PROCEEDINGS A

royalsocietypublishing.org/journal/rspa

Research

Cite this article: Alben S. 2022 Efficient bending and lifting patterns in snake locomotion. *Proc. R. Soc. A* **478**: 20220312. https://doi.org/10.1098/rspa.2022.0312

Received: 9 May 2022
Accepted: 7 September 2022

Subject Areas:
applied mathematics, mathematical modelling, mechanics

Keywords:
optimization, sidewinding, viscoelastic, friction, robotic, bioinspired

Author for correspondence:
S. Alben
e-mail: alben@umich.edu

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.6214757.

THE ROYAL SOCIETY PUBLISHING

Efficient bending and lifting patterns in snake locomotion

S. Alben

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA

SA, 0000-0002-8020-4999

We optimize three-dimensional snake kinematics for locomotor efficiency. We assume a general space-curve representation of the snake backbone with small-to-moderate lifting off the ground and negligible body inertia. The cost of locomotion includes work against friction and internal viscous dissipation. When restricted to planar kinematics, our population-based optimization method finds the same types of optima as a previous Newton-based method. With lifting, a few types of optimal motions prevail. We have an s-shaped body with alternating lifting of the middle and ends at small-to-moderate transverse friction. With large transverse friction, curling and sliding motions are typical at small viscous dissipation, replaced by large-amplitude bending at large viscous dissipation. With small viscous dissipation, we find local optima that resemble sidewinding motions across friction coefficient space. They are always suboptimal to alternating lifting motions, with average input power 10–100% higher.

1. Introduction

Snakes have a relatively simple body geometry but can perform a wide range of locomotor behaviours. They are useful for understanding the mechanics of locomotion in terrestrial and even aquatic [1] and aerial [2] environments. Snakes are also an important source of ideas for bioinspired robots [3–8], and their limblessness has advantages for control, adaptability, and navigation in complex and cluttered environments [9–11]. Although a wide range of motions are possible, four major modes of snake locomotion—serpentine, concertina, sidewinding and rectilinear—have been described and studied most often [12–15], though the true diversity of motions is greater [4,16,17].

Many common biological snake motions such as serpentine locomotion have been modelled successfully by assuming planar motions with a simple (Coulomb)

frictional model [18–26]. Serpentine and concertina-like motions were found to be optimally efficient among general time-periodic kinematics of three-link [27] and smooth bodies [28,29]. For certain body geometries and frictional anisotropies, other motions, beyond those observed biologically, were found to be optimal [30–32].

These models assume a resistive force law, local in the velocity, which is somewhat simpler than many fluid locomotion models [33–36]. Nonetheless, even the simplest planar models with Coulomb friction are too complicated to be solved theoretically with large-amplitude motions, so the physics of many motions are not well understood, and computational models and explorations of the types of efficient dynamics over parameter space play an important role.

Of the four major modes of locomotion, non-planar kinematics feature most strongly in sidewinding [12,14,37]. References [8,37] demonstrated that sidewinding can be represented as a pair of orthogonal body waves (vertical and horizontal) that can be independently modulated to achieve high manoeuverability and ascend inclines. The other major modes are not perfectly planar in real snakes or robots but can be approximated well by bodies that remain planar. The serpentine mode often features lifting at the curvature peaks, termed 'sinus-lifting' [3,4] and modelled in [22,23] using planar shapes together with a weight distribution function. [38] used planar shapes together with the weight distribution function approach of [22,23] to model sidewinding and other three-dimensional motions. Using a travelling-wave weight distribution function that is phase shifted from a travelling-wave body curvature they found a wide range of turning, slithering and sidewinding motions. Rieser *et al*. [26] used similar techniques to model sidewinding locomotion and found higher speeds with isotropic friction than with dominant transverse friction.

Here, instead of a planar curve with a weight distribution function, we represent the snake body as a time-dependent three-dimensional space curve. Given the space curve shape, the regions of contact and lifting arise naturally from solving the dynamical equations. We prefer this direct approach because it removes the question of which three-dimensional shape, if any, would produce a prescribed weight distribution function. The weight distribution depends on the equilibrium position of the space curve under gravity and may be sensitive to slight changes in the curve's shape.

Three-dimensionality opens up new possibilities for interesting mechanisms of locomotion. Some three-dimensional motions may have higher efficiency than similar two-dimensional motions, e.g. sinus-lifting versus lateral undulation, and three-dimensional motions may be necessary for traversing uneven ground and obstacles. Some three-dimensional studies have focused on modifications of sidewinding motions: [10] showed how sidewinding waveforms can be modulated in biological and robotic snakes to move around peg obstacles and [39] developed a geometric mechanics method to find optimal contact patterns for sidewinding robotic snakes for maximum speed. Other types of three-dimensional motions have also been studied: [11] studied experimentally how snakes use a combination of lateral and vertical bending to traverse uneven terrain made up of blocks with vertical and horizontal planar surfaces; [9] showed that body compliance can help snakes traverse obstacles stably.

As in our previous planar locomotion studies [27–32], we study non-planar motions using an optimization framework. It is not feasible to describe the full range of locomotor behaviours across the space of geometrical and physical parameters. Focusing on those that are optimally efficient is more manageable, even if they do not fully describe the possibilities. The optimal solutions indicate the trade-offs between the objectives and constraints in the problem, and how they depend on the parameters [40]. The optimal solutions can provide effective strategies and suggest general mechanisms for robotic locomotion [5,39]. The relationship between the optima and biological organisms is less clear, as the model omits many physiological aspects, and the importance of mechanical efficiency to reproductive success varies widely among organisms [41]. Locomotor efficiency on sandy surfaces is often mentioned in explanations of the sidewinding gait, but other benefits—avoiding predators, minimizing contact with hot surfaces and increasing sensory perception—are also discussed [42]. Comparing the optimal solutions to biological

locomotion indicates the relative importance of mechanical efficiency in determining the choice of locomotor mode [40].

In [28,31], we used quasi-Newton optimization methods; in [28,29] we used theoretical methods applicable to small body deflections; in [30] we used a commercial solver; and in [32], we used a stochastic population-based optimization method. Here, we also use a stochastic population-based method, because it is relatively simple and turns out to be effective and robust for the problem at hand. Convergence is generically slower than for Newton-based optimization methods, but those methods may have difficulties because the normal contact and tangential Coulomb friction forces where the body meets the ground are singular (or nearly singular in the regularized model we use), which also causes singularities in derivatives of the objective function.

Section 2 describes our model of a locomoting snake—a curve in three-dimensional space with curvature and torsion that vary with arc length and time. The curve moves according to Newton's Laws, with forces due to friction and normal contact with a planar ground. Unlike our previous optimization studies, here we include the body's internal viscous dissipation in the cost function. We use implicit time stepping to evolve the body's motion. Section 3 describes our optimization method. We compute the motions of a population of locomoting bodies, select a top-performing subset, and use random perturbations of its members to form the population at the next generation. After many generations, the population converges to the vicinity of a local optimum.

Section 4 describes the computed optima and how they vary with key parameters such as the numbers of modes describing the curvature and torsion (§4a). Section 4b shows how the algorithm behaves in the special case of planar motions, with and without internal viscosity. In the latter case, good agreement with [28] is found. Section 4c then shows how the optimal motions change when lifting is allowed. With small-to-moderate lifting, the effects of frictional anisotropy and damping are studied in §4d. A special class of optima that resemble sidewinding motions are shown in §4e, and special types of analytical optima are presented in §5. Section 6 gives the conclusion.

2. Model

In previous work, we assumed that the snake body was a curve in the two-dimensional plane, with a prescribed shape given by its curvature $\kappa(s,t)$, where s is arc length and t is time. Now, we assume the body is a curve in three-dimensional space (figure 1), so the shape is given by its curvature $\kappa(s,t)$ and torsion $\tau(s,t)$ [43]. We assume that $\kappa(s,t)$ and $\tau(s,t)$ are periodic in time with period T. Given κ and τ, we can integrate the Frenet–Serret formulae

$$\frac{\mathrm{d}\hat{\mathbf{s}}}{\mathrm{d}s} = \kappa\hat{\mathbf{n}}; \quad \frac{\mathrm{d}\hat{\mathbf{n}}}{\mathrm{d}s} = -\kappa\hat{\mathbf{s}} + \tau\hat{\mathbf{b}}; \quad \frac{\mathrm{d}\hat{\mathbf{b}}}{\mathrm{d}s} = -\tau\hat{\mathbf{n}} \tag{2.1}$$

to obtain the body tangent, normal and binormal vectors $\{\hat{\mathbf{s}}(s,t), \hat{\mathbf{n}}(s,t), \hat{\mathbf{b}}(s,t)\}$, if we know these vectors at one s value, say $s=0$ (the tail). The values of these vectors at the tail can be written as the columns of a three-dimensional rotation matrix $[\hat{\mathbf{s}}(0,t)\ \hat{\mathbf{n}}(0,t)\ \hat{\mathbf{b}}(0,t)] = \mathbf{R}_z(\alpha(t))\mathbf{R}_y(\beta(t))\mathbf{R}_x(\gamma(t))$, where $\mathbf{R}_z$, $\mathbf{R}_y$ and $\mathbf{R}_x$ are the matrices for counterclockwise rotations about the z-, y- and x-axes by (Euler) angles $\alpha(t)$, $\beta(t)$ and $\gamma(t)$, respectively [43].

After integrating the Frenet–Serret formulas, we have $\hat{\mathbf{s}}(s,t) = \partial_s\mathbf{X}(s,t)$, where $\mathbf{X}(s,t) = [x(s,t), y(s,t), z(s,t)]^T$ is the position of the body. Given the position of the body at the tail, $\mathbf{X}_0(t) \equiv \mathbf{X}(0,t)$, we integrate $\hat{\mathbf{s}}$ to obtain the position all along the body,

$$\mathbf{X}(s,t) = \mathbf{X}_0(t) + \int_0^s \hat{\mathbf{s}}(s',t)\,\mathrm{d}s'. \tag{2.2}$$

To summarize, given the shape of the body (κ and τ), and the six unknowns $\{\mathbf{X}_0(t), \alpha(t), \beta(t), \gamma(t)\}$ specifying its position and orientation at the tail, we can integrate to obtain $\mathbf{X}(s,t)$. At each t, we solve for the six unknowns by enforcing six equations, the linear and angular momentum

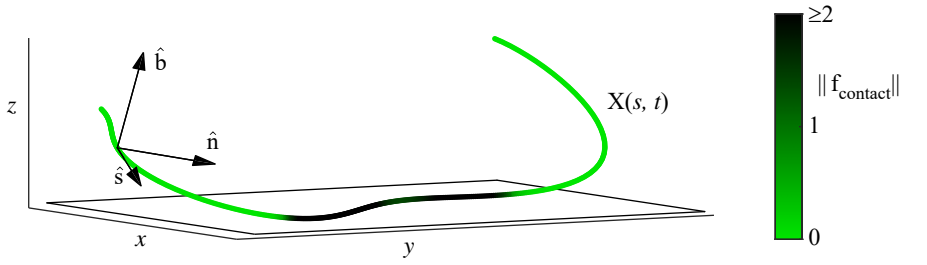

Figure 1. Schematic diagram of the three-dimensional space curve $\mathbf{X}(s,t)$ representing the snake backbone, together with the tangent, normal and binormal vectors $\{\hat{\mathbf{s}}, \hat{\mathbf{n}}, \hat{\mathbf{b}}\}$ at one location. The curve rests on the plane $z = 0$, the ground, indicated by a rectangle. The colour scale shows the magnitude of the ground contact force acting at each location. (Online version in colour.)

balances for the snake body as a whole

$$\rho \int_0^L \partial_{tt}\mathbf{X}\,\mathrm{d}s = \int_0^L \mathbf{f}_{\text{ext}}\,\mathrm{d}s; \quad \partial_t\left(\rho \int_0^L \mathbf{X}\times\partial_t\mathbf{X}\,\mathrm{d}s\right) = \rho\int_0^L \mathbf{X}\times\partial_{tt}\mathbf{X}\,\mathrm{d}s = \int_0^L \mathbf{X}\times\mathbf{f}_{\text{ext}}\,\mathrm{d}s. \tag{2.3}$$

Here, L is the body's length, ρ is its mass per unit length and $\mathbf{f}_{\text{ext}}$ is the external force on the body due to normal contact with the ground, gravity and friction

$$\mathbf{f}_{\text{ext}} = \mathbf{f}_{\text{contact}} + \mathbf{f}_{\text{gravity}} + \mathbf{f}_{\text{friction}}; \quad \mathbf{f}_{\text{contact}} = \rho g \mathbf{e}_z (H(z)\,\mathrm{e}^{-z/\delta_w} + (1-H(z))(1-z/\delta_w)) \tag{2.4}$$

and

$$\mathbf{f}_{\text{gravity}} = -\rho g \mathbf{e}_z; \quad \mathbf{f}_{\text{friction}} = -||\mathbf{f}_{\text{contact}}|| \frac{\mu_s(\partial_t\mathbf{X}_{2\text{D}}\cdot\hat{\mathbf{s}}_{2\text{D}})\hat{\mathbf{s}}_{2\text{D}} + \mu_n(\partial_t\mathbf{X}_{2\text{D}}\cdot\hat{\mathbf{s}}_{2\text{D}}^{\perp})\hat{\mathbf{s}}_{2\text{D}}^{\perp}}{\sqrt{||\partial_t\mathbf{X}_{2\text{D}}||^2 + \delta_s^2}}. \tag{2.5}$$

Here, g is gravitational acceleration and H is the Heaviside function. Instead of a hard or rigid contact force that rises in magnitude from 0 to $+\infty$ when the body penetrates the ground at $z = 0$, the contact force magnitude is very small when $z \gg \delta_w$ and very large when $z \ll -\delta_w$. We set $\delta_w = 10^{-3}L$, approximating the hard contact limit. Making δ_w much smaller than $10^{-3}L$ does not noticeably alter the body dynamics in our computations but can significantly increase the number of iterations required by our iterative solver of the nonlinear ODE system (2.3).

The frictional force model is an extension of the two-dimensional version used in several previous works [18–22,27]. Sliding friction opposes the component of velocity tangent to the ground, written here as $\partial_t\mathbf{X}_{2\text{D}}$, the projection of the body velocity in the x-y plane. The frictional force magnitude is proportional to the contact force magnitude and approximately independent of the velocity magnitude, following the Coulomb friction model. We allow for anisotropic friction that corresponds to easier sliding in certain directions. For example, snake scales allow for smaller friction when the snake slides toward the head, in the $\hat{\mathbf{s}}$ direction, instead of towards the tail or perpendicular to the body axis [19,22]. Stronger anisotropies can occur in robotic snakes due to wheels or active scales on the body surface [3,24,44]. In (2.5), the friction coefficients μ_s and μ_n are used for the components of velocity that are parallel to the ground and tangent to the backbone, or perpendicular to the backbone, respectively. Here, $\hat{\mathbf{s}}_{2\text{D}}$ is the projection of $\hat{\mathbf{s}}$ in the x-y plane, so the frictional force acts tangentially to the ground, and its magnitude decreases as $\hat{\mathbf{s}}$ becomes more vertical at a contact. Typically, $\hat{\mathbf{s}} \approx \hat{\mathbf{s}}_{2\text{D}}$ at a contact on the body's interior, because the body is smooth and almost tangent to the ground at a contact (otherwise it would penetrate the ground). At the ends of the body, $\hat{\mathbf{s}}$ may be somewhat more vertical, though in the computations we will limit the non-planarity of the snake body, so $\hat{\mathbf{s}} \approx \hat{\mathbf{s}}_{2\text{D}}$ at the ends also. In (2.5), $\hat{\mathbf{s}}_{2\text{D}}^{\perp}$ is $\hat{\mathbf{s}}_{2\text{D}}$ rotated 90° counterclockwise about the $\mathbf{e}_z$-axis. Thus $\hat{\mathbf{s}}_{2\text{D}}^{\perp}$ points in the x-y plane, in the direction transverse to the body tangent, and μ_n gives the coefficient of transverse friction. Note that $\hat{\mathbf{s}}_{2\text{D}}^{\perp} \neq \hat{\mathbf{n}}$, as $\hat{\mathbf{n}}$ points in the direction towards which the body curves, which could be vertical even when $\hat{\mathbf{s}}$ lies in the x-y plane.

The parameter δ_s in (2.5) is set to $10^{-3}L/T$, and smoothes a discontinuity in the frictional force that would occur when $||\partial_t \mathbf{X}|| = 0$, if δ_s were zero. Like δ_w, δ_s is used to make the iterative solver more robust without noticeably altering the dynamics. In [27], we noted that there are certain motions for which $\delta_s \neq 0$ is required for a solution to exist, but these are somewhat uncommon. The tangential friction coefficient μ_s takes the values μ_f or μ_b when the tangential body motion is in the forward or backward direction, respectively: $\mu_s = \mu_f H(\partial_t \mathbf{X}_{2D} \cdot \hat{\mathbf{s}}_{2D}) + \mu_b((1 - H(\partial_t \mathbf{X}_{2D} \cdot \hat{\mathbf{s}}_{2D}))$. Without loss of generality, we may assume $\mu_b \geq \mu_f$. Our focus here is on efficient body kinematics, defined similarly to previous locomotion studies [28,33,34,45,46]: among the body kinematics that result in a time-averaged centre-of-mass speed V, we find the one(s) that minimize the time-averaged power consumption. The power consumption is the rate of work done by the body against gravity, contact and frictional forces, and we also include the rate of internal viscous dissipation, assuming a linear viscoelastic model that we now describe.

So far, we have only described the centreline of the body, as a space curve with torsion τ and curvature κ. To quantify the internal viscous dissipation, we model the body as a rod of small but non-zero thickness with a viscoelastic response given by the Kelvin–Voigt model [47]. In §A of the electronic supplementary material, we show that the viscous dissipation per unit length of the rod is $\frac{1}{2}\eta_E I(\partial_t \kappa)^2$ with η_E the extensional viscosity (a function of the bulk and shear viscosities and the Poisson ratio [47]), E the Young's modulus and I the area moment of inertia.

In this work, we limit the size of τ and thus confine our attention to motions which are only moderately perturbed from planar motions. If instead we were to allow arbitrary $\tau(s,t)$, we would have essentially arbitrary three-dimensional motions which could involve complex falling and impact dynamics. Such motions are interesting but more challenging to compute accurately, and we do not address them here. With small-to-moderate τ, motions are relatively smooth in time, and $z(s,t)$ is close to time-periodic. Even with only moderately non-planar motions, it is difficult to consider the full range of possible motions, so we focus on those that are optimally efficient. If $z(s,t)$ is time-periodic (or simply has a finite long-time average), the average rate of work done against both the contact and gravity forces ((2.4) and (2.5)) is zero, because both correspond to potential energies that oscillate in time with constant long-time averages. In the simulations, we find that the rates of work done against gravity and contact forces are negligible. The time-averaged power consumption is then the sum of that due to dry friction with the ground and internal viscous dissipation

$$\langle P \rangle = \frac{1}{T}\int_0^T \int_0^L \mathbf{f}_{\text{friction}} \cdot \partial_t \mathbf{X}\, \mathrm{d}s\, \mathrm{d}t + \frac{1}{T}\int_0^T \int_0^L \eta_E I(\partial_t \kappa)^2\, \mathrm{d}s\, \mathrm{d}t. \tag{2.6}$$

We non-dimensionalize all quantities (e.g. in equations (2.3) and (2.6)) using T as the characteristic time, L as the characteristic length and $\rho g L$ as the characteristic force. Henceforth, the variables are assumed to be dimensionless, but we keep their names the same. The dimensionless versions of equations (2.3) are

$$\mathrm{Fr}\int_0^1 \partial_{tt}\mathbf{X}\, \mathrm{d}s = \int_0^1 \mathbf{f}_{\text{ext}}\, \mathrm{d}s; \quad \mathrm{Fr}\int_0^1 \mathbf{X} \times \partial_{tt}\mathbf{X}\, \mathrm{d}s = \int_0^1 \mathbf{X} \times \mathbf{f}_{\text{ext}}\, \mathrm{d}s, \tag{2.7}$$

where $\mathrm{Fr} = L/gT^2$ has been termed the Froude number [22]. As in many previous studies, we assume that Fr is sufficiently small that it can be approximated as zero. Thus we neglect the effect of body inertia, so we do not consider fast motions that involve significant body accelerations [31]. This simplifies the problem in two ways: it reduces the number of parameters under investigation, and it makes the solutions behave simply under time reparametrization. Now Fr is set to zero, and we can divide the equations (2.7) through by μ_f, so the dependence on friction coefficients is only through the ratios μ_n/μ_f and μ_b/μ_f. Having set Fr to zero, we have eliminated the acceleration terms in (2.7). The result is that body motions are invariant under reparametrization of time, if we assume that $\delta_s = 0$. In particular, if $\mathbf{X}(s,t)$ is the body motion that corresponds to $\kappa(s,t)$ and $\tau(s,t)$, then $\mathbf{X}(s,ct)$ is the body motion that corresponds to $\kappa(s,ct)$ and $\tau(s,ct)$, for any positive constant c. The invariance under time reparametrization was shown in appendix B of [27] and also occurs in many other locomotion models [48–50]. Here δ_s is sufficiently small that the rescaling

property holds to a very good approximation. We use the property as follows. We wish to find the kinematics that minimize the average power consumption among all those that achieve a given time-averaged centre-of-mass speed V. With the rescaling property, any kinematics that give a non-zero average speed can be rescaled in time so that the average speed is V. For a given motion $\mathbf{X}(s,t)$, let the norm of the centre-of-mass displacement after a period be D (positive in general). Then for this motion, we set $T = D/V$, so it has average speed V. The dimensionless average power is (2.6) divided by $\rho g L^2/T$, but is still denoted $\langle P\rangle$. We define $\tilde{P}$ to be the dimensionless average power divided by the optimal (i.e. minimal) value for a two-dimensional motion with average speed V, which was shown in [27] to be $\rho g \mu_{\min} VL$ (dimensional form) or $\mu_{\min} D/L$ (dimensionless form), where $\mu_{\min} = \min(\mu_f, \mu_b, \mu_n)$. This is the power expended by a planar body (i.e. $||\mathbf{f}_{\text{contact}}|| \equiv 1$) sliding uniformly in the direction of minimal friction. Our scaled average power is

$$\tilde{P} \equiv \frac{\langle P\rangle}{\langle P\rangle_{\text{opt.2D}}} = \frac{\langle P\rangle}{\mu_{\min} D/L} = \frac{L}{\mu_{\min} D}\int_0^1\int_0^1 \mathbf{f}_{\text{friction}} \cdot \partial_t \mathbf{X}\,\mathrm{d}s\mathrm{d}t + c_v \frac{L^2}{D^2}\int_0^1\int_0^1 (\partial_t \kappa)^2\,\mathrm{d}s\mathrm{d}t. \tag{2.8}$$

where $c_v = \eta_E I V/\rho g L^4 \mu_{\min}$ is a dimensionless constant that measures the ratio of internal viscous dissipation to external friction. The frictional and viscous power terms on the right side of (2.8) involve a time derivative and a time derivative squared, respectively, so they have the factors D^{-1} and D^{-2} to reflect the fact that the period (and hence time t, in units of the period) is scaled by D so that $V = D/T$ is constant across motions.

Using the zero-Fr rescaling property, we have reduced the dependence on viscosity and on V to the single parameter c_v. We give an order-of-magnitude estimate of c_v for a biological snake (for example, a corn snake [22]). We take $L = 30\,\text{cm}$, $g \approx 10\,\text{m s}^{-2}$, $\rho = 1\,\text{g cm}^{-3}$ and $\mu_{\min} = 0.1$ [23,51]. To estimate I, we assume a cylindrical body with thickness $h = 3\,\text{cm}$, so $I \approx 5\,\text{cm}^4$. A typical locomotion speed is $V = 10\,\text{cm s}^{-1}$. We are not aware of measurements of the effective viscosity of snake tissue (which would only approximate the true nonlinear viscoelastic behaviour), but an approximation comes from human muscle measurements [52], applied to a similar viscoelastic model of a saithe fish [53], with viscosity given as 10^4–10^5 poise. The net result is $c_v \approx 10^{-3}$–10^{-2}. Smaller viscosities of 10^1–10^2 poise were reported for human muscle and other tissues in [54], resulting in $c_v \approx 10^{-6}$–10^{-5}. Therefore, we vary c_v over a wide range, 10^{-6}–10^{-1}, to consider a wide range of possibilities.

Next, we describe an algorithm to the determine the kinematics that minimize $\tilde{P}$ among all kinematics that give locomotion at a given speed.

3. Stochastic optimization

We now describe the optimization algorithm. We write κ and τ as double Fourier–Chebyshev series

$$\kappa(s,t) = \sum_{j=0}^{N_f-1}\sum_{k=0}^{N_c-1}(A_{jk}\cos 2\pi jt + B_{jk}\sin 2\pi jt)T_k(s) \tag{3.1}$$

and

$$\tau(s,t) = (1-\mathrm{e}^{-(t/t_d)^2})\sum_{j=0}^{N_f-1}\sum_{k=0}^{N_c-1}(C_{jk}\cos 2\pi jt + D_{jk}\sin 2\pi jt)T_k(s), \tag{3.2}$$

where $T_k(s)$ is the Chebyshev polynomial of first kind of degree k. We initialize a population of 50 body kinematics, each given by $2(2N_f-1)N_c$ coefficients $\{A_{jk}, B_{jk}, C_{jk}, D_{jk}\}$ (excluding B_{0k}) with A_{jk} and B_{jk} drawn from a Gaussian distribution with standard deviation $W_{jk} = 1/j$ if $j \neq 0$, and 1 otherwise. The $1/j$ weight corresponds to piecewise continuous functions of time, a minimal type of regularity that we bias the solutions towards because we find empirically that it yields better optima [32]. C_{jk} and D_{jk} are drawn from uniform distributions on $\tau_{\text{amp}} W_{jk}[-1,1]$, where τ_{amp} is a non-negative parameter that limits the size of τ. We compute solutions to (2.7) from $t = 0$ to $t_{\text{final}} = 3$, using the second-order BDF method with time step $\Delta t = 0.005$. At each time step, we

Table 1. List of the main parameters and their values or ranges of values used in the present study.

t_{final}	Δt	Δs	δ_s, δ_w, Fr	Pop. Size	N_{gens}	p_{amp}	N_f	N_c	μ_n/μ_f	μ_b/μ_f	τ_{amp}	c_v
3	0.005	0.01	10^{-3}	50	300–600	0.1–0.3	3–8	3–9	0.1–10	1–5	0–1	10^{-6}–10^{-1}

solve the six nonlinear equations (2.7) for $\{\mathbf{X}_0(t), \alpha(t), \beta(t), \gamma(t)\}$ using Newton's method with the solution at the previous time step as an initial guess. The integrals in (2.7) are discretized using the trapezoidal rule with grid spacing $\Delta s = 0.01$. To obtain a good guess for the solution near the initial time, and to decrease the chances of jumping to other branches of solutions at later times, we use a non-zero (but small) Fr value, 10^{-3} in (2.7), and solve the problem as an initial value problem starting from zero tail position/angles and velocity. The initial condition is a good guess for the solution just after the initial time. Fr is small enough to give a very good approximation to the zero-Fr case, but provides a small amount of inertia that prevents the large accelerations that would occur with jumps to other solution branches. The exponential-in-time factor on the right side of (3.2) causes the body to ramp up from a planar shape over a short time given by $t_d = 0.2$. This improves convergence during the initial transient in which the body transitions from zero tail velocity to $O(1)$ tail velocity.

The algorithm runs for a number of generations $N_{\text{gens}} = 300$–600. In §B of the electronic supplementary material, we describe the method of selecting motions that achieve low $\tilde{P}$ simultaneously with negligible rotations about the z-axis—so the body's path is a straight line rather than a circle.

We list the values or ranges of values used here for the most important physical and numerical parameters in table 1. Next, we present the optimal solutions computed by the model and how they depend on the key parameters: N_f, N_c, μ_n/μ_f, μ_b/μ_f, τ_{amp} and c_v. To keep the presentation somewhat concise, we vary one or two parameters at a time, keeping the others fixed.

4. Computational results

(a) Numbers of modes

The numbers of Fourier and Chebyshev modes, $(2N_f - 1)$ and N_c, are important parameters in the optimization. Increasing them increases the range of possible kinematics, allowing for bending and lifting patterns that vary more sharply in space and time. The dimension of the parameter space (for κ and τ together) is $2(2N_f - 1)N_c$. It grows rapidly with N_f and N_c, which generally increases the number of local optima, and the chance of the population converging to a local optimum that greatly underperforms the global optimum [32]. We study the effect of varying N_f and N_c in §C of the electronic supplementary material. Based on these results, our approach is to gradually increase N_f and N_c from 3–6 and 3–7, respectively, as the generation number n increases, because this usually provides a lower $\tilde{P}$ than other simple choices. We also include results with N_f and N_c fixed at 4 and 5, respectively, in a few cases where this choice gives lower $\tilde{P}$ values.

(b) Planar optima

Next, we study how the algorithm performs in the case of no lifting, $\tau_{\text{amp}} = 0$, which corresponds to the class of planar motions studied in many previous works [21,22,28,32]. The most similar is [28], in which the same modal expansion of curvature was used, but with zero viscous damping only. When c_v is large, the friction term in $\tilde{P}$ becomes insignificant compared to the viscous term, but friction still plays a key role in setting the motion of the body through the dynamical evolution equations. We investigate the planar case first to show the results of the algorithm in this more familiar case. The planar case is also a benchmark against which the benefits of lifting can be assessed.

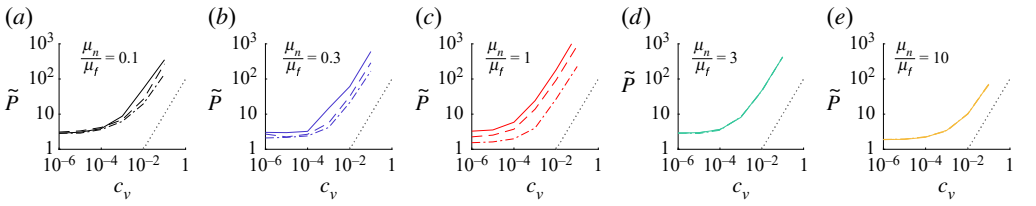

Figure 2. The dependence of $\tilde{P}$ on c_v for planar motions ($\tau_{\text{amp}} = 0$). Each panel shows data for a different μ_n/μ_f value, labelled at the top. Within each panel, three lines are given corresponding to $\mu_b/\mu_f = 1$ (solid line), 2 (dashed line) and 5 (dashed-dotted line). The dotted line shows the scaling $\tilde{P} \sim c_v$. (Online version in colour.)

In figure 2, we plot $\tilde{P}$ versus c_v for five different μ_n/μ_f (increasing from *a* to *e*), and three different μ_b/μ_f (solid, dashed and dashed-dotted lines within each panel). For each curve, $\tilde{P}$ is approximately independent of c_v near $c_v = 10^{-6}$, where viscous dissipation is negligible, and $\tilde{P}$ grows linearly with c_v near 10^{-1}, the large-c_v regime. Other aspects of these plots are discussed in §D of the electronic supplementary material.

Typical optimal planar motions are shown in figure 3. The minimal values of $\tilde{P}$ are plotted in friction coefficient space in *a* and *b*, with zero and large viscous damping respectively ($c_v = 0$). The magenta lines serve to divide friction coefficient space into six regions, each characterized by a typical motion, labelled C–H and Cv–Hv in *a* and *b*, respectively. Below these panels, each motion is shown as a set of 11 snapshots over a period of motion. The snapshots are uniformly rotated so that the (arbitrary) direction of time-averaged centre-of-mass displacement is up the page. The snapshots are shown twice, first in their true physical positions immediately adjacent to the letter label, as an overlapping set of shapes that vary from light grey to black as time increases. Then, extending rightward, the snapshots are spread out across the page with a fictitious horizontal displacement that makes it possible to see each snapshot individually (the true horizontal displacement over a period is zero).

In the two rightmost columns of *a*, $\mu_n/\mu_f = 3$ and 10, the optimal motions G and H are retrograde travelling-wave motions, also seen in [28]. The snapshots of G and H follow undulatory paths up the page. The more viscous optima Gv and Hv have larger deflection amplitudes and smaller curvatures, which increase D and decrease $\partial_t \kappa$ in the second term on the right side of (2.8), decreasing the viscous dissipation. For $\mu_n/\mu_f = 1$ and smaller, there are different types of optimal motions for $\mu_b/\mu_f = 1$ and $\mu_b/\mu_f > 1$. The isotropic case (E) is a type of concertina motion [27]—the body curls up at the front and rear, forming 'anchors' that can push the rest of the body forward. The viscous version (Ev) involves larger swinging motions and somewhat less curling, as well as more self-intersecting (which is not prevented for simplicity). As before, the curvatures are smaller while the net displacement is larger in the viscous case. Motions F and Fv exemplify the region $\mu_n/\mu_f = 1$ and $\mu_b/\mu_f > 1$—the body repeatedly flexes and unflexes as a curvature wave travels slightly forward. For $\mu_n/\mu_f < 1$, the optima with zero viscosity (C and D) are mostly perpendicular to the direction of locomotion, a low drag configuration. The large viscosity optima (Cv and Dv) have much larger degrees of flexing with a swinging motion (Cv) or a forward-propagating wave of curvature (Dv), giving a larger displacement than C and D, which is preferable at large viscosity. Additional information about the planar optima is given in §D of the electronic supplementary material.

(c) Lifting amplitude

We now consider how the optimal motions change when the body is allowed to lift off of the plane, i.e. when τ_{amp} is increased from zero. To limit the number of cases under discussion, we focus on just four choices of friction coefficient pairs $(\mu_n/\mu_f, \mu_b/\mu_f)$: (0.1, 1), (1, 1), (10, 1), (1, 5), listed along the left boundary of the colour plot at the top of figure 4. Along the bottom boundary, τ_{amp} varies from 0 to 1. The viscous damping constant c_v is set to 0. Moving from left to right in

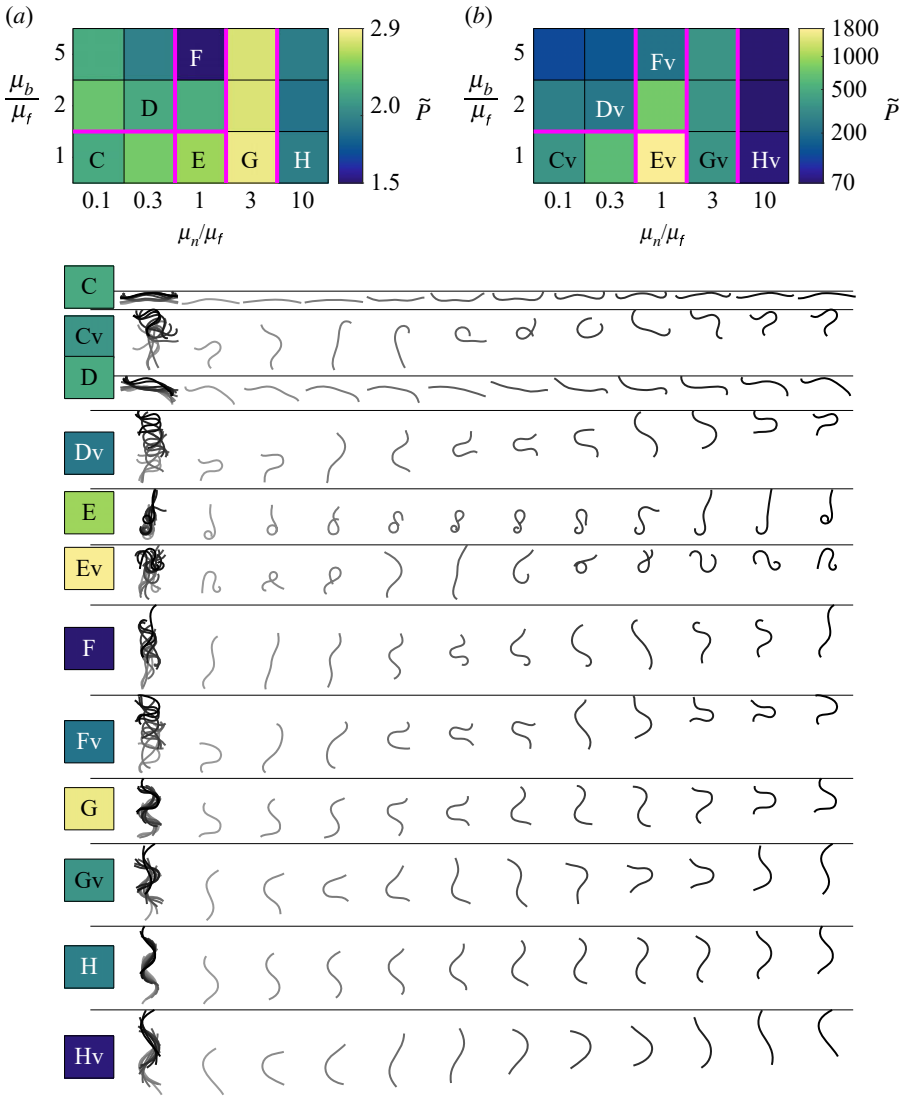

Figure 3. Values of input power $\tilde{P}$ for optimal planar motions with damping constant $c_v = 0$ (*a*) and 0.1 (*b*), across a 5-by-3 grid of values of the friction coefficient ratios μ_n/μ_f and μ_b/μ_f. In (*a*,*b*), the friction coefficient space is divided by magenta lines into six regions, where similar types of optimal motions occur. At the points C–H (*a*) and Cv–Hv (*b*), examples of these motions are shown by the sequences of snapshots in the lower portion of the figure. (Online version in colour.)

the colour plot, we see that the optimal $\tilde{P}$ drops with increasing τ_{amp} (except for the very slight increase just to the right of point G). It is reasonable that the optimal $\tilde{P}$ decreases, because the set of possible motions at a given τ_{amp} includes those at smaller τ_{amp}. Enlarging the set of possible states could cause the algorithm to find a worse local optimum, but not here. To show how the optimal motions change with lifting, we use A–L to label the subsets of cases with $\tau_{\text{amp}} = 0, 0.1,$ and 1 at each of the four friction pairs, and show the body snapshots below.

Moving from the planar motion A to the non-planar motion B, the type of motion changes completely, from a small amplitude deflection about a straight body, to a 'walking' type of motion. The snapshots are coloured according to the contact force magnitude (green-black scale above the snapshots), where lifted regions of the body are in green, and those strongly contacting the ground are in black. The colour maps at the right show, for each set of snapshots, the distribution of contact forces in the space of arc length s and time t, over one period. Motion B involves

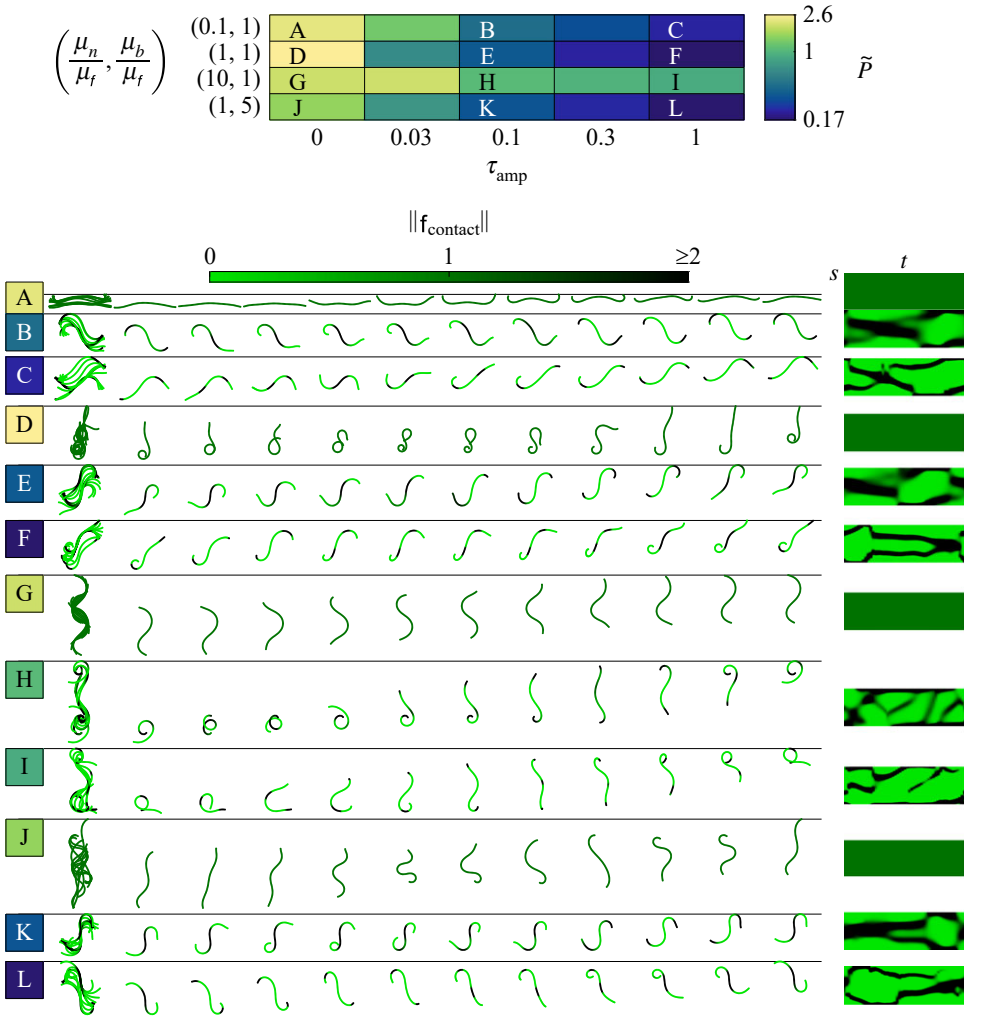

Figure 4. For zero internal viscosity ($c_v = 0$), minimum $\tilde{P}$ and corresponding motions for different lifting amplitudes τ_{amp} at four different friction coefficient ratio pairs. At the top is a colour plot showing the minimum $\tilde{P}$ found at each τ_{amp} (horizontal axis) and friction coefficient ratio pair (vertical axis). For the choices $\tau_{amp} = 0$, 0.1 and 1, the corresponding optimal motions are shown below (labelled A–L), with snapshots displaced horizontally. To the right of each row of snapshots is a colour plot showing the ground contact force density $\|f_{contact}\|$ versus arc length s and time t, over one period of time. (Online version in colour.)

small perturbations about an s-shaped curve. The contact map on the right shows that motion B involves two phases. The first phase is shown by two crossing diagonal black bands. Thus two contact regions move from the ends to the middle and back to the ends. In this phase, the snapshots show that the ends are lifted and moved in the forward direction. In the second phase, the black regions are confined to the ends. Here, the middle part of the body is lifted and moved in the forward direction. From A to B, $\tilde{P}$ is reduced by about a factor of five. Motion C, with τ_{amp} increased to 1, is similar to B, though the bands of lifting are sharper and $\tilde{P}$ is reduced by an additional factor of almost two. In principle, alternately lifting parts of the body and moving them only when they are off the ground can reduce the power done against friction to very small values. Motions D, E and F show the sequence of optima with increasing τ_{amp} when friction is isotropic. The snapshots and contact maps of E and F are fairly similar to those of B and C. The mean body configurations in E and F are more aligned with direction of motion than B and C,

probably because of the increase in μ_n/μ_f. The mean power $\tilde{P}$ is reduced by about 25% from B/C to E/F, perhaps due to slight differences in the relative amounts of tangential and normal motions.

Motions G, H and I correspond to $\mu_n/\mu_f = 10$ and $\mu_b/\mu_f = 1$. The planar motion (G) is a basic lateral undulation, a retrograde travelling wave. The lifted motions (H and I) have snapshots that start with the body curled up. Then part of the body straightens and moves forward. It is mostly lifted off the ground except at one end, its more forward point, that is in contact with the ground, and slides forward along the ground. In the last few snapshots (in H and I), the forward end curls up, drawing the rest of the body forward behind it. The overall motion is similar to a concertina motion, with an anchor formed by the curled region, and the straight region extended or retracted forward, mostly lifted, except at one end. From G to H to I, $\tilde{P}$ drops by 46% and then by 17%, smaller reductions from lifting than at the other friction pairs. Motions H and I combine lifting with sliding mainly in the tangential direction, in contrast to B/C and E/F, where the body is oriented transverse to the direction of motion, so the normal friction coefficient applies more strongly for the sliding that occurs. None of the optima computed with $\mu_n/\mu_f = 10$ and $\mu_b/\mu_f = 1$ and non-zero τ_{amp} resemble 'sinus-lifting'—lateral undulation with lifting at the peaks of the sinusoidal wave. We study this aspect further in §E of the electronic supplementary material.

Motions J, K and L apply with $\mu_n/\mu_f = 1$ and $\mu_b/\mu_f = 5$. K and L are very similar to E and F, both in the snapshots' shapes and in the values of $\tilde{P}$ (those for K/L are within a few percent of those for E/F). There is little backward motion, so the difference in μ_b/μ_f is not very important.

In some cases (D, H and I), there is apparently self-intersection at certain times. This partly due to the two-dimensional projection of the images, and in any case is not a major concern because with slight modifications the motions avoid self-intersection.

In figure 5, we consider the same quantities but instead of zero viscosity, we have large viscosity ($c_v = 0.1$). As with zero c_v, there is a large change in motions and $\tilde{P}$ values as τ_{amp} increases from 0 to 0.03 or 0.1, and a smaller change as τ_{amp} increases further to 1. This is particularly true for the friction coefficient pairs (0.1,1) and (1,1), and somewhat less so at (10,1) and (1,5). As in figure 4, motions B and C mainly involve an alternating pattern of lifting and moving of the outer regions and of the central region. Here, however, the body curves more and obtains a larger displacement over one period. The main advantage probably is to increase D in the c_v-term in (2.8), thereby decreasing $\tilde{P}$. Motions E and F are fairly similar, though interestingly E resembles C more than B, and F resembles B more than C, despite the different τ_{amp} values. Motion H qualitatively resembles G even though G has no lifting. H has an alternating 2-2 contact pattern as it flexes symmetrically to the left and to the right. The main contacts are at the head and midbody when the tail swings forward, and at the tail and midbody when the head swings forward. $\tilde{P}$ is reduced by about 30% from G to H. Motion I is similar to H but less symmetrical, and gives a further 14% reduction in $\tilde{P}$. With lifting, the friction coefficient pair (10,1) has the highest power with zero c_v (H–I at the top of figure 4) but the lowest power with large c_v (figure 5). Motions K and L are very different from each other, but have about the same $\tilde{P}$, about a factor of three less than that of the non-lifting optimum J. K involves large bending to the left and right, with some similarities to H and I. L has bending to one side only, with two almost-fixed contact regions near the head and the midbody. It resembles J in that it is a repeated bending and unbending motion, but to one side only. A combination of backward and normal friction acts at each contact region at different times, pushing the body forward.

For the first three of the friction coefficient pairs, the smallest $\tilde{P}$ is obtained at $\tau_{\text{amp}} = 0.3$, and for the fourth, the smallest value occurs at $\tau_{\text{amp}} = 1$, but it is within 7% of the values at $\tau_{\text{amp}} = 0.1$ and 0.3. Because friction is negligible in $\tilde{P}$ at large c_v, a large degree of lifting is less important in terms of avoiding frictional power dissipation. But a moderate degree of lifting is still much better than small lifting, because it changes the correspondence between body bending and locomotion. It allows for a smaller viscoelastic (bending) dissipation for a given average speed of locomotion.

To summarize, at both zero viscous damping and large viscous damping there are large changes in the optimal motions and large improvements in $\tilde{P}$ when τ_{amp} increases from 0 to 1. Most of the changes occur in the increase from 0 to 0.1, with smaller changes over the increase from 0.1 to 1. The optimal motions in the latter range are characteristic of the regime of moderate

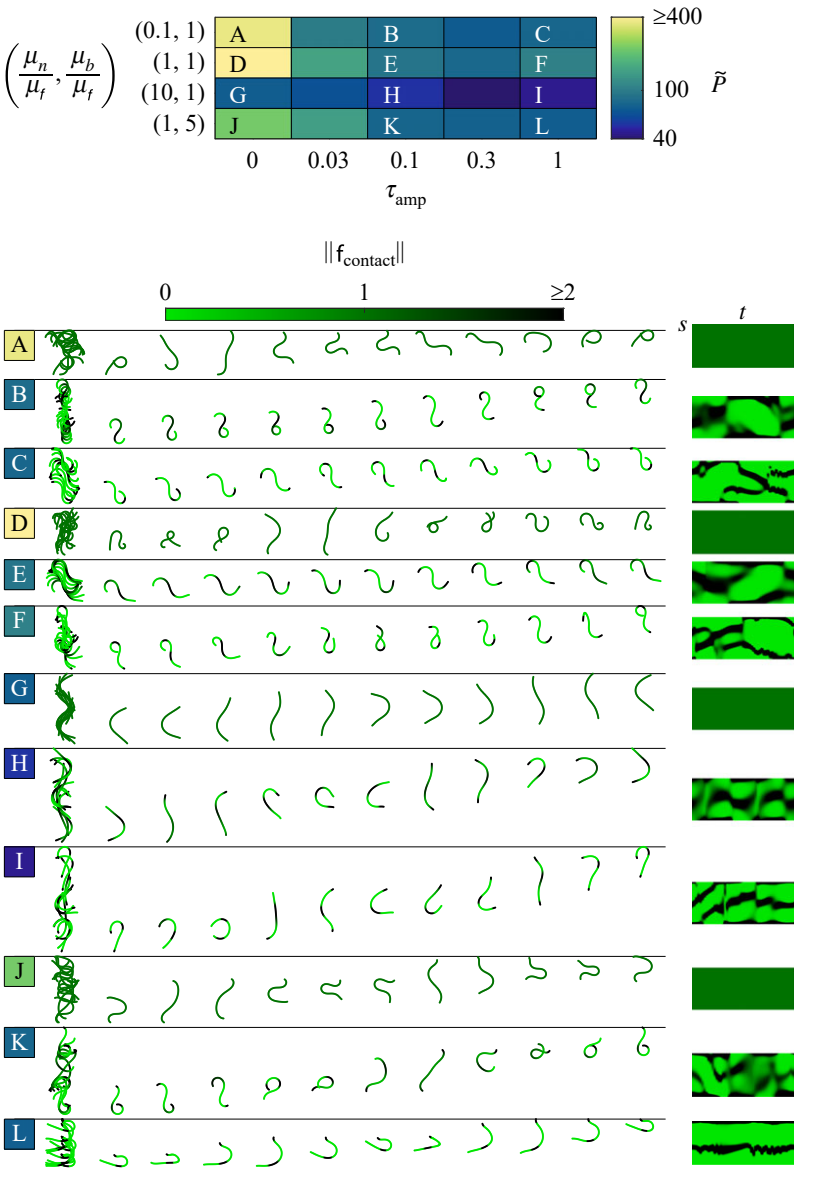

Figure 5. For a relatively large internal viscosity ($c_v = 0.1$), the same quantities as in figure 4. (Online version in colour.)

lifting, in which the body is mostly extended in the x-y plane with a z-extent that is much smaller, but large enough to lift completely off the ground at most points. In some cases, a lower $\tilde{P}$ is obtained at $\tau_{amp} = 0.3$ than at 1. In most cases, the optimization routine is faster and more robust at $\tau_{amp} = 0.3$ than at 1. At $\tau_{amp} = 1$, the lifting may be large enough that the body tips over due to gravity, and a tumbling motion occurs which is not resolved by the time-stepping algorithm. In order to avoid such cases while examining the effects of the friction coefficients in more detail, we therefore fix τ_{amp} at 0.3, which gives a good representation of motions with a moderate amount of lifting. We are not aware of biological or robotic motions that closely resemble the optima presented here, but we discuss possible connections further in §5.

(d) Effects of frictional anisotropies and damping

We now consider the optimal motions and $\tilde{P}$ values across the 5-by-3 grid of friction coefficient ratios. We use 10 random initializations with each of the (3-6,3-7) and (3-6$_w$,3-7$_w$) mode choices

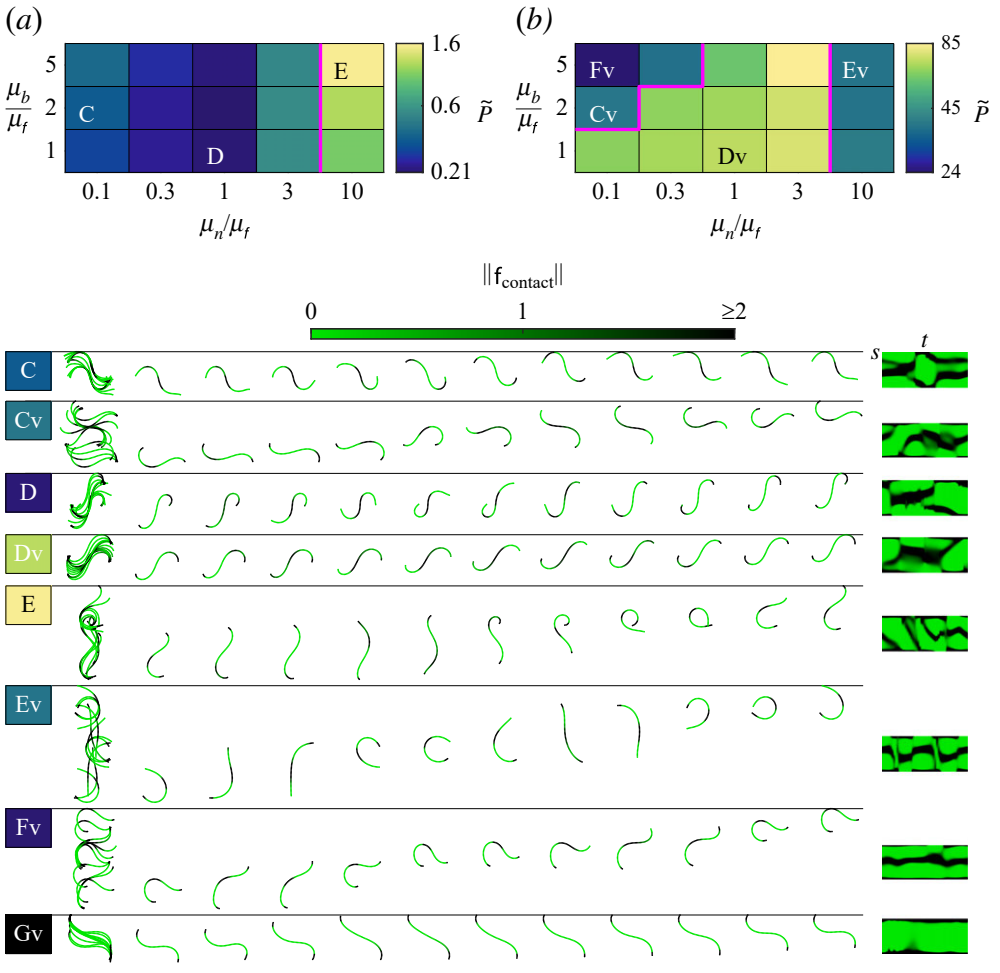

Figure 6. Values of input power $\tilde{P}$ for optimal motions with lifting amplitude $\tau_{\text{amp}} = 0.3$ and damping constants $c_v = 0$ (*a*) and 0.1 (*b*), across a 5-by-3 grid of values of the friction coefficient ratios μ_n/μ_f and μ_b/μ_f. In (*a*,*b*), the friction coefficient space is divided by magenta lines into different regions where similar types of optimal motions occur. At the points C–E (*a*) and Cv–Fv (*b*), examples of these motions are shown by the sequences of snapshots in the lower portion of the figure. (Online version in colour.)

(see §C of the supplementary information). We fix τ_{amp} at 0.3, and vary c_v from 0 to 0.1 (large). In figure 6, we plot the optima at the two extreme c_v only, 0 in *a* and 0.1 in *b*. We have also computed results at intermediate values, and we find that around $c_v = 10^{-3}$ they smoothly transition between the behaviours at the extremes. With zero c_v (*a*), there are essentially two types of optima, one for $\mu_n/\mu_f \leq 3$ and another at the largest $\mu_n/\mu_f = 10$. The two regimes are divided by the magenta line in *a*. For $\mu_n/\mu_f \leq 3$ and all μ_b/μ_f, the optimal solution is an s-shaped body with alternate lifting of the body ends and the body middle, similar to those shown in figure 4*b*,*c*,*e*,*f*,*k*, and *l*, with modest changes depending on the friction coefficient ratios. Examples are shown by the motions marked C and D in figure 6. The values of $\tilde{P}$ are almost independent of μ_b/μ_f at $\mu_n/\mu_f = 1$ and 3, but have a noticeable variation at $\mu_n/\mu_f = 0.1$ and 0.3. At $\mu_n/\mu_f = 10$, the optimum (E) is a combination of undulation and curling/uncurling, together with lifting, as in figure 4*h*,*i*.

With large viscosity, $c_v = 0.1$ (*b*), the friction coefficient space divides into three types of optimal motions rather than two. At $\mu_n/\mu_f = 1$ and 3 the alternating lifting motion (shown by Dv) is again

optimal. At $\mu_n/\mu_f = 10$, a symmetric bending motion with alternating pairs of contacts (Ev) is optimal, as in figure 5*h*,*i*. At small μ_n/μ_f and $\mu_b/\mu_f > 1$, different optima are seen, shown by Cv and Fv. The first four snapshots of Cv show the middle part of the body lifted and placed forward. Snapshots 4–7 show curling and uncurling of the right side of the body, which pushes the left side of the body forward. Snapshots 8–11 show curling and uncurling of the left side of the body, which pushes the right side of the body forward. Motion Fv is a similar pattern of curling and uncurling, but by the left side of the body only. Curling pushes the left side forward, and uncurling pushes the right side forward. Meanwhile, there are three almost fixed contact regions at the two ends and the middle of the body, which resembles the contact pattern at certain times in Cv, but not others.

At the bottom of the figure is an example of a different type of optimal motion, labelled Gv. The friction coefficients are the same as for E and Ev ($\mu_n/\mu_f = 10$, $\mu_b/\mu_f = 5$), but $c_v = 10^{-4}$, intermediate between the values in *a* and *b*. The contacts occur exclusively at the two ends throughout the motion. Because these are oriented in the tangential direction, frictional dissipation occurs mainly with the minimal drag coefficient μ_f. There is also backward movement at the contacts, but it is much smaller than the forward movement because $\mu_b/\mu_f = 5$.

To summarize, alternating lifting motions predominate across friction coefficient space both with small viscosity and large viscosity. However, other motions become optimal when friction is very anisotropic.

(e) Optima with travelling-wave contacts

One of the most interesting types of lifting motions in biological snakes is the sidewinding motion, which consists of a travelling wave of curvature synchronized with a travelling wave of lifting, with a phase difference between the two [12,14,26,37]. Interestingly, for some random initializations our optimization algorithm converges to these types of motions, across friction coefficient space. This only occurs for low viscous damping, $c_v = 0$ or 10^{-6} usually. So far in this paper, we have presented only the best local optimum we have found at a given parameter set, and these are never the sidewinding-type motions. We now show, in figure 7, examples of the sidewinding-type local optima. Six different examples are shown, labelled A, C, E, G, I and K in blue at left. Each case has a different pair of friction coefficient ratios, in parentheses below and to the right of the blue letter label. Below each of the six cases is the best local optimum found at the same friction coefficient ratios, labelled B, D, F, H, J and L at left. These are alternating lifting motions (B, D, F, H) or curling and sliding motions (J and L) that are similar to those already discussed. Next to each letter label is the value of $\tilde{P}$ for that motion, and the values for the sidewinding motions vary from slightly higher (K versus L) to about a factor of two higher (A versus B, C versus D etc.) than those of the best local optima.

To the right of each set of snapshots are the contact force maps and curvature maps in *s*-*t* space. For the sidewinding motions (A, C, E, G, I and K), both maps have diagonal bands with about the same slopes, showing unidirectional travelling waves of contact force and curvature that are approximately synchronized. The optimal motions (B, D, F, H, J and L) do not show unidirectional travelling waves except for motion J, and there the diagonal bands in the contact and curvature maps have different slopes, corresponding to waves moving at different speeds. Another distinctive feature of the sidewinding motions can be seen by examining the sets of snapshots in physical space, just to the right of the $\tilde{P}$ values. The black regions (where the body contacts the ground) of all the snapshots together trace out approximately continuous line segments on the ground. Between the black regions are almost-parallel arrays of green regions, lifted parts of the snake body that are being moved from one black region to the next. When the motion is repeated over multiple periods, the sets of black line segments form a series of parallel tracks on the ground, as for biological sidewinding snakes [12,14]. In A, C, E, G, I and K, the tracks are almost orthogonal to the direction of locomotion (up the page), whereas biological tracks are typically at an oblique angle [12,14,15]. In figure 7, the sidewinding motions separate

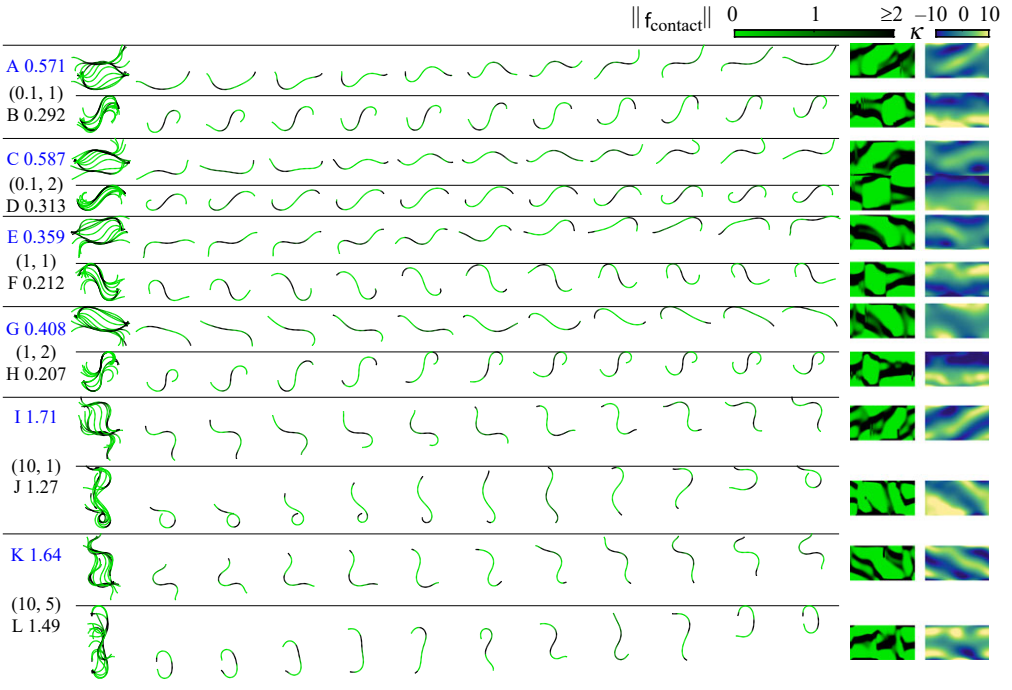

Figure 7. Six examples of local optima with travelling-wave contact force distributions (labelled in blue at left, A, C, E, G, I and K), paired with the global optimum (B, D, F, H, J and L, respectively) found at that set of friction coefficient ratios, listed in parentheses at the left. Also listed at left are the $\tilde{P}$ values. c_v is 10^{-6} in B, F, H, J, K and L, and 0 in the other six cases. To the right of each row of snapshots are plots showing the contact force distributions and curvature distributions versus s and t. (Online version in colour.)

into two groups: $\mu_n/\mu_f \leq 1$ (ACEG), where the lifted (green) body segments are at an oblique angle to the tracks and to the direction of locomotion, and $\mu_n/\mu_f = 10$, where the lifted segments are perpendicular to the tracks and parallel to the direction of locomotion. Here, the tracks are shorter and the body has a larger displacement per period.

It is not obvious why sidewinding optima occur commonly in the computations yet are never the global optima. Perhaps sidewinding motions require relatively more modes to achieve high efficiency than the global optima.

5. Simple theoretical motions

Our computed optima include many cases of alternating lifting (walking) or sidewinding, which involves contact regions rolling on and off the ground. We now demonstrate theoretical walking and rolling motions that locomote with essentially zero work done against friction. After describing the motions, we will discuss their differences with the computed optima and biological motions.

Figure 8*a* shows a sequence of four snapshots of a walking type of motion, similar in principle to the optima computed for $\mu_n/\mu_f \lesssim 1$. In each snapshot of *a*, the body is bent into a shape with the x-z plane as a plane of symmetry, in order to extend its base of support along the y direction, and make it stable to small perturbations that involve a rotation about the x-axis. The body has regions in contact with the ground—the 'feet'—and an elevated region that is used to shift the centre of mass from one foot to the other. Two dashed red lines are used as fixed guides to show how far the body locomotes. In snapshot 1, the left foot, highlighted in yellow, is in contact with the ground, and the body's centre of mass, a black dot, lies within its convex hull, so the body is

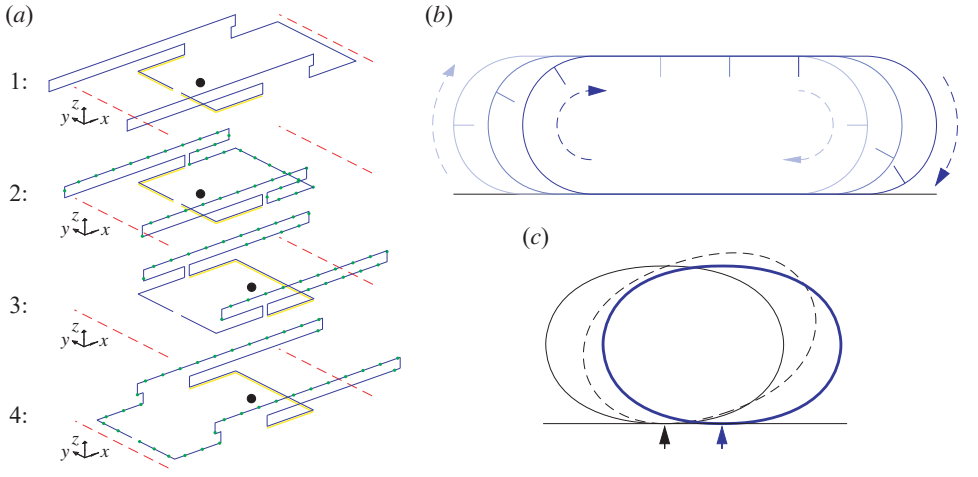

Figure 8. Examples of motions with very little work done against friction. (*a*) Walking motion, shown by four snapshots from top to bottom. (*b*) Tank-treading rolling motion, show by a sequence of three snapshots (light, medium and dark blue). (*c*) Falling-centre-of-mass rolling motion, shown by a sequence of three snapshots (solid black line, dashed black line and blue line). (Online version in colour.)

stable. From snapshot 1 to 2, the right foot is lifted slightly, then brought closer to the left foot, and meanwhile the elevated portion shifts rightward to keep the centre of mass fixed. Small green dots in snapshots 2–4 show the regions of the body that have moved from the previous snapshot. From snapshot 2 to 3, the body puts both feet on the ground, then shifts the elevated region so the centre of mass lies over the right foot, then lifts the left foot. From snapshot 3 to 4, the left foot is moved leftward, and simultaneously the elevated region shifts rightward to keep the centre of mass fixed. Now the body can return from snapshot 4 to snapshot 1 (but translated leftward), shifting the elevated region and the centre of mass leftward, and repeat the process. As with the computed optima, the body is a single continuous segment here. The computed optima, which involved lifting the middle portion and ends of an s-shaped body, are a smoother type of walking motion, which have smaller viscous dissipation than in *a* (or even a slightly regularized version that removes the sharp corners). With a finite number of modes, the computed optima probably cannot have a frictional dissipation that is precisely zero, as it is for the motion in *a*.

In (*b*,*c*), we examine two rolling types of motions, described as kinematic and dynamic rolling in [55]. A vast number of shape-changing rolling robots have been created using these types of motions, e.g. [56–59]. Panel *b* shows three successive snapshots of a body performing a tank-treading motion. Portions of the body at the left are rolled off the ground into the semicircle at the left end while portions at the right are rolled out of the semicircle at the right end, onto the ground. The dashed lines with arrows show the directions of rolling, and the hash marks show how fixed material points move between the three snapshots. In the tank-treading motion, the region in contact with the ground is stabilized by static friction, which is approximated by the sliding friction model with the small δ_s term we have used (equation (2.5)) as described in [27]. The tank-treading motion is called kinematic rolling, because the body is always at a stable equilibrium and inertia plays no role. In *c*, dynamic rolling, the body starts in a stable elliptical shape (solid black line), and then changes its shape to the dashed line. If friction is sufficiently large, the ground contact remains approximately fixed. The dashed line shape is gravitationally unstable, so the body rolls rightward to reduce the height of its centre of mass, resulting in the blue shape. With inertia, the body will perform a rocking oscillation about the blue shape, damped by internal or frictional dissipation in real situations. With further shape changes, the body may continue rightward with steady or variable velocity [57]. Like walking, these rolling motions do essentially no work against friction in the ideal case. They can also be made into stable three-dimensional

shapes by adding a mirror-symmetric portion, as for the motion in *a*. The rolling motions can be made to have relatively small changes in curvature, decreasing viscous dissipation.

Many of our computed optima resemble the walking or rolling strategies. The sidewinding motions involve a continuous rolling on and off the ground at certain contact regions, although the body does not assume a circular shape. A small number of other organisms have been found to roll passively (without shape change), including tumbleweed, a type of salamander, a spider, and a woodlouse [56]. Organisms that use active rolling to locomote include a type of shrimp and a caterpillar [56]. There may be various reasons why sidewinding is preferable to a more vertical rolling configuration for snakes. Sidewinding may allow for a better view of the snake's surroundings including prey and predators. In the sidewinding motion, the snake is closer to the ground, which is better for stealth when stalking prey or avoiding predators [15]. Body configurations that are closer to the ground are often more gravitationally stable. Another advantage is that less work is done is lifting portions of the body high off the ground. Although gravitational potential energy can be recovered in a periodic motion, there may be more losses during large conversions of potential to kinetic energy [60].

Another possible advantage of sidewinding over the rolling motions in figure 8 is that it preserves the snake's upright orientation. Snakes have an upright orientation in most studies of locomotion over approximately flat surfaces, and their belly scales seem to be adapted for contact with surfaces [15]. One notable and unusual case of an upside-down posture is a specialized antipredator behaviour of the hognose snake [61]. Upright postures are also preferred by many other organisms such as fish [62,63]. It is interesting that sidewinding optima occur in our computations even though our model omits biological considerations beyond mechanical efficiency, as well as details of snake physiology that may favour certain body postures such as sidewinding. Sidewinding may be the most efficient rolling type of motion that fits the constraint of small-to-moderate lifting. It is also interesting that sidewinding optima occur only with a very small viscous damping constant. However, the more-efficient alternating lifting or 'walking' optima are also optimal in many cases with large viscous dissipation, which include high-speed motions.

6. Conclusion

We have developed a model and computational method to find three-dimensional motions that optimize the mechanical efficiency of snake-like locomotion with small-to-moderate lifting off of the ground. The key physical parameters are the two ratios of friction coefficients ($\mu_n/\mu_f, \mu_b/\mu_f$), the lifting amplitude τ_{amp}, and the viscous damping parameter c_v. Our stochastic population-based optimization method finds motions that minimize the average input power $\tilde{P}$ while locomoting with very small net rotation per period. The same types of optima are found when we vary the numbers of modes in the Chebyshev–Fourier basis that describe the body shapes, so the results are robust with respect to changing the numbers of modes. Although we have mainly focused on the best computed optima, in most cases, the second-best computed optimum is very similar to the best, so the algorithm finds the same optima from different initializations. Presumably the best computed optimum has a sizeable basin of attraction and is not very difficult to locate.

For planar (and non-planar) locomotion, the transition from negligible to dominant viscous damping occurs near $c_v = 10^{-3}$. The optima in the planar case with zero viscous damping are very similar to those computed in [28] with a Newton-based optimization method. With non-zero viscous damping, the body shapes are often similar but smoother, and achieve larger displacements per period.

When the allowed lifting amplitude τ_{amp} is increased from zero, the optimal motions change dramatically. For $\mu_n/\mu_f \leq 3$, a motion with an s-shaped body and alternating lifting of the midbody and the ends is typical, though there are other types of motions at certain friction coefficients. For $\mu_n/\mu_f = 10$, the optimal motions are a combination of lifting with curling and sliding at zero c_v. At large c_v, the motions are a combination of lifting with large amplitude

bending and sliding, and somewhat resemble lateral undulation but combined with lifting that alternates between the ends. Other types of optima are seen with moderate and large damping and small or large μ_n/μ_f and $\mu_b/\mu_f > 1$, including repeated curling and uncurling with the contacts almost fixed and oriented in the direction of minimal friction.

At all friction coefficient values, optima that resemble sidewinding are seen, but these underperform the optima with an s-shaped body and alternating lifting. The sidewinding optima have a unidirectional wave of contact synchronized with a wave of curvature, tracing out a contiguous contact region (or 'track') on the ground. At different friction coefficient ratios, the body orientation in the lifted region varies. Finally, we discussed the resemblance of the computed optima to theoretical walking and rolling motions with essentially zero frictional dissipation. Future work may consider additional physical effects such as a non-planar substrate, as well as the possible benefits of passive flexibility on efficient non-planar locomotion [46].

Data accessibility. Source codes are provided in electronic supplementary material [64].

Conflict of interest declaration. I declare I have no competing interests.

Funding. This research was supported by the US NSF Mathematical Biology award DMS-1811889.

References

1. Shine R, Cogger HG, Reed RR, Shetty S, Bonnet X. 2003 Aquatic and terrestrial locomotor speeds of amphibious sea-snakes (Serpentes, Laticaudidae). *J. Zool.* **259**, 261–268. (doi:10.1017/S0952836902003242)
2. Socha JJ. 2002 Kinematics: gliding flight in the paradise tree snake. *Nature* **418**, 603–604. (doi:10.1038/418603a)
3. Hirose S. 1993 *Biologically inspired robots: snake-like locomotors and manipulators*. Oxford, UK: Oxford University Press.
4. Transeth AA, Pettersen KY, Liljebäck P. 2009 A survey on snake robot modeling and locomotion. *Robotica* **27**, 999–1015. (doi:10.1017/S0263574709005414)
5. Hopkins JK, Spranklin BW, Gupta SK. 2009 A survey of snake-inspired robot designs. *Bioinspir. Biomim.* **4**, 021001. (doi:10.1088/1748-3182/4/2/021001)
6. Liljebäck P, Pettersen KY, Stavdahl Ø, Gravdahl JT. 2012 *Snake robots: modelling, mechatronics, and control*. London, UK: Springer.
7. Hatton RL, Choset H. 2010 Generating gaits for snake robots: annealed chain fitting and keyframe wave extraction. *Autonom. Rob.* **28**, 271–281. (doi:10.1007/s10514-009-9175-2)
8. Astley HC *et al.* 2015 Modulation of orthogonal body waves enables high maneuverability in sidewinding locomotion. *Proc. Natl Acad. Sci. USA* **112**, 6200–6205. (doi:10.1073/pnas.1418965112)
9. Fu Q, Li C. 2020 Robotic modelling of snake traversing large, smooth obstacles reveals stability benefits of body compliance. *R. Soc. Open Sci.* **7**, 191192. (doi:10.1098/rsos.191192)
10. Astley HC, Rieser JM, Kaba A, Paez VM, Tomkinson I, Mendelson JR, Goldman DI. 2020 Side-impact collision: mechanics of obstacle negotiation in sidewinding snakes. *Bioinsp. Biomim.* **15**, 065005. (doi:10.1088/1748-3190/abb415)
11. Fu Q, Astley H, Li C. 2022 Snakes combine vertical and lateral bending to traverse uneven terrain. *Bioinsp. Biomim.* **17**, 036009. (doi:10.1088/1748-3190/ac59c5)
12. Gray J. 1946 The mechanism of locomotion in snakes. *J. Exp. Biol.* **23**, 101–120. (doi:10.1242/jeb.23.2.101)
13. Gans C. 1970 How snakes move. *Sci. Am.* **222**, 82–99. (doi:10.1038/scientificamerican0670-82)
14. Jayne BC. 1986 Kinematics of terrestrial snake locomotion. *Copeia* **1986**, 915–927. (doi:10.2307/1445288)
15. Lillywhite HB. 2014 *How snakes work: structure, function and behavior of the world's snakes*. Oxford, UK: Oxford University Press.
16. Gans C. 1984 Slide-pushing: a transitional locomotor method of elongate squamates. In *Symp. Zool. Soc. London*, vol. 52, London, UK, 26–27 May 1983, pp. 12–26. London, UK: Academic Press.
17. Jayne BC. 2020 What defines different modes of snake locomotion?. *Int. Comp. Biol.* **60**, 156–170. (doi:10.1093/icb/icaa017)
18. Ma S. 2001 Analysis of creeping locomotion of a snake-like robot. *Adv. Robot.* **15**, 205–224. (doi:10.1163/15685530152116236)

19. Sato M, Fukaya M, Iwasaki T. 2002 Serpentine locomotion with robotic snakes. *IEEE Cont. Sys. Mag.* **22**, 64–81. (doi:10.1109/37.980248)
20. Chernousko FL. 2005 Modelling of snake-like locomotion. *Appl. Math. Comput.* **164**, 415–434. (doi:10.1016/j.amc.2004.06.057)
21. Guo ZV, Mahadevan L. 2008 Limbless undulatory propulsion on land. *Proc. Natl Acad. Sci.* **105**, 3179. (doi:10.1073/pnas.0705442105)
22. Hu DL, Nirody J, Scott T, Shelley MJ. 2009 The mechanics of slithering locomotion. *Proc. Natl Acad. Sci. USA* **106**, 10081. (doi:10.1073/pnas.0812533106)
23. Hu DL, Shelley M. 2012 Slithering locomotion. In *Natural Locomotion in Fluids and on Surfaces*, pp. 117–135. New York, NY: Springer.
24. Aguilar J *et al.* 2016 A review on locomotion robophysics: the study of movement at the intersection of robotics, soft matter and dynamical systems. *Rep. Prog. Phys.* **79**, 110001. (doi:10.1088/0034-4885/79/11/110001)
25. Yona T, Or Y. 2019 The wheeled three-link snake model: singularities in nonholonomic constraints and stick–slip hybrid dynamics induced by Coulomb friction. *Nonlinear Dyn.* **95**, 2307–2324. (doi:10.1007/s11071-018-4693-0)
26. Rieser JM, Li TD, Tingle JL, Goldman DI, Mendelson III JR. 2021 Functional consequences of convergently evolved microscopic skin features on snake locomotion. *Proc. Natl Acad. Sci. USA* **118**, e2018264118. (doi:10.1073/pnas.2018264118)
27. Alben S. 2019 Efficient sliding locomotion with isotropic friction. *Phys. Rev. E* **99**, 062402. (doi:10.1103/PhysRevE.99.062402)
28. Alben S. 2013 Optimizing snake locomotion in the plane. *Proc. R. Soc. A* **469**, 1–28. (doi:10.1098/rspa.2013.0236)
29. Wang X, Osborne MT, Alben S. 2014 Optimizing snake locomotion on an inclined plane. *Phys. Rev. E* **89**, 012717. (doi:10.1103/PhysRevE.89.012717)
30. Jing F, Alben S. 2013 Optimization of two- and three-link snake-like locomotion. *Phys. Rev. E* **87**, 022711. (doi:10.1103/PhysRevE.87.022711)
31. Alben S, Puritz C. 2020 Intermittent sliding locomotion of a two-link body. *Phys. Rev. E* **101**, 052613. (doi:10.1103/PhysRevE.101.052613)
32. Alben S. 2021 Efficient sliding locomotion of three-link bodies. *Phys. Rev. E* **103**, 042414. (doi:10.1103/PhysRevE.103.042414)
33. Childress S. 1981 *Mechanics of swimming and flying*. Cambridge, UK: Cambridge University Press.
34. Sparenberg J. 1994 *Hydrodynamic propulsion and its optimization: (Analytic Theory)*, vol. 27. Dordrecht, The Netherlands: Kluwer Academic Pub.
35. Alben S. 2008 Optimal flexibility of a flapping appendage in an inviscid fluid. *J. Fluid Mech.* **614**, 355–380. (doi:10.1017/S0022112008003297)
36. Alben S. 2009 On the swimming of a flexible body in a vortex street. *J. Fluid Mech.* **635**, 27–45. (doi:10.1017/S0022112009990619)
37. Marvi H *et al.* 2014 Sidewinding with minimal slip: Snake and robot ascent of sandy slopes. *Science* **346**, 224–229. (doi:10.1126/science.1255718)
38. Zhang X, Naughton N, Parthasarathy T, Gazzola M. 2021 Friction modulation in limbless, three-dimensional gaits and heterogeneous terrains. *Nat. Commun.* **12**, 1–8. (doi:10.1038/s41467-021-26276-x)
39. Chong B, Wang T, Lin B, Li S, Blekherman G, Choset H, Goldman DI. 2022 Moving sidewinding forward: optimizing contact patterns for limbless robots via geometric mechanics. In *Robotics: science and systems*.
40. Alexander RM. 1996 *Optima for animals*. Princeton, NJ: Princeton University Press.
41. Langerhans RB, Reznick DN. 2010 Ecology and evolution of swimming performance in fishes: predicting evolution with biomechanics. *Fish Locomotion* **200**, 248. (doi:10.1201/b10190-8)
42. Tingle JL. 2020 Facultatively sidewinding snakes and the origins of locomotor specialization. *Integr. Comp. Biol.* **60**, 202–214. (doi:10.1093/icb/icaa011)
43. Guggenheimer H. 2012 *Differential geometry*. New York, NY: Dover Publications.
44. Marvi H, Hu DL. 2012 Friction enhancement in concertina locomotion of snakes. *J. R. Soc. Interface* **9**, 3067–3080. (doi:10.1098/rsif.2012.0132)
45. Schultz WW, Webb PW. 2002 Power requirements of swimming: do new methods resolve old questions?. *Integr. Comp. Biol.* **42**, 1018–1025. (doi:10.1093/icb/42.5.1018)
46. Wang X, Alben S. 2018 Dynamics and locomotion of flexible foils in a frictional environment. *Proc. R. Soc. A* **474**, 20170503. (doi:10.1098/rspa.2017.0503)

47. Linn J, Lang H, Tuganov A. 2013 Geometrically exact Cosserat rods with Kelvin–Voigt type viscous damping. *Mech. Sci.* **4**, 79–96. (doi:10.5194/ms-4-79-2013)
48. Tam D, Hosoi AE. 2007 Optimal stroke patterns for Purcell's three-link swimmer. *Phys. Rev. Lett.* **98**, 68105. (doi:10.1103/PhysRevLett.98.068105)
49. Hatton RL, Ding Y, Choset H, Goldman DI. 2013 Geometric visualization of self-propulsion in a complex medium. *Phys. Rev. Lett.* **110**, 078101. (doi:10.1103/PhysRevLett.110.078101)
50. Gutman E, Or Y. 2015 Symmetries and gaits for Purcell's three-link microswimmer model. *IEEE Trans. Rob.* **32**, 53–69. (doi:10.1109/TRO.2015.2500442)
51. Wu W, Yu S, Schreiber P, Dollmann A, Lutz C, Gomard G, Greiner C, Hölscher H. 2020 Variation of the frictional anisotropy on ventral scales of snakes caused by nanoscale steps. *Bioinspir. Biomim.* **15**, 056014. (doi:10.1088/1748-3190/ab9e51)
52. Schneck DJ. 1992 *Mechanics of muscle*. New York, NY: New York University Press.
53. Cheng JY, Pedley T, Altringham J. 1998 A continuous dynamic beam model for swimming fish. *Phil. Trans. R. Soc. Lond. B* **353**, 981–997. (doi:10.1098/rstb.1998.0262)
54. Al Mayah A. 2018 *Biomechanics of soft tissues: principles and applications*. Boca Raton, FL: CRC Press.
55. Sastra J, Chitta S, Yim M. 2009 Dynamic rolling for a modular loop robot. *Int. J. Rob. Res.* **28**, 758–773. (doi:10.1177/0278364908099463)
56. Armour RH, Vincent JF. 2006 Rolling in nature and robotics: a review. *J. Bionic Eng.* **3**, 195–208. (doi:10.1016/S1672-6529(07)60003-1)
57. Puopolo MG, Jacob JD. 2016 Velocity control of a cylindrical rolling robot by shape changing. *Adv. Rob.* **30**, 1484–1494. (doi:10.1080/01691864.2016.1246262)
58. Li WB, Zhang WM, Gao QH, Guo Q, Wu S, Zou HX, Peng ZK, Meng G. 2021 Electrically activated soft robots: speed up by rolling. *Soft Rob.* **8**, 611–624. (doi:10.1089/soro.2020.0012)
59. Shah DS, Powers JP, Tilton LG, Kriegman S, Bongard J, Kramer-Bottiglio R. 2021 A soft robot that adapts to environments through shape change. *Nat. Mach. Intell.* **3**, 51–59. (doi:10.1038/s42256-020-00263-1)
60. Biewener A, Patek S. 2018 *Animal locomotion*. Oxford, UK: Oxford University Press.
61. Hemken BS. 1974 *Defensive Behavior of the Hognose Snake (Heterodon platyrhinos)*. PhD thesis. Eastern Illinois University.
62. Eidietis L, Forrester T, Webb P. 2002 Relative abilities to correct rolling disturbances of three morphologically different fish. *Can. J. Zool.* **80**, 2156–2163. (doi:10.1139/z02-203)
63. Tasoff H. 2017 Why don't fish swim upside down? *Hakai Magazine*.
64. Alben S. 2022 Efficient bending and lifting patterns in snake locomotion. Figshare. (doi:10.6084/m9.figshare.c.6214757)