqslant 1} |f(x)|. \tag{4.16}$$

The significance of ε_0 is that it serves as an effective radius of convergence for the lubrication model – the model is valid as long as $\varepsilon \ll \varepsilon_0$, where $\varepsilon = H/W$; see Appendix C. Minimizing ℓ is effective in keeping ε_0^{-1} as small as possible for a given attainable pair $(V, V/P)$, but ε_0^{-1} blows up as we approach $\left(\frac{3}{2}, \frac{3}{4}Z_0\right)$; see figure 14. This means that for a fixed value of $\varepsilon = H/W$, we can only approach $\left(\frac{3}{2}, \frac{3}{4}Z_0\right)$ so far before $\varepsilon/\varepsilon_0$ ceases to be small and the lubrication model breaks down.

This does not mean we should abandon our approach: the key forces involved in this problem *are* lubrication forces, and optimization *does* produce fast, efficient wave profiles even for the full Stokes equations; however, it would be helpful to incorporate some mechanism in our lubrication model to penalize cusp-like singularities and indicate when to stop optimizing. As a first step, we take the two-parameter family of optimal curves found entirely within lubrication theory, but then compute second- and fourth-order corrections to the swimming speed and efficiency for various choices

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

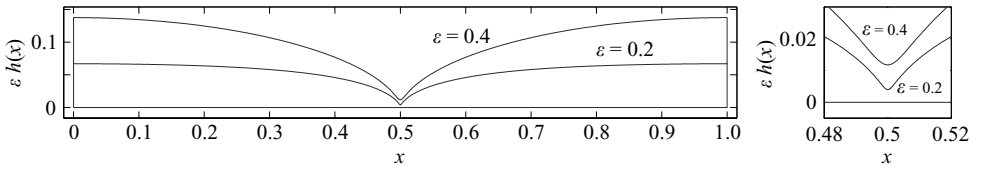


FIGURE 15. Wave profiles identified in figure 14 as optimal when corrections up to fourth order in ε are included in the calculation of speed and efficiency. The plot on the right is an enlargement.

of ε. We find that as the parameters $(V, V/P)$ approach $\left(\frac{3}{2}, \frac{3}{4}Z_0\right)$ and ε_0^{-1} increases, these corrections eventually dominate the calculation and cause the attainable region to "fold over" itself, penalizing wave profiles that are too cusp-like (see figure 14*b*). A smaller choice of ε will suppress these corrections, allowing us to get closer to $(\frac{3}{2}, \frac{3}{4}Z_0)$ before the lubrication model breaks down.

In figure 15, we show the optimal wave profiles obtained in this way using $\varepsilon = 0.2$ and $\varepsilon = 0.4$. These curves correspond to the points $(V, V/P)$ indicated in figure 14 – they are interior points of the attainable region in the zeroth-order theory, but are mapped to boundary points (along the fold) when fourth-order corrections are taken into account. To show correct aspect ratios, we plot $\varepsilon h(x)$ rather than $h(x)$; these curves may be interpreted as viewing $\hat{h}(\hat{x})$ with the x- and y-axes measured in units of W. Within the zeroth-order theory, the $\varepsilon = 0.2$ curve has a cost due to fluid loss that is exactly half that of the $\varepsilon = 0.4$ curve, and a cost due to mechanical energy that is $2(0.1045)/(0.1158) = 1.805$ times larger than that of the $\varepsilon = 0.4$ curve. Here 0.1045 and 0.1158 are the dimensionless efficiencies of these curves in figure 14, and the factor of 2 is due to the change in ε. Thus, reducing ε decreases the cost of fluid production and increases the mechanical energy expended, but also allows a more efficient wave profile to be chosen from the two-parameter family of dimensionless optimal solutions.

Our use of higher-order corrections to deform the attainable region of the parameter space is intended only to provide intuition – by the time these corrections are large enough to dominate the calculation, there is no reason to trust them as solutions of the Stokes equations. A better procedure would be to use a finite element method to compute the swimming speed and efficiency of these curves; we expect that if this were carried out, the result would be qualitatively similar to what we saw above using higher-order corrections to the lubrication model. Finally, one would actually like to *optimize* the curves in the regime where lubrication theory fails using the full Stokes equations as the underlying fluid model. It should be possible to do this using a variant of the spline minimization algorithm described in § 4.1. The adjoint problems (Pironeau 1984) that must be solved to compute variations in $F(1, 0)$, $F(0, 1)$, $P(1, 0)$ and $P(0, 1)$ in § 2.1 with respect to perturbations of the curves turn out to be the Stokes equations again with similar boundary conditions; however, it is difficult to avoid aggressive line searches in the spline space that introduce oscillations in $h(x)$ that cannot be resolved accurately by the finite element method.

5. Discussion

In summary, we have investigated optimal shapes for sheets swimming near boundaries that maximize speed and efficiency while minimizing arclength and fluid loss. We also considered several simplified versions of the problem to gain intuition, study singular limits, and highlight generic difficulties such as non-existence

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

of solutions, multiple local extrema, and optimal solutions that leave the realm of validity of the model.

In the first of these simplified variants, we found that solutions that maximize speed subject to arclength and fluid volume constraints approach inverted parabolas separated by self-similar kinks as the ratio A_0/ℓ_0 approaches the critical value $1/\sqrt{6}$. These solutions are interesting because they show that the lubrication approximation can be used successfully even in the limit of fairly sharp profiles as long as the gap thickness is small in the region where the curvature is large. Unfortunately, the power required to maintain this maximum speed diverges as $V \to 3/2$, so although lubrication theory remains valid for all the wave profiles in this one-parameter family, a choice with V around 1 or 1.2 better balances speed and efficiency.

Next we considered maximizing efficiency without regard to speed. Here we found that in order to approach the optimal efficiency, the wave profile must approach a constant function, which violates the arclength constraint. Furthermore, as the optimal efficiency is approached, the swimming speed goes to zero. Thus we need to optimize swimming speed and efficiency simultaneously to obtain a physically relevant solution – maximizing either variable separately drives the other to zero.

Finally, we considered the full problem involving speed, efficiency, arclength and fluid volume (or fluid loss). We found that any of these variables can be considered the objective function with the others treated as constraints, and settled on the problem of minimizing arclength to find the smoothest wave profile to achieve a given goal of speed, efficiency and fluid loss. (A smooth wave profile is desirable to avoid singularities in the solution of the Stokes equations, and for the lubrication model to remain valid.) Scale invariance of the lubrication model allows us to reduce the dimension of the parameter space by absorbing the fluid loss constraint into the vertical length scale H of the physical problem. This normalizes the remaining two-parameter family of optimal solutions (with speed and efficiency as parameters) to have the same cost in terms of energy associated with fluid production, making mechanical efficiency comparisons meaningful.

In studying this two-parameter family, we encountered an interesting dilemma: optimization within lubrication theory drives us out of its realm of applicability. As the speed and efficiency constraints approach their maximal values, the optimal solutions develop self-similar, cusp-like singularities that cause the error in the lubrication approximation to grow without bound for any fixed choice of $\varepsilon = H/W$. This problem disappears as soon as we incorporate higher-order corrections into the lubrication model. Doing this causes the region of attainable speeds and efficiencies to fold over itself, penalizing cusp-like singularities and causing non-singular wave profiles to emerge as those that strike the optimal balance between speed, efficiency and fluid loss. However, once we leave the realm of lubrication theory, the problem becomes much more complicated: the parameter space becomes genuinely three-dimensional and the variational problem can no longer be solved using ODE methods.

To the best of our knowledge, there are currently no data from mechanical systems that we can use to test our predictions. Both the swimming LCEs and RoboSnail were developed as proof-of-concept designs and were not optimized for speed or efficiency. Hence both use waves that are close to sinusoidal as such shapes are relatively simple to generate mechanically. It is even more challenging to compare these results with biological systems as quantitative measurements are difficult to come by. Furthermore, it is not at all obvious that biological systems are optimized and, even if they are, under what constraints. However, despite these difficulties, there is some evidence to

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

suggest that profiles consistent with those calculated in this study are indeed observed in nature. One type of structure that arises repeatedly in our calculations consists of arcs connected by sharper (but still blunt) cusps – the details and scaling of the arcs and cusps may differ depending on the objective function and constraints, but this general form of smooth curves separated by relatively sharp "narrow gap" regions appears to be robust. Trueman (1975, p. 33) reproduces shape data taken from the underside of a crawling snail's (*Patella*) foot. Although the data are noisy, the one striking feature of the profile that emerges is wide, smooth arcs separated by sharper cusps. (Note that the data in Trueman's plot are inverted with the ventral (bottom) of the snail above and the dorsal (top) of the snail below; i.e. the cusps are pointing in a direction consistent with our predictions.) While this comparison is admittedly inconclusive and, at best, qualitative, it is encouraging and suggests that further study is warranted. Of particular interest would be the robustness of this arc/cusp structure – does it persist when other biologically relevant cost functions and constraints are applied? It is also interesting that the optimal solutions in our model turned out to be symmetric about their extrema, a situation that could change if the rheology of the fluid were more complicated.

In this study, we have presented only the first steps in optimizing low-Reynolds-number swimmers activated at the surfaces of thin films. It is hoped that these results can be used as guidelines in designing efficient mechanical swimmers and crawlers as well as increase our understanding of biological systems. The latter, understanding optimization in biology, is a much more ambitious goal as appropriate cost functions, which can often be well-defined in engineering systems, are not at all obvious in biological counterparts. Furthermore, in optimizing mechanical shapes, we have only scratched the surface of the space of relevant parameters; equally relevant questions, such as optimizing material properties of the fluid and including the cost of bending and stretching the foot, remain largely unexplored.

J. W. was supported in part by the Director, Office of Science, Computational and Technology Research, US Department of Energy under Contract No. DE-AC02-05CH11231, and by the National Science Foundation through grant DMS-0101439. A. E. H. gratefully acknowledges support from the National Science Foundation through grant CCF-0323672.

Appendix A. Swimming speed and power dissipation

In this section, we derive formulae for the driving force F and power $V_1 P$ required to maintain the boundary conditions illustrated in figure 1 for the Stokes equations. Recall from (2.2) that we have defined

$$P(V_1, V_2) = \int_{\Gamma_1} \hat{\boldsymbol{t}} \cdot \boldsymbol{\sigma} \hat{\boldsymbol{n}} \, \mathrm{d}s, \qquad F(V_1, V_2) = -\int_{\Gamma_1} \hat{\boldsymbol{x}} \cdot \boldsymbol{\sigma} \hat{\boldsymbol{n}} \, \mathrm{d}s = \int_{\Gamma_2} \hat{\boldsymbol{x}} \cdot \boldsymbol{\sigma} \hat{\boldsymbol{n}} \, \mathrm{d}x, \tag{A 1}$$

where $\boldsymbol{\sigma}$ is the Newtonian stress tensor, $\mathrm{d}s = \sqrt{1 + h_x^2}\, \mathrm{d}x$, and $\hat{\boldsymbol{t}}$, $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{n}}$ are unit vectors oriented as shown in in figure 1:

$$\boldsymbol{\sigma} = -p\boldsymbol{I} + \mu \begin{pmatrix} 2u_x & u_y + v_x \\ u_y + v_x & 2v_y \end{pmatrix}, \quad \hat{\boldsymbol{t}}\big|_{\Gamma_1} = \frac{(1, h_x)}{\sqrt{1 + h_x^2}}, \quad \hat{\boldsymbol{n}}\big|_{\Gamma_1} = \frac{(-h_x, 1)}{\sqrt{1 + h_x^2}}, \quad \hat{\boldsymbol{n}}\big|_{\Gamma_2} = (0, -1). \tag{A 2}$$

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

Substituting (A 2) into (A 1) and using the continuity equation $u_x + v_y = 0$, we obtain

$$P(V_1, V_2) = \int_{\Gamma_1} \mu \frac{-4u_x h_x + (u_y + v_x)(1 - h_x^2)}{\sqrt{1 + h_x^2}}\, \mathrm{d}x, \tag{A 3}$$

$$F(V_1, V_2) = \int_{\Gamma_1} \left[-ph_x + \mu \left[2u_x h_x - (u_y + v_x) \right] \right] \mathrm{d}x = -\int_{\Gamma_2} \mu(u_y + v_x)\, \mathrm{d}x. \tag{A 4}$$

Here x and y subscripts denote partial derivatives. We then use the no-slip boundary conditions

$$\left.\begin{aligned} u(x, h(x)) &= V_1 \left(1 + h_x^2\right)^{-1/2}, \\ v(x, h(x)) &= V_1 h_x \left(1 + h_x^2\right)^{-1/2}, \end{aligned}\right\} \tag{A 5}$$

to conclude that, on Γ_1,

$$\left.\begin{aligned} u_x &= -(u_y + \kappa V_1) h_x, \\ v_x &= \kappa V_1 - v_y h_x = \kappa V_1 + u_x h_x = \kappa V_1 (1 - h_x^2) - u_y h_x^2, \end{aligned}\right\} \tag{A 6}$$

where $\kappa = h_{xx}(1 + h_x^2)^{-3/2}$ is the curvature of the interface. Substituting these into (A 3) and (A 4) gives

$$P(V_1, V_2) = \int_{\Gamma_1} \mu(u_y + \kappa V_1) \left(1 + h_x^2\right)^{3/2} \mathrm{d}x, \tag{A 7}$$

$$F(V_1, V_2) = -\int_{\Gamma_1} \left[ph_x + \mu(u_y + \kappa V_1) \left(1 + h_x^2\right) \right] \mathrm{d}x = -\int_{\Gamma_2} \mu u_y\, \mathrm{d}x. \tag{A 8}$$

The terms involving κ are derivatives of periodic functions and vanish when integrated, thus

$$P(V_1, V_2) = \int_{\Gamma_1} \mu u_y \left(1 + h_x^2\right)^{3/2} \mathrm{d}x, \tag{A 9}$$

$$F(V_1, V_2) = -\int_{\Gamma_1} \left[ph_x + \mu u_y \left(1 + h_x^2\right) \right] \mathrm{d}x = -\int_{\Gamma_2} \mu u_y\, \mathrm{d}x. \tag{A 10}$$

These expressions are used frequently throughout the manuscript to evaluate swimming speed and mechanical power consumption.

Appendix B. Numerical methods

In this Appendix we will fill in some of the essential details that were omitted when we described our numerical optimization methods in § 4.

B.1. *Direct minimization using periodic cubic splines*

The key to the algorithm described in § 4.1 is finding an efficient way to compute the gradient of the discretized system. Let us denote the spline basis functions by φ_i. Each φ_i is a periodic piecewise cubic C^2 function satisfying $\varphi_i(x_j) = \delta_{ij}$ at the knots x_j. Each $q \in \mathbb{R}^n$ is associated with a unique spline function $h_q(x) = \sum_{i=1}^n q_i \varphi_i(x)$, which satisfies $h_q(x_i) = q_i$. Our goal is to minimize the function

$$f(q) = G[h_q], \tag{B 1}$$

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

where $G[h]$ was given in (3.5). The gradient $g = \nabla f$ is given by

$$g_i = \int_0^1 \frac{\delta G}{\delta h} \frac{\partial h_q}{\partial q_i} \, \mathrm{d}x = \int_0^1 \frac{\delta G}{\delta h} \varphi_i \, \mathrm{d}x. \tag{B 2}$$

For computational efficiency, we evaluate this integral in two steps: first, we integrate $\delta G/\delta h$ against the compactly supported B-splines (de Boor 2001), which span the same space as the φ_i; then we solve a tridiagonal system to obtain all the g_i at once. This reduces the cost of a gradient evaluation from $O(n^2)$ to $O(n)$, where n is the number of knots. To compute $\delta G/\delta h$ in (B 2), we evaluate (2.13) and (3.5)–(3.8) using a 30-point Gaussian quadrature rule on each spline segment to compute the integrals to machine precision. In practice, we define $h_q = \left(\sum q_i \varphi_i\right)^{-1}$ to prevent the BFGS algorithm from doing an aggressive line search which causes h to become negative.

We now describe our choice of spline knots and the starting guess on the coarsest grid. The solution requires more resolution in the region where $h(x)$ is close to zero because the integrals I_k involve negative powers of h. We use the interval $[-1/2, 1/2]$ as the computational domain and choose a knot spacing which is clustered at the origin:

$$x_i = \begin{cases} -3(i/n - 1/2)^2 - 2(i/n - 1/2)^3, & i = 0, 1, \ldots, n/2, \\ 3(i/n - 1/2)^2 - 2(i/n - 1/2)^3, & i = n/2, \ldots, n. \end{cases} \tag{B 3}$$

Since the Euler–Lagrange equation (3.9) is invariant under translation $x \to x + a$ and reflection $x \to -x$, solutions will be translationally invariant and symmetric about their extrema. Our goal is to find the solution (among all translations) which has a minimum at $x = 0$, where the resolution of the grid is highest. Suppose h is an even function. Since $\delta G/\delta h$ in (3.7) depends on x only through its dependence on h and h_{xx}, it will also be an even function of x. Since the knots are symmetric about the origin, the same is true at the discrete level: if $q_{n/2+i} = q_{n/2-i}$ for $1 \leqslant i < n/2$, then the gradient g also has this property. The BFGS Hessian update formula (Nocedal & Wright 1999) respects this symmetry; hence, by choosing a starting vector q such that h_q is symmetric about $x = 0$ and has a minimum there, we expect the final answer (and intermediate line search results) to have this property as well. Numerical roundoff can cause a slight drift in symmetry in the last few digits of the solution, but each time we refine the mesh, we symmetrize the solution before using it as a starting guess; on the finest meshes (when the starting guess is very good), the solution tends to be exactly symmetric in spite of roundoff error. On the first iteration, we use B-splines to smooth out a vertical translation of the critical parabola discussed in § 3 and use it for the initial guess.

B.2. *Computing the Jacobian for the Levenberg–Marquardt method*

In § 4.2 we use the Lebenberg-Marquardt method to vary the parameters $q = (b, I_1, I_2, I_3, I_4) \in \mathbb{R}^5$ in the Euler–Lagrange equation (4.3), which has the form

$$h'' = f(h, I_1, \ldots, I_4), \qquad h(0) = b, \qquad h'(0) = 0, \tag{B 4}$$

until the constraints $r(q) = 0$ are satisfied, where

$$r_0 = h'(1/2), \qquad r_{\{k=1,2,3\}} = \frac{2}{I_k} \int_0^{1/2} h^{-k} \, \mathrm{d}x - 1, \qquad r_4 = \frac{2}{A_0} \int_0^{1/2} h \, \mathrm{d}x - 1. \tag{B 5}$$

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

To do this, we must be able to compute the Jacobian $J = \nabla r = (\delta r_i/\delta q_j)$. Let us define the functions

$$g_0(x) = \frac{\partial}{\partial b} h(x), \qquad g_{\{k=1,2,3,4\}} = \frac{\partial}{\partial I_k} h(x). \tag{B 6}$$

It is well known (Coddington & Levinson 1984) that since h satisfies (B 4), the g_k satisfy the variational equations

$$g_0'' = \frac{\partial f}{\partial h}(h, I) g_0, \qquad g_0(0) = 1, \quad g_0'(0) = 0, \tag{B 7}$$

$$g_k'' = \frac{\partial f}{\partial h}(h, I) g_k + \frac{\partial f}{\partial I_k}(h, I), \quad g_k(0) = 0, \quad g_k'(0) = 0. \tag{B 8}$$

By solving these equations, we determine the rate at which the curve h changes with respect to each component of q. Using variational calculus and the chain rule, we obtain the Jacobian

$$J = \begin{pmatrix} [g_0'(1/2),\ g_1'(1/2),\ g_2'(1/2),\ g_3'(1/2),\ g_4'(1/2)] \\ 2\displaystyle\int_0^{1/2} \begin{pmatrix} -1/I_1 h(x)^2 \\ -2/I_2 h(x)^3 \\ -3/I_3 h(x)^4 \\ 1/A_0 \end{pmatrix} [g_0(x), \ldots, g_4(x)]\,\mathrm{d}x \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & a_1 & 0 & 0 & 0 \\ 0 & 0 & a_2 & 0 & 0 \\ 0 & 0 & 0 & a_3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{B 9}$$

where $a_k = -(2/I_k^2)\int_0^{1/2} h^{-k}\,\mathrm{d}x$. We compute $r(q)$ and $J(q)$ numerically using a fourth-order Runge–Kutta method to evaluate $h(x)$ and each $g_k(x)$ at nodes spaced appropriately to resolve the solutions near the origin, where they vary most rapidly, and to compute the integrals via Gaussian quadrature rules. (The nodes are organized into 300 groups of 20, where at the coarse level the spacing grows exponentially with x, and at the fine level it is a 20-point Gaussian quadrature rule.) Our approach for solving (4.8) is similar, so we omit details.

Appendix C. Validity and limitations of lubrication theory

In this Appendix we outline the steps required to compute higher-order corrections of the lubrication approximation; we compare the swimming speed and power dissipation predicted by these expansion solutions to finite element simulations of the full Stokes' equations for a family of test cases; and we discuss recent *a priori* error estimates for such expansion solutions. Our goal is to understand how the error in the approximation depends on the curve $h(x)$, which is particularly important in this work as the most interesting profiles, namely those that develop singularities in curvature or slope as they approach touchdown, have a competing mechanism controlling the size of this error (singularity formation vs. nearness to the substrate). Moreover, optimization within lubrication theory drives us out of its realm of applicability; hence, error bounds need to be established so that we can decide when to stop optimizing.

The lubrication approximation gives the zeroth-order term in an expansion of the solution of the Stokes equations in powers of the small parameter $\varepsilon = H/W$. As explained in § 2.2, the problem is non-dimensionalized by choosing different scales for the x- and y-coordinates to keep the aspect ratio of the re-scaled domain of order

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

unity. This introduces factors of ε in the equations which are used to 'bootstrap' successive terms of the solution.

In more detail, for fixed boundary data V_1 and V_2, we may expand the velocity (u, v), the pressure p, and the fluid flux $Q = \int_0^h u \, dy$ in powers of ε, e.g.

$$u(x, y) = u^{(0)}(x, y) + \varepsilon u^{(1)}(x, y) + \varepsilon^2 u^{(2)}(x, y) + \cdots, \tag{C 1}$$

and match like powers of ε in the equations

$$p_x = \varepsilon^2 u_{xx} + u_{yy}, \qquad p_y = \varepsilon^4 v_{xx} + \varepsilon^2 v_{yy}, \qquad v_y = -u_x \tag{C 2}$$

and boundary conditions

$$(u, v)\big|_{\Gamma_1} = V_1[1 + \varepsilon^2 h_x^2]^{-1/2}(1, h_x), \qquad (u, v)\big|_{\Gamma_2} = (V_2, 0) \tag{C 3}$$

to obtain a recursion expressing $u^{(2k)}$, $v^{(2k)}$, $p^{(2k)}$, and $Q^{(2k)}$ in terms of the lower-order (known) terms. Each stage of this procedure is very similar to the zeroth-order derivation of § 2.2, except that $p^{(2k)}$ is no longer independent of y when $k \geqslant 1$ and $v^{(2k)}$ must be taken into account when we desire more than $O(\varepsilon)$ accuracy in the physical velocity $(\hat{u}, \hat{v}) = U(u, \varepsilon v)$. These expansions may also be derived using a stream function formulation

$$\psi_{yyyy} = -2\varepsilon^2 \psi_{xxyy} - \varepsilon^4 \psi_{xxxx}, \qquad u = \psi_y, \qquad v = -\psi_x, \tag{C 4}$$

from which it is easy to argue (taking the form of the boundary conditions into account) that only even-order terms appear in the expansions of u, v, ψ, Q, p, and $\omega = \varepsilon^2 v_x - u_y$ (see Wilkening 2006).

Once the expansions for u, v and p are known up to a given order, we may use them to compute expansions for the power and force required to maintain the steady motion of the boundaries. From (A 9) and (A 10), we have

$$P(V_1, V_2) = \int_{\Gamma_1} (u_y)(1 + \varepsilon^2 h_x^2)^{3/2} \, dx = P_0^{(0)}(V_1, V_2) + \varepsilon^2 P_0^{(2)}(V_1, V_2) + \cdots, \tag{C 5}$$

$$F(V_1, V_2) = -\int_{\Gamma_2} u_y \, dx \qquad\qquad = F_0^{(0)}(V_1, V_2) + \varepsilon^2 F_0^{(2)}(V_1, V_2) + \cdots. \tag{C 6}$$

The zero subscripts distinguish these expansions (with V_1 and V_2 held fixed) from the composite expansions

$$P(V_1, V_2^{(0)} + \varepsilon^2 V_2^{(2)} + \varepsilon^4 V_2^{(4)} + \cdots) = P^{(0)} + \varepsilon^2 P^{(2)} + \varepsilon^4 P^{(4)} + \cdots, \tag{C 7}$$

$$F(V_1, V_2^{(0)} + \varepsilon^2 V_2^{(2)} + \varepsilon^4 V_2^{(4)} + \cdots) = F^{(0)} + \varepsilon^2 F^{(2)} + \varepsilon^4 F^{(4)} + \cdots. \tag{C 8}$$

The self-propelled swimmer must have $F^{(2k)} = 0$ for $k \geqslant 0$, which allows us to compute

$$V_2^{(0)} + \varepsilon^2 V_2^{(2)} + \varepsilon^4 V_2^{(4)} + \cdots = -\left(\frac{F_0^{(0)}(1, 0) + \varepsilon^2 F_0^{(2)}(1, 0) + \cdots}{F_0^{(0)}(0, 1) + \varepsilon^2 F_0^{(2)}(0, 1) + \cdots} \right) V_1. \tag{C 9}$$

Finally, we obtain corrections to the swimming speed $V = (W/L)V_1 - V_2$ by expanding the arclength $L = W \int_0^1 (1 + \varepsilon^2 h_x^2)^{1/2} \, dx$ and using (C 9) for V_2. The results up to fourth

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

order (with $V_1 = 1$) are

$$V^{(0)} = \frac{6(1-\zeta)}{4-3\zeta}, \tag{C 10}$$

$$V^{(2)} = \frac{1}{10(4-3\zeta)^2}\left[(8+10\zeta)\frac{E_1}{I_1} + 48\zeta\frac{E_3}{I_3} - (44\zeta+15\zeta^2)\frac{E_2}{I_2}\right] - \frac{E_0}{2I_0}, \tag{C 11}$$

$$\begin{aligned} V^{(4)} = {} & \frac{1}{12600(4-3\zeta)^3}\left[(-10752-13440\zeta)\frac{E_1^2}{I_1^2} - 193536\zeta\frac{E_3^2}{I_3^2}\right. \\ & + (-88704\zeta - 30240\zeta^2)\frac{E_1E_3}{I_1I_3} + (-16896 + 26112\zeta - 10080\zeta^2)\frac{G_1}{I_1} \\ & + (2304\zeta - 1728\zeta^2)\frac{G_3}{I_3} + (-14304 + 1488\zeta + 6930\zeta^2)\frac{J_1}{I_1} \\ & + (91392\zeta + 60480\zeta^2)\frac{E_1E_2}{I_1I_2} + (-82944\zeta + 62208\zeta^2)\frac{J_3}{I_3} \\ & + (177408\zeta + 254016\zeta^2)\frac{E_2E_3}{I_2I_3} + (-32256\zeta - 193536\zeta^2 - 30240\zeta^3)\frac{E_2^2}{I_2^2} \\ & \left. + (59344\zeta - 17208\zeta^2 - 20475\zeta^3)\frac{J_2}{I_2} + (23808\zeta - 38016\zeta^2 + 15120\zeta^3)\frac{G_2}{I_2}\right] \\ & + \frac{E_0^2}{4I_0^2} + \frac{J_0}{8I_0}, \end{aligned} \tag{C 12}$$

and

$$P^{(0)} = \frac{12(1-\zeta)I_1}{4-3\zeta} = 2I_1V, \tag{C 13}$$

$$P^{(2)} = \frac{6I_1}{5(4-3\zeta)^2}\left[(8-10\zeta+5\zeta^2)\frac{E_1}{I_1} - (4\zeta+5\zeta^2)\frac{E_2}{I_2} + 8\zeta\frac{E_3}{I_3}\right], \tag{C 14}$$

$$\begin{aligned} P^{(4)} = {} & \frac{I_1}{1050(4-3\zeta)^3}\left[(-8064\zeta - 10080\zeta^2)\frac{E_1E_3}{I_1I_3} + (-672 - 1680\zeta - 1050\zeta^2)\frac{E_1^2}{I_1^2}\right. \\ & + (384\zeta - 288\zeta^2)\frac{G_3}{I_3}(-13824\zeta + 10368\zeta^2)\frac{J_3}{I_3} + (16128\zeta + 52416\zeta^2)\frac{E_2E_3}{I_2I_3} \\ & + (-2016\zeta - 23856\zeta^2 - 13230\zeta^3)\frac{E_2^2}{I_2^2} + (3824\zeta + 6232\zeta^2 - 6825\zeta^3)\frac{J_2}{I_2} \\ & + (6144 - 18048\zeta + 16800\zeta^2 - 5040\zeta^3)\frac{G_1}{I_1} + (7392\zeta + 11760\zeta^2 + 3150\zeta^3)\frac{E_1E_2}{I_1I_2} \\ & + (-8544 + 15648\zeta - 11550\zeta^2 + 3465\zeta^3)\frac{J_1}{I_1} \\ & \left. + (8448\zeta - 13056\zeta^2 + 5040\zeta^3)\frac{G_2}{I_2} - 32256\zeta\frac{E_3^2}{I_3^2}\right], \end{aligned} \tag{C 15}$$

where $\zeta = I_2^2/(I_1I_3)$ and

$$I_k = \int_0^1 \frac{1}{h^k}\,\mathrm{d}x, \quad E_k = \int_0^1 \frac{h_x^2}{h^k}\,\mathrm{d}x, \quad J_k = \int_0^1 \frac{h_x^4}{h^k}\,\mathrm{d}x, \quad G_k = \int_0^1 \frac{h^2h_{xx}^2}{h^k}\,\mathrm{d}x. \tag{C 16}$$

Note that each term in the formulae (C 10)–(C 15) is a weighted average of powers of h_x or hh_{xx}; for example, J_3/I_3 is the average value of h_x^4 with respect to the

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

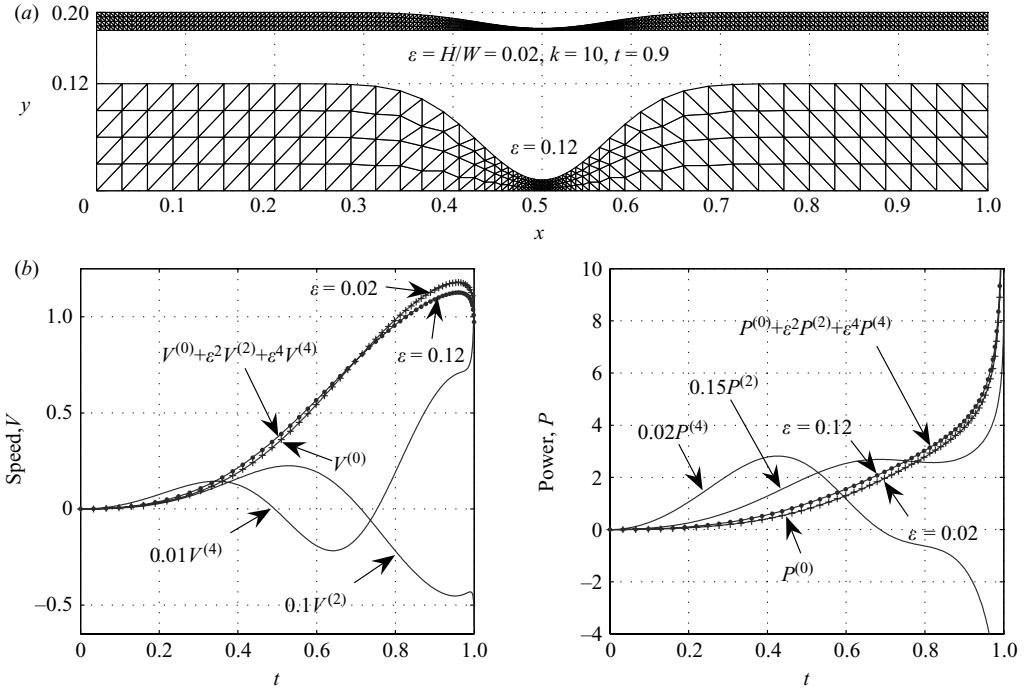


FIGURE 16. (*a*) Two of the mesh geometries (coarsened for visibility) used to compute the dependence of speed and power on t and H. (*b*) Comparison of the finite element results for speed and power (denoted by + for $\varepsilon = 0.02$ and dots for $\varepsilon = 0.12$) with the lubrication theory (solid line under the +) and the fourth-order correction (solid line under the •). Note that the zeroth-order solution has no dependence on ε; hence, the solid line (under the +) represents the zeroth-order approximation for both the $\varepsilon = 0.02$ and the $\varepsilon = 0.12$ cases. Also shown are the expansion coefficients used to compute the fourth-order correction.

weight function h^{-3}. Thus, if we have a family of curves $h(x, \delta)$ such that for all δ the slope h_x remains bounded and the curvature h_{xx} becomes large only where h is small (so that the product hh_{xx} is bounded), these terms will remain uniformly bounded in δ. This gives a hint as to how h and its derivatives should enter into rigorous, *a priori* error bounds for truncated expansion solutions of the Stokes equations in thin geometries. A detailed error analysis along these lines (including a study of the algebraic structure of the expansions to arbitrary order) has recently been carried out by one of the authors (Wilkening 2006), and will be discussed further below.

Before attempting to say anything rigorous, it is instructive to compare the predictions of these expansion solutions to finite element calculations of the full Stokes equations for a family of test cases. The wave profiles considered are of the form

$$\hat{h}(x) = Hh(x), \qquad h(x) = b + a[1 - (\sin^2 \pi x)^k] \qquad (a = t,\ b = 1 - t). \tag{C 17}$$

Here H varies between 0.015 and 0.15, t varies between 0 and 1, and $k = 10$ is held fixed. Note that when $t = 0$, $h(x) \equiv 1$ and when t approaches 1, $h(x)$ approaches touchdown with a locally parabolic profile. In the finite element calculations, we use a logically rectangular grid with nodes $\boldsymbol{x}_{ij} = (x_i, \hat{h}(x_i)j/N)$, $0 \leqslant i \leqslant M$, $0 \leqslant j \leqslant N$ (see figure 16). The mesh is triangulated by cutting quadrilaterals along their shortest diagonal. The x_i are spaced to keep the aspect ratios of the triangles as close to

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

1 as possible. In our calculations, N ranged from 48 to 96 while M ranged from 768 to 13152 (to maintain proper spacing in the x-direction). We used a variant of the First Order System Least Squares finite element method (Cai, Manteuffel & McCormick 1997) to solve the Stokes equations (2.1) using quadratic elements (with curved boundaries on the top wall) for all six variables $w=(u, v, p, \omega, \gamma, \tau)$. Here ω is the vorticity, and γ and τ are shear stresses. We solved the equations using a multigrid preconditioned conjugate gradient method. For each geometry, we solve the Stokes equations twice: once with boundary conditions $V_1=1$, $V_2=0$, and once with boundary conditions $V_1=0$, $V_2=1$. Using (A 9) and (A 10), these computations give $F(1,0)$, $P(1,0)$, $F(0,1)$, $P(0,1)$, which are sufficient to compute the swimming speed and power

$$V=\frac{W}{L}+\frac{F(1,0)}{F(0,1)}, \qquad \text{power}=P(1,0)-\frac{F(1,0)}{F(0,1)}P(0,1). \tag{C 18}$$

We normalize these solutions by multiplying the power by H/μ. Recall that the actual power is related to the non-dimensionalized power via $\hat{V}_1\hat{P}=\mu U^2(W/H)V_1P$; the finite element solutions have already been partially non-dimensionalized by setting $U=1$, $V_1=1$, $W=1$.

In figure 16, we show the results of two sequences of finite element calculations in which t varies from 0 to 1, one with $\varepsilon=0.02$ and one with $\varepsilon=0.12$. The $H=0.02$ computation agrees nicely with the lubrication prediction (shown with a solid line) even for parameter ranges in which the swimming speed exceeds the wave speed. The $H=0.12$ computation deviates visibly from the lubrication prediction, but agrees well with the fourth-order correction, which is also shown with a solid line. Note that the zeroth-order solution does not depend on ϵ and hence predicts the same velocities for both the $\varepsilon=0.02$ and the $\varepsilon=0.12$ case. The final two curves in each graph show the coefficients $V^{(2)}$, $V^{(4)}$, $P^{(2)}$, $P^{(4)}$ used to compute the fourth order correction, using (C 10)–(C 15).

In figure 17, we hold $t=0.9$ fixed and vary H from 0.015 to 0.15. As expected, we see that the error term associated with using the zeroth-order lubrication theory is $O(\varepsilon^2)$; using the second- and fourth-order corrections to the lubrication theory lead to errors which are $O(\varepsilon^4)$ and $O(\varepsilon^6)$, respectively. Numerical error in the finite element computation is responsibly for the scatter in the computed error of the fourth-order correction when ε is small; in this regime, it is more accurate to use the expansion solutions instead of finite elements.

It is remarkable that the lubrication theory works so well since the geometries we tested against the finite element simulations deviate substantially from the flat profiles envisioned in the derivation of the theory. Stated differently, ε is quite large in several of these test cases – the meshes in figure 16 are drawn to scale. This can be roughly explained by arguing that the largest boost from the pressure gradient occurs in the vicinity of the touchdown, where we expect lubrication to do well; the details of the outer flow (where lubrication may fail), are unimportant in calculating crawling velocities. The problem with this reasoning is that proximity to the substrate does not guarantee good performance if the curvature is too large (as was seen in Problem 3a above); and conversely, sometimes lubrication theory gives good predictions in the outer region as well. Moreover, by using higher-order corrections to monitor the error in the Reynolds approximation, we have implicitly assumed that these corrections improve the solution, which is not obvious (or even true in general). A more quantitative explanation of how the error in such expansion solutions depends on the geometry $h(x)$ is desirable.

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

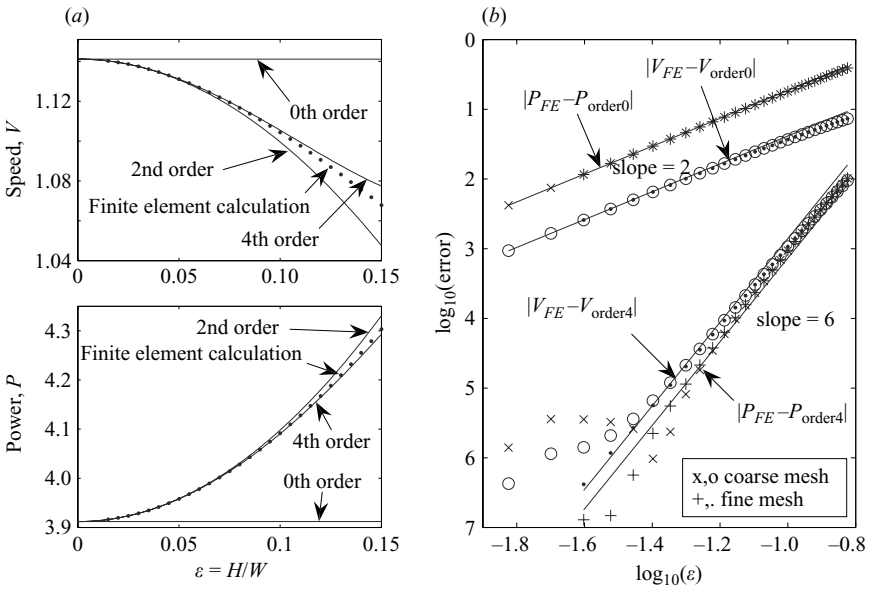


FIGURE 17. (*a*) Comparison of the lubrication approximation for speed and power and its higher-order corrections to the finite element results as H ranges from 0.015 to 0.15 with $t = 0.9$ held fixed. (*b*) Log-log plot of the error versus ε in (*a*).

It has recently been shown (Wilkening 2006) that for $m \geqslant 0$, the error in the expansion of ψ, u, v, p, ω and Q up to order $2m$ (keeping in mind that only even powers of ε occur in these expansions) is $O(\varepsilon^{2m+2})$, with a constant that depends on $h(x)$ only through its first and third inverse moments $I_k = \int_0^1 h(x)^{-k}\,\mathrm{d}x$, and via the ℓth roots of the (maximum) norms $\left\|(1/\ell!)h^{\ell-1}\partial_x^\ell h\right\|_\infty$, $1 \leqslant \ell \leqslant 2m+2$. For example, the error in vorticity $\omega_{err}^{(2m)} = \omega_{exact} - \left[\omega^{(0)} + \varepsilon^2\omega^{(2)} + \cdots + \varepsilon^{2m}\omega^{(2m)}\right]$ satisfies the following error estimate over the re-scaled domain $\Omega = \{(x, y) \,:\, 0 \leqslant x \leqslant 1,\ 0 < y < h(x)\}$:

$$\left\|\omega_{err}^{(2m)}\right\|_{L^2(\Omega)} \leqslant \sqrt{I_1}\left(|V_1| + |V_2|\right)\left[1 + \theta_m \frac{\varepsilon}{\varepsilon_m}\sqrt{\frac{I_3}{I_1}}\right]\left(\frac{\varepsilon}{\rho_m \varepsilon_m}\right)^{2m+2}, \tag{C 19}$$

where $\varepsilon_m = (\max_{1\leqslant\ell\leqslant 2m+2}\{\|(1/\ell!)h^{\ell-1}\partial_x^\ell h\|_\infty^{1/\ell}\})^{-1}$, $I_k = \int_0^1 h(x)^{-k}\,\mathrm{d}x$, and ρ_m and θ_m are constants independent of h that can be computed once and for all, e.g.:

m	0	1	2	3	10	20
ρ_m	0.197	0.210	0.252	0.288	0.232	0.137
θ_m	1.34	0.101	0.0167	0.00358	1.43×10^{-12}	1.02×10^{-34}

The variables ψ, u, εv and Q satisfy identical estimates with the L^2 norm replaced by appropriate weighted Sobolev norms (or an absolute value in the case of Q). For technical reasons, the L^2 error of the pressure seems to require an additional factor of $\max(9h_0^{-1/2}, \frac{9}{4}h_0^{-3/2})(\varepsilon_m + \varepsilon_m^{-1})^2$, where h_0 satisfies $0 < h_0 \leqslant h(x) \leqslant 1$ for $0 \leqslant x \leqslant 1$; however, our finite element experiments indicate that $\|p_{err}^{(2m)}\|_{L^2}$ is generally similar in size to $\|\omega_{err}^{(2m)}\|_{L^2}$, so it may be possible to eliminate this factor using an alternative proof.

The main content of this theorem is that the most important quantities that govern the validity of these expansion solutions are the ℓth roots of the maximum norms of

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

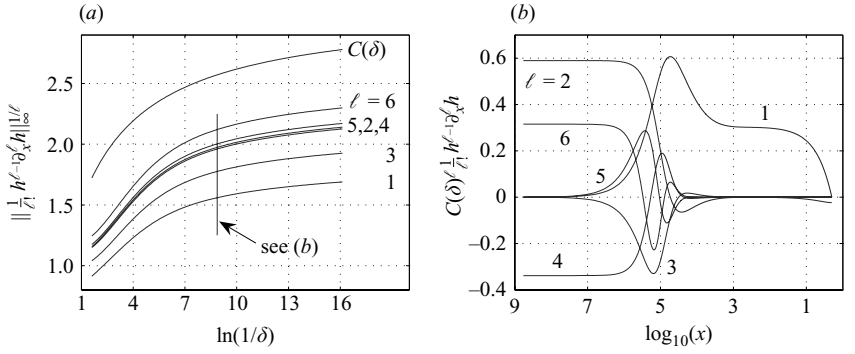


FIGURE 18. (*a*) Plot of the quantities in (C 20) vs. $\ln(1/\delta)$ together with the common bound $C(\delta)$ defined after (C 20) for the optimal solutions of Problem 1 in Section 3. (*b*) Plot of the functions $C(\delta)^{-\ell}(1/\ell!)h(x)^{\ell-1}\partial_x^\ell(x)$ for $1\leqslant\ell\leqslant 6$ over the range $10^{-9}\leqslant x\leqslant 0.5$ with $\delta=\delta_{100}=1.41\times10^{-4}$ held fixed. The ℓth roots of the absolute maxima in this plot correspond to the values at the cross-section shown in (*a*), divided by $C(\delta)$.

the homogeneous products h_x, $\frac{1}{2}hh_{xx}$, $\frac{1}{6}h^2h_{xxx}$, etc., which enter into the definition of the ε_m. The denominator $\rho_m\varepsilon_m$ in (C 19) may be up to of as an effective radius of convergence; if $\varepsilon\ll\rho_m\varepsilon_m$, the expansion up to order $2m$ will be very accurate (as long as I_1 and I_3 are not too large). But if $\varepsilon\gg\rho_m\varepsilon_m$, the terms in the expansion tend to be very large, giving nonsensical answers. This is the case in Problem 3a above. As δ approaches zero, the optimal curve $h(x,\delta)$ becomes self-similar with a scaling $h(x,\delta)\approx\delta^2y(x/\delta^3)$, so $\varepsilon_0=\max(\|h_x\|_\infty,\ \|\frac{1}{2}hh_{xx}\|_\infty^{1/2})^{-1}$ converges to zero as $\delta\to0$, and no single choice of ε can remain small in comparison to the shrinking radius of convergence $\rho_0\varepsilon_0\to0$. The higher-order corrections are even worse since $\varepsilon_m\leqslant\varepsilon_0$ for $m\geqslant0$ and the (large!) numbers $\varepsilon/(\rho_m\varepsilon_m)$ are being raised to higher and higher powers.

By contrast, in Problem 1 above, since the solution becomes self-similar as $\delta\to0$ when the x- and y-axes are re-scaled via the same function of δ, the functions h_x, $\frac{1}{2}hh_{xx}$, $\frac{1}{6}h^2h_{xxx}$, etc. remain uniformly bounded as $\delta\to0$. Indeed, writing $h(x,\delta)\approx\delta\ln(1/\delta)\,y(x/(\delta\ln(1/\delta)))$ we see that the factors of $\delta\ln(1/\delta)$ from the powers of h are cancelled by the inverse factors arising from differentiation. To verify this claim, in figure 18 we plot the six quantities

$$\|h_x\|_\infty,\ \left\|\frac{1}{2}hh_{xx}\right\|_\infty^{1/2},\ \ldots,\ \left\|\frac{1}{6!}h^5h_{xxxxxx}\right\|_\infty^{1/6}\qquad\left(\|f\|_\infty=\max_{0\leqslant x\leqslant1}|f(x)|\right)\qquad\text{(C 20)}$$

that enter into ε_m with $m=4$ in the bound (C 19) for the optimal solutions $h(x,\delta_k)$ of Problem 1 as δ_k ranges from $\delta_0=0.2$ to $\delta_{200}=9.9\times10^{-8}$. We observe that these functions grow more slowly than $C(\delta)=[\ln(1/\delta)+1]/[.32\ln(1/\delta)+1]$, which remains bounded as $\delta\to0$. This asymptotic form would be expected from a similarity analysis in which we re-scale the x- and y-axes differently at the next order, e.g. via $h(x)\approx(\delta\ln(1/\delta)+a_1\delta)\,y\big(x/(\delta\ln(1/\delta)+a_2\delta)\big)$ with $a_1=1$, $a_2=(0.32)^{-1}$. Since the norms (C 20) appear to remain uniformly bounded as $\delta\to0$, we expect the fourth-order expansion to give an accurate measure of the error in the Reynolds approximation. Although the factors $\sqrt{I_1}$ and $\sqrt{I_3/I_1}$ in the error estimate (C 19) also depend on δ and diverge as $\delta\to0$, the norms of the solutions of the Stokes equations diverge at a similar rate. We believe the relative error of the expansion solutions remains small uniformly in δ as $\delta\to0$, but we have not carried out the analysis necessary to prove this claim.

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press

REFERENCES

Avron, J., Gat, O. & Kenneth, O. 2004 Optimal swimming at low reynolds numbers. *Phys. Rev. Lett*. **93 (18)**, 186001.

Becker, L., Koehler, S. & Stone, H. 2003 On self-propulsion of micro-machines at low Reynolds number: Purcell's three-link swimmer. *J. Fluid Mech.* **490**, 15–35.

de Boor, C. 2001 *A Practical Guide to Splines*. Springer.

Cai, Z., Manteuffel, T. A. & McCormick, S. F. 1997 First-order system least squares for the Stokes equations, with application to linear elasticity. *SIAM J. Numer. Anal.* **34**, 1727–1741.

Camacho-Lopez, M., Finkelmann, H., Palffy-Muhoray, P. & Shelley, M. 2004 Fast liquid-crystal elastomer swims into the dark. *Nature Materials* **3**, 307–310.

Chan, B., Balmforth, N. J. & Hosoi, A. E. 2005 Building a better snail: Lubricaiton and adhesive locomotion. *Phys. Fluids* **17**, 113101.

Childress, S. 1981 *Mechanics of Swimming and Flying*. Cambridge University Press.

Coddington, E. A. & Levinson, N. 1984 *Theory of Ordinary Differential Equations*. Malabar, Florida: Krieger Publishing Company.

Denny, M. W. 1980 The role of gastropod pedal mucus in locomotion. *Nature* **285**, 160.

DiLuzio, W. R., Turner, L., Mayer, M., Garstecki, P., Weibel, D. B., Berg, H. C. & Whitesides, G. M. 2005 *Escherichia coli* swim on the right-hand side. *Nature* **435**, 1271–1274.

Felderhof, B. U. 2006 The swimming of animalcules. *Phys. Fluids* **18**, 063101.

Hancock, G. J. 1953 The self-propulsion of microscopic organisms through liquids. *Proc. R. Soc. Lond.* A **217**, 96–121.

Lauga, E., DiLuzio, W. R., Whitesides, G. M. & Stone, H. A. 2006 Swimming in circles: Motion of bacteria near solid boundaries. *Biophys. J.* (to appear).

Lighthill, J. 1976 Flagellar hydrodynamics: The John von Neumann lecture, 1975. *SIAM Rev.* **18**, 161–230.

Lobaton, E. & Bayen, A. 2007 Modelling and optimization analysis of single flagellum bacterial motion. *Proceedings of the 2007 American Control Conference* pp. 455–461.

Nocedal, J. & Wright, S. J. 1999 *Numerical Optimization*. New Springer.

Oron, A., Davis, S. H. & Bankoff, S. G. 1997 Long-scale evolution of thin liquid films. *Rev. Mod. Phys.* **69**, 931–980.

Pedley, T. J. & Kessler, J. O. 1992 Hydrodynamic phenomena in suspensions of swimming microorganisms. *Annu. Rev. Fluid Mech.* **24**, 313–358.

Pironeau, O. 1984 *Optimal Shape Design for Elliptic Systems*. Springer.

Shapere, A. & Wilczek, F. 1989 Efficiencies of self-propulsion at low Reynolds number. *J. Fluid Mech.* **198**, 587–599.

Taylor, G. 1951 Analysis of swimming microscopic organisms. *Proc. R. Soc. Lond.* A **209**, 447–461.

Trueman, E. R. 1975 *The Locomotion of Soft-Bodied Animals*. Elsevier.

Vlès, F. 1907 Sur les ondes pediuses de mollosques reptateurs. *C. R. Hebd. Seances Acad. Sci.* **145**, 276.

Wilkening, J. 2006 Convergence of Reynolds' lubrication approximation and its higher order corrections on a periodic domain. *Tech. Rep.* LBNL-61019. Lawrence Berkeley National Laboratory.

https://doi.org/10.1017/S0022112008000384 Published online by Cambridge University Press