

Research



Cite this article: Giardina F, Mahadevan L. 2021 Models of benthic bipedalism. *J. R. Soc. Interface* **18**: 20200701.
<https://doi.org/10.1098/rsif.2020.0701>

Received: 29 August 2020

Accepted: 8 December 2020

Subject Category:

Life Sciences–Physics interface

Subject Areas:

biomechanics, biomimetics

Keywords:

bipedalism, robotics, benthic walking

Author for correspondence:

L. Mahadevan

e-mail: lmahadev@g.harvard.edu

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5250907>.

Models of benthic bipedalism

F. Giardina¹ and L. Mahadevan^{1,2,3}

¹John A. Paulson School of Engineering and Applied Sciences, ²Department of Physics, and ³Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA

FG, 0000-0002-2660-5935; LM, 0000-0002-5114-0519

Walking is a common bipedal and quadrupedal gait and is often associated with terrestrial and aquatic organisms. Inspired by recent evidence of the neural underpinnings of primitive aquatic walking in the little skate *Leucoraja erinacea*, we introduce a theoretical model of aquatic walking that reveals robust and efficient gaits with modest requirements for body morphology and control. The model predicts undulatory behaviour of the system body with a regular foot placement pattern, which is also observed in the animal, and additionally predicts the existence of gait bistability between two states, one with a large energetic cost for locomotion and another associated with almost no energetic cost. We show that these can be discovered using a simple reinforcement learning scheme. To test these theoretical frameworks, we built a bipedal robot and show that its behaviours are similar to those of our minimal model: its gait is also periodic and exhibits bistability, with a low efficiency mode separated from a high efficiency mode by a ‘jump’ transition. Overall, our study highlights the physical constraints on the evolution of walking and provides a guide for the design of efficient biomimetic robots.

1. Introduction

The transition of vertebrates from water to land is thought to have occurred around 400 million years ago and required fundamental changes in morphology and behaviour. Aquatic vertebrates living in near-neutral buoyancy had to adapt to the effects of gravity on land, which required a change in locomotion strategy. Fins provided a basic form of legs that helped early land-dwelling vertebrates to support their body weight and switch from undulatory swimming (by lateral bending of the body) to ambulatory walking (by swinging their limbs). A common view of the transition from swimming to walking is that it occurred gradually [1], consistent with observations in contemporary tetrapods that show vestiges of both swimming and legged locomotion, e.g. the salamander [2]. Walking requires the coordinated motion of limbs. While the development of limbs and legs can be traced to the fossil record, the origins of neural circuits giving rise to the control required for ambulatory locomotion are unclear. However, recent work [3] suggests that the neural circuits required for limb control can be found in aquatic vertebrates who are distant relatives to the first tetrapods, indicating that the neuromuscular basis for legged locomotion was present in all vertebrates with paired fins. These observations raise the question: Was a walking gait actively used in the earliest finned vertebrates, or did it emerge just prior to terrestrialization? Given the plethora of extant benthic (living near the bottom of aquatic environments) fish and species that can walk in benthic and littoral regions [4–7], it is conceivable that their ancestors with similar morphologies might have used these ancient neural circuits to walk. A particularly striking extant example in this regard is the little skate *Leucoraja erinacea* (figure 1a,b) that is incapable of much undulation due to its stiff spinal column [3,8] and uses a benthic gait consisting of left-right alternating walking. These observations are strongly suggestive that walking in aquatic environments can emerge without a prior undulatory gait and that the requirement for the evolution of walking is the capability to independently control each fin along with a control strategy that synchronizes the walking motion. While there is strong evidence that

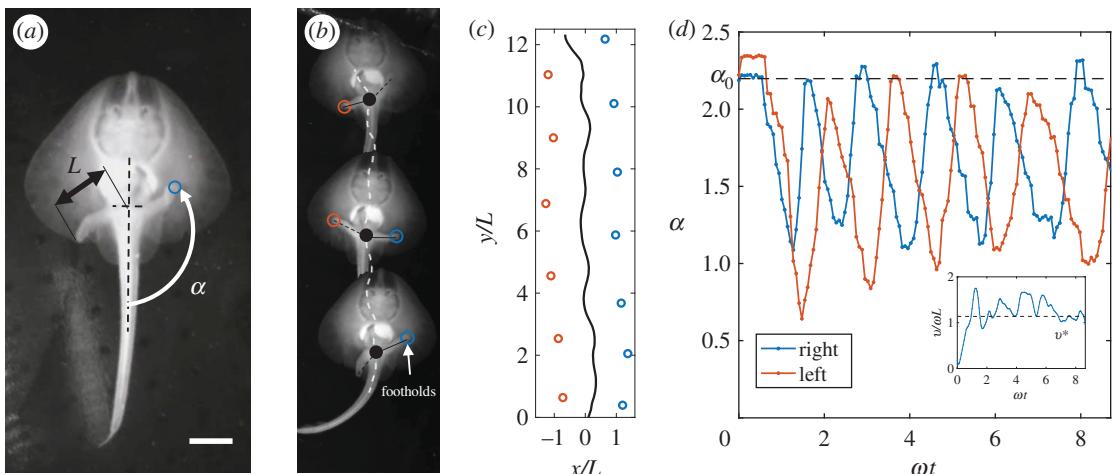


Figure 1. Little skate locomotion behaviour movies in [3] (with permission from the authors). (a) Ventral view of little skate indicating leg length L , footsteps and leg angle α measured relative to the centreline of the body. Scale bar, 1 cm. (b) Sequence of walking gait indicating trajectory of the pelvic girdle (dashed white line), active legs (solid black lines), passive legs (dashed black lines) and footsteps. (c) Trajectory of the pelvic girdle (black line) as a function of dimensionless position with footsteps (circles). (d) Left and right leg angles α as a function of dimensionless time and mean foot placement angle α_0 . The inset shows the dimensionless speed of the pelvic girdle as a function of dimensionless time with v^* (dashed line) the approximate lower speed bound during steady-state locomotion.

independent leg control was present in early vertebrates with paired fins [3], the form of the gait and the concomitant control strategy for stable and efficient legged locomotion remain open questions.

Inspired by published video data [3] on the dynamics of walking in the little skate, here we study the questions raised by the stability, energy efficiency and control complexity of early benthic locomotion in terms of a minimal mathematical model, a computational realization and a physical instantiation. We compare the most efficient gaits predicted by the model to the kinematics of the little skate and show that both are characterized by a left-right alternating leg pattern with an undulating centre of mass and a regular foot placement pattern. Closed-form expressions for the dynamics of the model show that the most energy efficient gait is associated with no energetic cost of locomotion and merely requires a simple open-loop control strategy. In addition, the model also predicts the coexistence of a second gait with much lower efficiency. To complement this explicit dynamic model, we use a reinforcement learning (RL) strategy and show that a little skate-like gait is the preferred solution in this framework, suggesting that an evolutionary process can converge to this in nature. Finally, to test our results with hardware, we created a simple bipedal robot and show that it also exhibits bistable behaviour for certain control parameters, qualitatively consistent with our theory.

2. Mathematical model of bipedalism

Published videos of the little skate [3] allowed us to extract foot placements, trajectories and leg angles as shown in figure 1c,d (see also electronic supplementary material, SI, and video). In steady locomotion, we observed that as the pelvic girdle position of the skate undulates during locomotion, only one foot is in contact with the ground at any time, and there is little slip between the leg and the ground during contact. This leads to a periodic foot placement as shown in figure 1c, accompanied by periodic dynamics of the leg angle α (measured with respect to the centreline of the body), as shown in figure 1d.

Based on these observations, a minimal model of locomotion suggests modelling the aquatic body as a neutrally buoyant mass m with rotational moment of inertia I , which can move and rotate freely in the plane if no legs are placed on the ground, as shown in figure 2. Since the legs of length L are very light relative to the massive body, we neglect their mass. We further assume them to be directly attached to the centre of mass, with the ability to switch their contact state with the ground at a frequency ω . The foot-ground contact was modelled as a perfectly inelastic impact, which dissipates any velocity component that violates the constraint at foot placement, consistent with a range of previously used simple models of locomotion [9–11]. In the absence of slip, upon leg contact with the ground, the velocity of the body v is perpendicular to the leg vector r_c , which points from body to contact point c . It may be useful here to contrast our model with other commonly adopted models for walking and running on land [12,13] and punting in water [14] that rely on spring-mass systems, known collectively as SLIP models. These spring models are particularly useful in their ability to model energy recovery in such instances as running dynamics. However, our kinematic analysis of the little skate data shown in figure 1c indicates no evidence of a spring-like leg and instead suggests that the body pivots rigidly around a fixed contact point. Therefore, we do not account for leg compliance, but recognize that spring-like legs might facilitate other interesting gait dynamics in benthic environments. During the impulse-free phases, a torque T accelerates the body along a circular path around the active leg with length L and the current contact point c . For simplicity, we assumed the torque to be constant during the duration of a step i , i.e. $T_i(t) = T$.

The resulting dynamics of this model can be then described in terms of generalized coordinates $q = [x, y, \theta]^T$, with x, y defining the position in the plane and θ the body angle with respect to a fixed direction in an inertial frame of reference. In addition, a constraint $r_c^T v = 0$ enforces pivoting of the centre of mass around the current ground contact point. The minimal picture outlined earlier is consistent with a model for the formation and breaking of bilateral contacts with the ground studied in non-smooth mechanics [15].

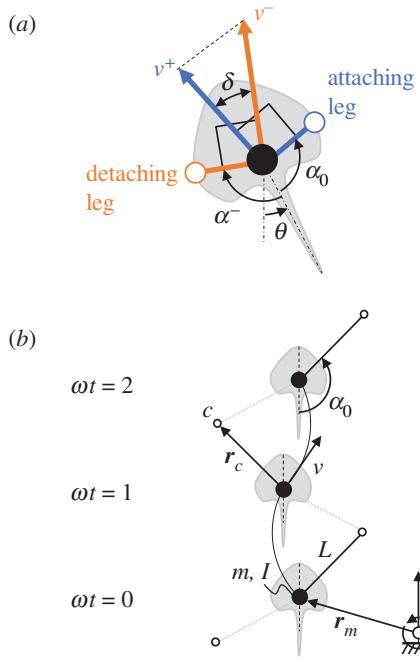


Figure 2. Model sketch. (a) The leg transition is modelled with an inelastic impact where the post-impact speed v^+ is obtained by a projection of the pre-impact speed v^- to the line perpendicular to the attaching leg direction. (b) Model sketch at three subsequent dimensionless times. Point mass m with moment of inertia I is constrained in its motion by a connected massless leg with length L to rotate around current active contact point c . The system cannot slip, and velocity losses can only occur due to inelastic impact at leg transition. A constant torque T applied to the leg can accelerate the system. The leg placement angle is given with $\pm\alpha_0$ at transition.

Ignoring any dissipative effects during a step (see the following discussion for a consideration of these effects), we may write the rotational equation of motion for the body as $\ddot{\theta}(t) = T/I$. Integrating this equation of motion for an alternating leg sequence with constant frequency ω and torque T , the body rotates during a step i (index i indicating the leg transition time, i.e. $\theta_i = \theta(i/\omega)$) by $\Delta\theta_i = \pm T/2I\omega^2 + \theta_{i-1}/\omega$ where $\dot{\theta}_i = \dot{\theta}_{i-1} \pm T/I\omega$. We note that the transition from active ground contact in one leg to the other is symmetric. Furthermore, with initial conditions $\dot{\theta}_0 = \mp T/2I\omega$, $\theta_0 = 0$ (positive sign for right leg, negative for left) the body angle at transition is always zero, i.e. $\theta_i = 0 \forall i$, and we can ignore the evolution of the body rotation in the following analysis. This is also indicated in the stepping sequence in figure 2b where all body orientations at leg transitions are aligned.

The dynamics of the inertial body coupled with the bilateral constraint leads to body translation along an undulatory path. Converting leg angular motions to body linear velocities following the geometry shown in figure 2a, we see that pre- and post-impact velocities are perpendicular to the leg direction (defined by the angles α^- for the detaching leg and α_0 for the attaching leg relative to the overall direction of the body) due to the pivoting motion. Denoting the magnitude of the centre of mass velocity at step i before impact by v_i^- and after impact by v_i^+ , for rigid legs, we note that at the leg transition the projection of the velocity of the detaching leg will not be necessarily in the direction perpendicular to the attaching leg. This incompatibility will lead to an impulse associated with impact, which we assume to be perfectly inelastic and allows us to define an inelastic impact law at step i purely as

a function of the leg angles at collision as follows:

$$v_i^+ = |\cos \delta_i| v_i^-, \quad (2.1)$$

where $\delta_i = \alpha_0 - \alpha_i^- - \pi$ is the difference in leg angles at the transition, as shown in figure 2a. During the leg contact phase, the gain in velocity of the body due to the constant torque T over the period $1/\omega$ is

$$\Delta v = v_i^- - v_{i-1}^+ = \frac{T}{\omega m L}, \quad (2.2)$$

which follows from the acceleration of a point mass m around a circular path with radius L . This condition then defines the mapping of the previous step's post-impact velocity, v_{i-1}^+ , to the post-impact velocity of the current step via (2.1)

$$v_i^+ = |\cos \delta_i|(v_{i-1}^+ + \Delta v). \quad (2.3)$$

Equation (2.3) has an implicit dependence on α_i^- through δ_i . In turn, this difference itself depends on v_{i-1}^+ via the relation

$$\alpha_i^- = -\alpha_0 + \left(\frac{v_{i-1}^+}{\omega L} + \frac{T}{2\omega^2 L^2 m} \right) \quad (2.4)$$

that follows from computing the leg angle by integrating (2.2). Then, the evolution of the centre of mass velocity v_i^+ over one step is given in terms of its initial velocity and a prescribed torque and follows the equations (2.2)–(2.4), and the little skate's trajectory and velocity are fully defined given the leg touchdown angle α_0 , the leg torque T , the body mass m and the leg switching frequency ω .

Before analysing the dynamical system defined earlier, we note that the experimentally observed kinematic data in figure 1 and electronic supplementary material suggest that, after a short transient phase, the leg torque T , frequency ω and leg placement angle α_0 are constant over consecutive steps. This suggests a minimal open-loop *bipedal control strategy* for locomotion: alternate leg activation and keep the applied torque T , the driving frequency ω and the gait angle α_0 constant over all steps. Since all control parameters are constant, the speeds at which loss due to collision and gain due to torque T balance and result in no variation in speed over subsequent steps, i.e. the fixed points of the system v^* , are given by the implicit equation (2.3). By using non-dimensional fixed-point velocities $\bar{v}^* = v^*/\omega L$, we obtain

$$\bar{v}^*(1 - |\cos \delta|) = \gamma |\cos \delta|, \quad (2.5)$$

where $\gamma = T/\omega^2 L^2 m$ is the non-dimensional torque.

To reveal the possible behaviours and gaits of the model, we searched for solutions of the discrete map (2.5) and determine their stability. Finding a solution (i.e. a fixed point \bar{v}^*) in the model reveals the periodic gaits of the system. For a constant foot placement angle α_0 , the fixed-point velocity \bar{v}^* is solely defined by the non-dimensional torque γ . The stability of the fixed points of the discrete map (2.3) follows from a simple linear stability analysis, i.e. searching for v^* , such that

$$\left| \frac{dv_i^+}{dv_{i-1}^+} \right|_{v_{i-1}^+ = v^*} < 1. \quad (2.6)$$

Figure 3 shows the solutions of (2.5) for a fixed foot placement angle α_0 , plotted in a bifurcation diagram. Stability analysis of the solutions revealed that there are three stable regions of interest in the diagram. For $\gamma \leq 0.2$, two solution branches coexist, one at low speed and the other one at high speed connected by an unstable region. A system on

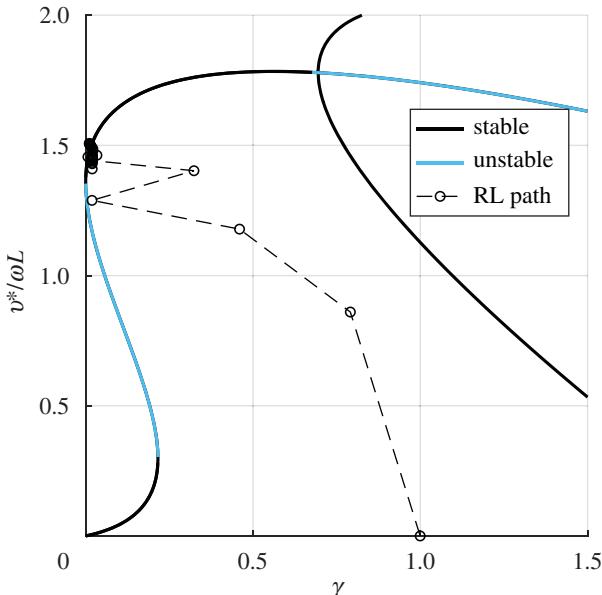


Figure 3. Dynamics of bipedal locomotion strategy. Bifurcation diagram with non-dimensional torque $\gamma = T/\omega^2 L^2 m$ for a fixed foot placement angle $\alpha_0 = 2.25$. Black lines are stable fixed-point velocities $\bar{v}^* = v^*/\omega L$ and blue lines unstable ones. The bifurcation around $\gamma \approx 0.75$ is a period-doubling bifurcation, and the solution jumps between the lower and upper black branch. Black circles show consecutive steps of the optimized reinforcement learning policy.

the edge of the lower branch will experience a sudden jump in its attracting fixed point as the non-dimensional torque γ is increased, consistent with the existence of a saddle-node bifurcation [16]. For $\gamma \in [0.2, 0.75]$, only one stable fixed point exists, which is the continuation of the upper branch. Finally, around $\gamma \approx 0.75$, a period-doubling bifurcation occurs, and the system jumps between two solution branches for $\gamma > 0.75$. This gait is asymmetric with a sideway component relative to the body orientation.

Gaits with small γ are biologically most plausible as they reduce energetic cost. In fact, there exists a point corresponding to $\gamma = 0$ on the upper branch in the bifurcation diagram for which the post-impact velocity is non-zero, indicating a gait with zero energetic cost. This was confirmed by considering the kinetic energy over time: energy fluctuations vanished when $\gamma = 0$, which required the leg vectors r_c of the left and right leg at transition to be parallel, i.e. $\delta = 0$. To characterize the speed of this point, which corresponds to the energy-optimal gait, we note that for small γ , the right-hand side of (2.5) is negligible, and we have the solutions $\bar{v}^* = 0$ and $|\cos \delta| = 1$. The first fixed-point speed provides the trivial solution on the lower branch. For the second solution, recall that $\delta = 2\alpha_0 - \bar{v}^* - \gamma/2 - \pi$, which implies $\bar{v}^* = 2\alpha_0 - n\pi$, $n \in \mathbb{Z}$. For $n = 1$, the solution corresponds to the energy-optimal gait, with $\delta = 0$, i.e. the legs are parallel at the transition eliminating any dissipation due to impact. Varying the integer n leads to an infinite number of energy-optimal gaits, where the transition from one leg to the next occurs after $|n|/2$ rotations around the same ground contact point, effectively resulting in a pirouette-like rotation before the leg transition. In the context of our investigation of directed bipedal locomotion, we do not analyse this further.

In real physical systems that must operate with $\gamma > 0$ to ensure stability, the upper branch for $\gamma < 0.2$ is not reachable

from rest with the bipedal control strategy, and the system will always converge to the lower, slower and less efficient branch. To switch to the upper branch, one has to literally ‘leap’ onto it using an impulsive start. Such motions are observed in the little skate as it takes off from rest by punting forward with a powerful stroke using both legs at the same time, followed by immediate switching to the alternating gait (see figure 1d and electronic supplementary material).

We now turn to understand the role of the leg placement angle α_0 on the behaviour of the bifurcation diagram. We find that the two saddle-node bifurcations move, but the bifurcation diagram remains the same qualitatively. At the lowest sensible value of $\alpha = \pi/2$, when the legs are perpendicular to the body, the two bifurcation points merge at the origin of the coordinate system, from which the stable upper branch extends that transitions to the period-doubling bifurcation at $\gamma \approx 1$. This configuration is the slowest, as the energy efficient gait is effective at $v^* = 0$. At the highest reasonable value of $\alpha = \pi$, when the leg is parallel and in front of the body, the fastest energy efficient gait can be found at $v^* = \pi$. This configuration’s upper branch is rather short, as the period-doubling bifurcation occurs at $\gamma \approx 0.5$ and the upper branch’s basin of attraction does not allow for a convergence from zero initial velocity, unlike the case shown in figure 3. Observations of the little skate show a range of $\alpha_0 \in [2\pi/3, 3\pi/4]$, which strikes a balance between speed and robustness of the upper branch, and our analysis therefore focuses on this range of values.

Motion underwater is resisted by the effect of fluid drag. To understand the role of this on the fixed points \bar{v}^* , we note that the Reynolds number for the motion of the little skate is $O(10^3)$, where a crude approximation allows us to write the drag force as $F_d = 1/2\rho v^2 C_d A$, with ρ the fluid density, v the body velocity, C_d the drag coefficient and A the reference area. The velocity of the body in this case obeys the differential equation

$$\frac{dv}{dt} = \frac{T}{mL} - \frac{1}{2}\rho v^2 C_d A, \quad (2.7)$$

which has the solution

$$v(t) = \sqrt{\frac{2T}{L\rho C_d A}} \tanh\left(\sqrt{\frac{\rho C_d A T}{2m^2 L}}(c_1 mL + t)\right) \quad (2.8)$$

with c_1 a constant to be defined by initial conditions. This then changes the form of the map (2.3) for Δv , and the velocity increases over the contact phase of one leg. The fixed points now depend on the drag coefficient C_d . As expected, the original drag-free solution is recovered for $C_d \rightarrow 0$, and figure 4 shows the bifurcation diagram for the case with fluid drag for various values of C_d and choices of the reference area estimated from specimen dimensions [3]. As the drag coefficient increases, the regime of bistability decreases. Estimates of the drag coefficient for a benthic ray (*Raja clavata*) suggest a drag coefficient $C_d \sim 0.1$ (see [17]), while that for a sphere of a similar size would suggest $C_d \approx 0.5$. Even for this conservative case, we found that bistability persists in a regime that is qualitatively similar to the no-drag case although the non-dimensional torque γ required for a similar speed increases significantly. It is only when $C_d \approx 1.08$ that we see that the bistable region disappears completely and is replaced by a monostable region before transitioning to a period-doubling bifurcation.

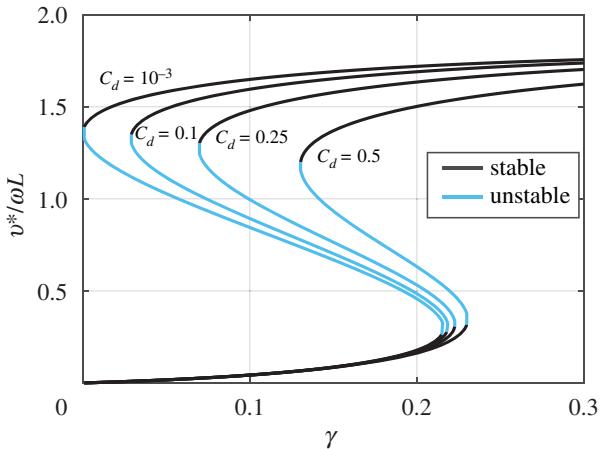


Figure 4. Bifurcation diagram with fluid drag for various drag coefficients. Bifurcation diagram as a function of nondimensional torque $\gamma = T/\omega^2 L^2 m$. Black lines are stable fixed-point velocities $v^*/\omega L$ and blue lines are unstable ones. Equations for the fixed points are obtained by using (2.8) in (2.3). As $C_d \rightarrow 0$, the bifurcation diagram approaches the result shown in figure 3.

3. Reinforcement learning of bipedalism

Our analysis so far shows that aquatic bipedalism has few requirements in terms of morphology (rudimentary legs) and control (constant torque, touchdown angle and frequency), but we need to impose an alternating leg sequence with fixed foot placement angle and torque for the analysis. This then naturally raises the question: Can the neural control for this bipedal gait be discovered given an aquatic organism with rudimentary appendages? Is it optimal when the constraints associated with our simple choice of constant torque etc. are relaxed?

The search for a favourable gait relates to the field of gait selection and optimization [18–21], while the learning of motor control relates to the notion of RL [22]. With the objective of maximizing the travelled distance and minimizing the required energy (or equivalently, minimizing energy consumption for a travelled distance), we trained an RL agent [23,24] on the model to obtain a given locomotion speed v_T . The framework has four state parameters (planar position and velocity of the point mass) and four control parameters (three continuous ones for T , ω , α_0 and a binary one for the leg case l)—for details, see Methods.

We observed, in most instances of the learning routine, e.g. figure 5, a one-legged locomotion strategy emerges after only four episodes, two-legged locomotion emerges around episode 200 and periodic locomotion with alternating leg cycles emerges via RL around episode 600. The periodic gait prevails as the most efficient one, and other explored strategies have a worse cumulative reward. We ran ~ 50 instances of learning for 5000 episodes with changing learning parameters and weights for the reward function and found that the left-right alternating gait emerged in 70% of instances and generated the highest reward. The best solution matches the little skate's observed walking gait in figure 1, and we see an undulatory motion of the centre of mass and a left-right alternating leg sequence (see figure 2c and electronic supplementary material, video).

For comparison with the analytically obtained bifurcation diagram for gaits, we plot the evolution of the best learned RL policy in the parameter space as the dashed lines in

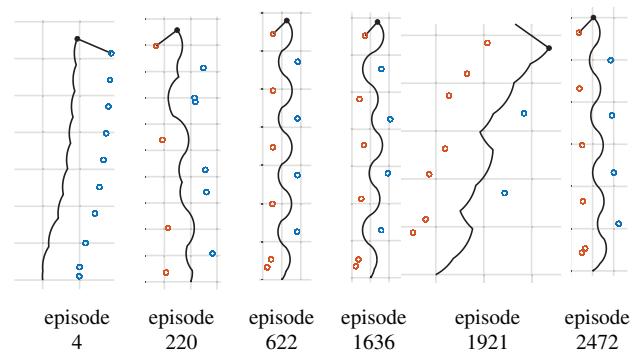


Figure 5. Learning progress. Training progress of one instance of the reinforcement learning agent for the little skate model with centre of mass trajectories and footsteps at different episodes during learning.

figure 3, starting with no forward speed and $\gamma = 1$. The first step uses a large γ and over subsequent steps minimizes it while increasing v^* , eventually settling close to the upper saddle-node bifurcation point at $\gamma = 0$, confirming the optimality of the solution discovered by RL. In the context of our model, these results suggest that, despite the vast solution space of gaits, a left-right alternating bipedal control strategy can and will be discovered and is the optimal solution for energy efficient locomotion.

4. Robotic model of bipedalism

Inspired by a range of recent robotic studies such as mimicking the legged locomotion of mudskippers [25], the reconstruction of feasible tetrapod gaits in extinct species [26], high-frequency undulatory swimming [27] and robophysical studies in general [28], we ask if it is possible to mimic benthic bipedalism experimentally to test our theoretical ideas. To start, we used a supported (simulating neutrally buoyant environments) robotic biped as shown in figure 6a. The legs were mechanically constrained to satisfy $\alpha \in [0, 2.15]$ rad, and we fixed ω , L , m and varied T to change γ . The design of the robot aims to test the planar dynamics of aquatic walking, restricting vertical oscillations of the body (but not the vertical displacement of legs) for simplicity; an unsupported system would require stabilization of vertical body attitude and position by using a tail or pectoral fins that generate lift.

Figure 6b shows a bottom view of a series of snapshots of two experiments at different times. The system was initialized from rest and with $\gamma = 0.123$ and $\gamma = 0.148$. The solid line corresponds to the centre of mass trajectory of the robot and the dots to the footstep of the snapshots. The coexistence of fixed points at $\gamma < 0.13$ was tested by initializing the robot from rest and with a flying start, i.e. an initial non-zero velocity. As expected, we observed two steady solutions: a slow and a fast gait. Note, however, as γ increases, we no longer require a non-zero velocity initialization to reach the fast solution branch, effectively demonstrating that gait transition (resting to walking), acceleration and stabilization are performed without the need for additional control. As observed in skate experiments and our model, the robot also exhibits undulatory behaviour and a regular footstep pattern. The observed versus predicted locomotion speeds are shown in figure 6c. The observed locomotion speeds are low when started from rest for $\gamma < 0.13$, but converge to the upper branch of the bifurcation diagram for $\gamma > 0.13$, with the

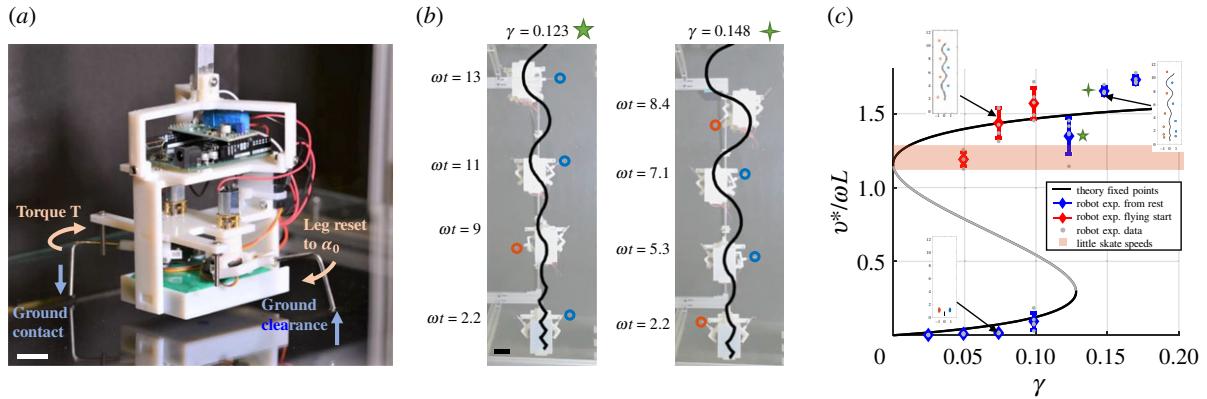


Figure 6. Robot experiments. (a) Image of supported robot with right leg in contact with the ground and left leg resetting to initial angle α_0 . Scale bar, 5 cm. (b) Bottom-up view of sequence of robot walking gait for two cases initialized from rest. The black line indicates the centre of mass trajectory of the robot. Circles indicate closed-leg contact points for corresponding picture. Scale bar, 5 cm. (c) Experimental fixed-point velocities as a function of non-dimensional torque γ . The shaded region is the range of observed little skate speeds as obtained from the video analysis in [3]. Insets show a selection of experimental trajectories of the centre of mass (black line) as a function of dimensionless position with footsteps (circles).

exception of the $\gamma = 0.123$ case where a convergence to the upper branch occurs at marginally lower γ than predicted. Due to frictional losses in the robotic system, it is not possible to replicate the zero-energy solution. Any source of friction will move the upper saddle-node bifurcation towards higher torques, analogous to the case with fluid drag shown in figure 4. The most energy efficient point in the robotic system is the gait on the upper branch at $\gamma = 0.05$ as no steady locomotion gait could be found below this non-dimensional torque on the upper branch. This is also the configuration where legs were most parallel at the leg switching point, indicative of a smooth transition with low impact losses. Altogether, the gait is completely determined in terms of a constant rate of motion ω , range of motion α_0 and energetic cost as determined by the constant torque T .

5. Discussion

Our combined theoretical, computational and robotic study has shown that in neutrally buoyant environments, organisms with rudimentary leg-like appendages can converge onto a left-right alternating bipedal locomotion strategy that is stable, energy efficient, learnable and robotically realizable. That such a simple control strategy leads to robust and efficient behaviour is reminiscent of passively stabilized dynamics in a slinky ‘walking’ down a slope and other passive dynamic walkers [29]. Such systems demonstrate that the appropriate morphology for a particular environment often leads to the most efficient behaviour with simple or no active feedback control. In the same vein, in aquatic locomotion, anaesthetized fish spontaneously exhibit undulatory gaits in a vortex street [30] that enable them to swim upstream. Our little skate robot is passive in the sense that it exhibits sustained locomotion with a constant energy source without feedback control, but also different from the previous examples as the energy is provided by an internal actuator and not an external source like gravity.

From a neuromechanical perspective, the bipedal control strategy has minimal demands: a body characterized by a simple morphology (leg length L and body inertia m), an oscillator capable of left-right alternating gait with a frequency ω , an actuating muscular force generator that leads to a torque

T and a proprioceptive sensor that can measure a geometric quantity α . The energy optimal gait in our model has a scaled velocity $v^*/\omega L \sim O(1)$, i.e. the frequency and leg length determine the locomotion speed v^* . Thus, the length and the frequency must be matched to the neuromuscular capabilities of the fins to generate a torque T , resulting in a frequency ω , which suggests limitations on the maximal value of v^* for a given leg length L . The leg length of the little skate ~ 1 cm and its leg switching frequency ~ 1 Hz thus set the locomotion speed ~ 1 cm s $^{-1}$. In aquatic locomotion, fish tailbeat frequencies range between 1 and 25 Hz and linearly correlate with the swimming speed over body length [31]. We see that the energy efficient gait as predicted in our model has the same correlation with respect to leg length. The bipedal walker also needs to have the capability to sense leg placement angle α_0 ; while this can be obtained proprioceptively in organisms, it can be enforced by mechanical constraints in our robot. This reduces the number of constraints required for a feasible bipedal strategy in benthic environments by finned vertebrates to a kinematically and dynamically matched fin length, a neuromuscular torque generator and a switch that alternates between feet.

From an energetic perspective, the reasons for a walking gait in the little skate may be a consequence of the need to forage in benthic environments and the increased cost of transport for swimming near walls [32]. While metabolic rate measurements of walking skates are yet to be recorded and will provide further insights into energy expenditure as a driver of gait selection, the passive bipedal gait presented here might help explain the energetic benefits of benthic walking in aquatic environments. It must be added that the little skate uses an alternative legged gait called punting [8], which it uses, for example, to kick-start the left-right alternating locomotion strategy. Punting uses two legs simultaneously and was not discovered in our optimization framework, but it may be the preferred gait when fast acceleration is rewarded over energy efficient locomotion.

Our study complements earlier work on the theoretical existence of zero-energy gaits in terrestrial walking [33] by showing how it arises in a minimal theoretical model for aquatic bipedalism and approximately in robot experiments. In particular, it requires the legs to be collinear at the end/beginning of every footstep, effectively reducing the degrees

of freedom in the problem. Instead of controlling the legs individually, one leg can simply mirror the motion of the other leg, reminiscent of mirror algorithms used in other impulsive robotic tasks like juggling [34]. This type of gait can also be realized using a rigid body with two attached rigid legs; walking then corresponds to alternate rotations about a vertical axis, centred about one of two legs. This is similar to the waddling gait of penguins, where lateral undulation is thought to improve the energetics of locomotion [35]. Of course such zero-energy models do not account for losses due to internal damping, cost of leg swing, contribution of leg mass to collision, fluid drag, etc. Adding fluid drag to our model, we found no qualitative difference in dynamics. Comparing the observed steady locomotion speed in the little skate $v^* \in [1.1, 1.26]$, we found that it is generally lower than our measured speeds in robots and might correspond to the gait with no energy loss (see figure 1 and electronic supplementary material).

Together, our results demonstrate the minimal requirements in a neural control strategy (constant force input, stability, learnability) while obtaining high energy efficiency (zero-cost gait) and are also consistent with biological observations in the little skate. Our study reinforces the idea that the physical environment, the morphology of the organism and the neural substrates synergistically produce a coordinated walking gait, linking to fundamental questions in passive dynamics, self-organization and embodiment [36]. Indeed, the combination of the passive dynamics associated with a minimal legged morphology that are ancient [37] and the presence of conserved neural circuits that are now known to be equally ancient [3] may well have helped pave the way for legged gaits before our aquatic ancestors transitioned to *terra firma*. Understanding how the brain, body and environment worked together in heterogeneous aquatic and terrestrial environments likely also needed proprioceptive feedback. But in reliably homogeneous environments, perhaps the simple strategy quantified here was where it all started.

6. Methods

6.1. Animal data

Kinematic data from little skates were obtained from supplementary movies in [3] with permission from the authors. Centre of mass position, body orientation and leg positions were extracted using the software Kinovea. Some characteristics of the extracted data are presented in table 1.

The animals were tested in a tank with a textured polydimethylsiloxane surface for traction of the legs with the substrate. Slip per step of the leg during stance phase varied across individuals and ranged between 0.1 mm and 1 mm, which corresponds to 0.5–5% of the step length. Angle plots of α (figure 1d; electronic supplementary material) were obtained from measuring the angle between the centreline (from tracking the connecting line between pelvic girdle and mouth) and vectors pointing from the pelvic girdle to the leg tips. Velocities of pelvic girdle as a function of time (insets in figure 1d; electronic supplementary material) were computed from filtered trajectories (local regression using weighted linear least squares and a second-degree polynomial model using a span of 10% of the total number of data points) and numerically differentiating them with respect to time. Data were made dimensionless with a leg length $L = 1.15$ cm and a frequency $\omega = 1.1$ Hz, which were extracted from the movies.

Table 1. Mean (italic) and standard deviation (parentheses) of kinematic data from three individual skates. Data averaged over 10 steps in experiment excluding initial acceleration. v_m , mean non-dimensional velocity; ϕ , phase difference of legs; α_p , peak leg angle; slip/step, normalized by leg length.

skate	v_m	ϕ	α_p	slip/step
1	1.3 (0.2)	162° (34°)	2.19 (0.7)	0.088 (0.1)
2	1.38 (0.16)	170° (10°)	2.21 (0.1)	0.12 (0.14)
3	1.24 (0.3)	186° (10°)	2.33 (0.02)	0.006 (0.06)

6.2. Reinforcement learning framework

For the model-based optimization of the little skate gait, we used an RL framework due to the obvious links between episodic and biological learning. Other optimization methods such as trajectory optimization can also find the optimal solution, but would not provide insight into the learnability of the walking gait via processes related to biological reinforcement [22]. We chose a deep deterministic policy gradient (DDPG) RL agent for the optimization of the little skate gait. DDPG [23] has the advantage of accepting continuous control inputs, which is commensurate with the biological control capabilities of the little skate. The dynamics for the RL environment are obtained by computing the next step position after placing leg l at angle α_0 on the ground and applying a torque T for $1/\omega$ seconds. This provides the new position coordinates x, y and velocities $v = (\dot{x}, \dot{y})^T$. We ignored the rotational degree of freedom of the little skate centre of mass for simplicity. At every episode, the centre of mass is placed at the origin and its velocity set to zero. The reward of step i was defined as follows:

$$R_i = -|v_y - v_T| + \Delta y \left(1 - \frac{T}{T_{\max}}\right). \quad (6.1)$$

The first term on the left-hand side penalizes variations of the end-of-step vertical component of the velocity from the target speed v_T . This term drives the system towards a constant locomotion speed v_T . The second term accounts for optimization of the cost of transport, in that it rewards the product of travelled distance Δy and negative normalized torque. T_{\max} is the maximum applicable torque in the system defined as a bound in the RL problem. The bounds for control parameters were $T \in [0, 1]$, $\omega \in [1, 1000]$, $\alpha_0 \in [0, \pi]$ and $l \in [-1, 1]$. We used Matlab's reinforcement learning toolbox to train the critic and actor networks with two fully connected layers with 400 and 300 nodes and rectifiers as activation functions (except for the actor output where we used a tanh function). The leg case (-1 left, 1 right) was treated as a continuous control variable, which was put through a signum function before its use. To test the effects of learning parameters on the converged solution, we ran combinations of values for noise variance {0.1, 0.2, 0.3}, discount factor {0.8, 0.9, 0.99} and learning rates { 5×10^{-2} , 5×10^{-3} , 5×10^{-4} }. We ran the RL routine for 5000 episodes (an episode was ended after a maximum of 30 steps or if the centre of mass surpassed the bounds at $x = \pm l$) for all combinations of parameter values and found convergence to the optimal bipedal gait in 17 of 27 cases; one of them is shown in figure 5 (all routines with learning rates of 5×10^{-2} did not converge). We further asked whether changing the relative weight of the two terms in the reward function (6.1) had an effect on the optimal solution of the gait. We ran 20 learning routines of 5000 episodes each and weighted the terms 1:3, 1:1 and 3:1. The solution yielding the highest reward was again of the type shown in figure 5 (left-right alternating strategy) and was found in 16 of 20 cases.

6.3. Robot experiment

We developed a supported legged robotic system to systematically test the model predictions. The robot body was created using PolyJet technology using VeroWhitePlus material. The robot is powered by a 6 V nickel-metal hydride battery and digitally controlled with an Arduino Uno microcontroller. A motor driver (pololu max14870) operated two 6 V DC motors (pololu 50:1 micro metal gear motor medium power) to allow for leg rotation. A servo motor per leg ensured ground contact and clearance of the leg tips (Power HD Sub-Micro Servo HD-1440A). Small rubber pads were glued to the leg tips to reduce slip. Two pins were mounted to the robot structure which prevent the legs from exceeding the angle mechanically, and the main robot structure prevented the leg angle from becoming negative, i.e. we always have $\alpha \in [0, 2.15]$. The mass of the robot was $m = 350$ g and leg length $L = 8$ cm. The robot was connected with a long and stiff aluminium bar to a ball bearing, which moves on a linear 1 m steel rod. The steel bar was mounted at an angle of 0.5° at which the ball bearing could slide along the steel rod. This resulted in a decrease in height of the bar position in direction of travel of the robot. Although this decrease in gravitational potential along the bar could potentially be used as a source for acceleration of the robot, friction inside the ball bearing resulted in a marginal velocity loss if the system is started with an initial speed v_0 . Note that this is a conservative set-up as our model predicts no velocity loss over time in the case of no leg collisions with the ground. The robot is hanging above a glass plate 90 cm in length onto which the feet could push against when activated to close ground contact. A webcam recorded the locomotion behaviour from the bottom of the glass plate at 30 fps and centre of mass trajectories obtained by tracking a blue marker on the bottom of the robot were subsequently extracted by analysing the videos using Matlab. See the electronic supplementary material for an illustration of the set-up.

The control strategy for walking was implemented as follows. Both legs are initialized at $\alpha_0 = 2.15$ before every trial. Leg switching frequency ω was set to 1.3 Hz. At switching time, both DC motors reverse their rotation direction and servo motors change their state from lifted to contact or from contact to lifted. The parameter γ was tuned by changing the leg

torque exerted by the DC motors, which was controlled by adjusting the supplied voltage set by the motor controller. See supplementary videos for various trials with a selection of γ values.

The data generated for figure 6c were obtained from five independent robot experiments per error bar. All experiments were initialized from rest except the three error bars on the upper branch in the bistable region, which were initialized with a non-zero velocity. The non-dimensional initial speed of all flying starts is $v = 2.4$ on average with standard deviation 0.4, which was large enough to push the system to a state that is attracted by the upper branch but not too large such that the speed can converge within the limited number of steps. The experiment lasted 20 steps, and the velocities correspond to the average of the last three leg transitions in the camera's view. Experiments with non-zero initial velocity moved out of the camera's field of view within only eight steps and we observed that transient behaviour had decayed by the last three steps in most cases. Trials at $\gamma \approx 0.05$, which were started with non-zero velocity often converged to the lower branch and slow velocities. The results shown for this case are the five successful cases where terminal velocity in the camera's field of view did not vanish. However, we cannot guarantee that these cases have converged or if they would further decay in a larger experimental set-up, which may explain the prediction error. In the case of $\gamma \approx 0.123$, we observe slower speeds than expected, which can be explained with the fact that the gait has not completely converged to the steady-state speed. These long transient phases were observed in simulations where γ is close but past the end of the bistable region, which corresponds well to the position of $\gamma \approx 0.05$.

Data accessibility. Animal data were retrieved from [3].

Authors' contributions. F.G. and L.M. conceived of the study; F.G. and L.M. designed the study; F.G. wrote the code; F.G. designed and built the robotic system; F.G. ran experiments and made the figures; F.G. and L.M. wrote the paper; F.G. and L.M. edited the paper.

Competing interests. We declare we have no competing interests.

Funding. We acknowledge support from the Swiss National Science foundation (F.G., grant P400P2-191115) and a MacArthur Fellowship (L.M.).

References

- Grillner S, Jessell TM. 2009 Measured motion: searching for simplicity in spinal locomotor networks. *Curr. Opin. Neurobiol.* **19**, 572–586. (doi:10.1016/j.conb.2009.10.011)
- Chevallier S, Ijspeert AJ, Ryczko D, Nagy F, Cabelguen J-M. 2008 Organisation of the spinal central pattern generators for locomotion in the salamander: biology and modelling. *Brain Res. Rev.* **57**, 147–161. (doi:10.1016/j.brainresrev.2007.07.006)
- Jung H *et al.* 2018 The ancient origins of neural substrates for land walking. *Cell* **172**, 667–682. (doi:10.1016/j.cell.2018.01.013)
- King HM, Shubin NH, Coates MI, Hale ME. 2011 Behavioral evidence for the evolution of walking and bounding before terrestrialization in sarcopterygian fishes. *Proc. Natl Acad. Sci. USA* **108**, 21 146–21 151. (doi:10.1073/pnas.1118669109)
- Flammang BE, Suvarnaraksha A, Markiewicz J, Soares D. 2016 Tetrapod-like pelvic girdle in a walking cavefish. *Sci. Rep.* **6**, 23711. (doi:10.1038/srep23711)
- Macesic LJ, Kajiura SM. 2010 Comparative punting kinematics and pelvic fin musculature of benthic batoids. *J. Morphol.* **271**, 1219–1228. (doi:10.1002/jmor.10865)
- Lucifora LO, Vassallo AI. 2002 Walking in skates (chondrichthyes, rajidae): anatomy, behaviour and analogies to tetrapod locomotion. *Biol. J. Linnean Soc.* **77**, 35–41. (doi:10.1046/j.1095-8312.2002.00085.x)
- Koester DM, Spirito CP. 2003 Punting: an unusual mode of locomotion in the little skate, leucoraja erinacea (chondrichthyes: Rajidae). *Copeia* **203**, 553–561. (doi:10.1643/CG-02-153R1)
- Garcia M, Chatterjee A, Ruina A, Coleman M. 1998 The simplest walking model: stability, complexity, and scaling. *J. Biomech. Eng.* **120**, 281–288. (doi:10.1115/1.2798313)
- Goswami A, Thuilot B, Espiau B. 1998 A study of the passive gait of a compass-like biped robot: symmetry and chaos. *Int. J. Rob. Res.* **17**, 1282–1301. (doi:10.1177/027836499801701202)
- Usherwood JR, Bertram JE. 2003 Understanding brachiation: insight from a collisional perspective. *J. Exp. Biol.* **206**, 1631–1642. (doi:10.1242/jeb.00306)
- Blickhan R, Full R. 1993 Similarity in multilegged locomotion: bouncing like a monopode. *J. Comp. Physiol. A* **173**, 509–517. (doi:10.1007/BF00197760)
- Holmes P, Full RJ, Koditschek D, Guckenheimer J. 2006 The dynamics of legged locomotion: models, analyses, and challenges. *SIAM Rev.* **48**, 207–304. (doi:10.1137/S0036144504445133)
- Chellapurath M, Stefanni S, Fiorito G, Sabatini AM, Laschi C, Calisti M. 2020 Locomotory behaviour of the intertidal marble crab (*Pachygrapsus marmoratus*) supports the underwater spring loaded

- inverted pendulum as fundamental model for punting in animals. *Bioinspir. Biomim.* **15**, 055004. (doi:10.1088/1748-3190/ab968c)
15. Brogliato B. 1999 *Nonsmooth mechanics*. London, UK: Springer.
 16. Strogatz SH. 2018 *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. Boca Raton, FL: CRC Press.
 17. Webb PW. 1989 Station-holding by three species of benthic fishes. *J. Exp. Biol.* **145**, 303–320.
 18. Hoyt DF, Taylor CR. 1981 Gait and the energetics of locomotion in horses. *Nature* **292**, 239–240. (doi:10.1038/292239a0)
 19. Ruina A, Bertram JE, Srinivasan M. 2005 A collisional model of the energetic cost of support work qualitatively explains leg sequencing in walking and galloping, pseudo-elastic leg behavior in running and the walk-to-run transition. *J. Theor. Biol.* **237**, 170–192. (doi:10.1016/j.jtbi.2005.04.004)
 20. Srinivasan M, Ruina A. 2006 Computer optimization of a minimal biped model discovers walking and running. *Nature* **439**, 72–75. (doi:10.1038/nature04113)
 21. Alexander RM. 1980 Optimum walking techniques for quadrupeds and bipeds. *J. Zool.* **192**, 97–117. (doi:10.1111/j.1469-7998.1980.tb04222.x)
 22. Glimcher PW. 2011 Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc. Natl Acad. Sci. USA* **108**(Supp. 3), 15 647–15 654. (doi:10.1073/pnas.1014269108)
 23. Lillicrap TP, Hunt JJ, Pritzel A, Heess N, Erez T, Tassa Y, Silver D, Wierstra D. 2015 Continuous control with deep reinforcement learning. (<http://arxiv.org/abs/1509.02971>)
 24. Sutton RS, Barto AG. 2018 *Reinforcement learning: an introduction*. Cambridge, MA: MIT Press.
 25. McInroe B, Astley HC, Gong C, Kawano SM, Schiebel PE, Rieser JM, Choset H, Blob RW, Goldman DL. 2016 Tail use improves performance on soft substrates in models of early vertebrate land locomotors. *Science* **353**, 154–158. (doi:10.1126/science.aaf0984)
 26. Nyakatura JA *et al.* 2019 Reverse-engineering the locomotion of a stem amniote. *Nature* **565**, 351–355. (doi:10.1038/s41586-018-0851-2)
 27. Zhu J, White C, Wainwright D, Di Santo V, Lauder G, Bart-Smith H. 2019 Tuna robotics: a high-frequency experimental platform exploring the performance space of swimming fishes. *Science Robotics* **4**, eaax4615. (doi:10.1126/scirobotics.aax4615)
 28. 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)
 29. McGeer T *et al.* 1990 Passive dynamic walking. *I. J. Robotic Res.* **9**, 62–82. (doi:10.1177/027836499000900206)
 30. Beal D, Hover F, Triantafyllou M, Liao J, Lauder GV. 2006 Passive propulsion in vortex wakes. *J. Fluid Mech.* **549**, 385–402. (doi:10.1017/S0022112005007925)
 31. Bainbridge R. 1958 The speed of swimming of fish as related to size and to the frequency and amplitude of the tail beat. *J. Exp. Biol.* **35**, 109–133.
 32. Di Santo V, Kenaley CP. 2016 Skating by: low energetic costs of swimming in a batoid fish. *J. Exp. Biol.* **219**, 1804–1807. (doi:10.1242/jeb.136358)
 33. Gomes M, Ruina A. 2011 Walking model with no energy cost. *Phys. Rev. E* **83**, 032901. (doi:10.1103/PhysRevE.83.032901)
 34. Buhler M, Koditschek DE, Kindlmann PJ. 1990 A family of robot control strategies for intermittent dynamical environments. *IEEE Control Syst. Mag.* **10**, 16–22. (doi:10.1109/37.45789)
 35. Griffin TM, Kram R. 2000 Biomechanics: penguin waddling is not wasteful. *Nature* **408**, 929. (doi:10.1038/35050167)
 36. Pfeifer R, Lungarella M, Iida F. 2007 Self-organization, embodiment, and biologically inspired robotics. *Science* **318**, 1088–1093. (doi:10.1126/science.1145803)
 37. Standen EM, Du TY, Larsson HC. 2014 Developmental plasticity and the origin of tetrapods. *Nature* **513**, 54–58. (doi:10.1038/nature13708)