

Unifying constructal theory for scale effects in running, swimming and flying

Adrian Bejan^{1,*} and James H. Marden²

¹Duke University, Department of Mechanical Engineering and Materials Science, Durham, NC 27708-0300, USA
 and ²Pennsylvania State University, Department of Biology, University Park, PA 16802, USA

*Author for correspondence (e-mail: dalford@duke.edu)

Accepted 8 November 2005

Summary

Biologists have treated the view that fundamental differences exist between running, flying and swimming as evident, because the forms of locomotion and the animals are so different: limbs and wings vs body undulations, neutrally buoyant vs weighted bodies, etc. Here we show that all forms of locomotion can be described by a single physics theory. The theory is an invocation of the principle that flow systems evolve in such a way that they destroy minimum useful energy (exergy, food). This optimization approach delivers in surprisingly direct fashion the observed relations between speed and body mass (M_b) raised to $1/6$, and between frequency (stride, flapping) and $M_b^{-1/6}$, and shows why these relations hold for running, flying and swimming. Animal locomotion is an optimized two-step intermittency: an optimal balance is achieved between the vertical loss of useful energy (lifting the body weight, which later drops), and the horizontal loss caused

by friction against the surrounding medium. The theory predicts additional features of animal design: the Strouhal number constant, which holds for running as well as flying and swimming, the proportionality between force output and mass in animal motors, and the fact that undulating swimming and flapping flight occur only if the body Reynolds number exceeds approximately 30. This theory, and the general body of work known as constructal theory, together now show that animal movement (running, flying, swimming) and fluid eddy movement (turbulent structure) are both forms of optimized intermittent movement.

Key words: design in nature, animal locomotion, optimality theory, optimal speed, maximum range speed, optimal frequency, stride frequency, wing beat frequency, Strouhal number, force output, scaling, allometry, turbulence, gravitational wave, constructal theory.

Introduction

Running, flying and swimming occur in very different physical environments. Not surprisingly then, the mechanics of moving a body on legs that contact solid ground are vastly different from what is required to achieve weight support in air, or to move a neutrally buoyant body through liquid (Alexander, 2003). Despite these differences there are strong convergences in certain functional characteristics of runners, swimmers and fliers. The stride frequency of running vertebrates (Heglund et al., 1974; Alexander and Maloio, 1984; Heglund and Taylor, 1988; Gatesy and Biewener, 1991) scales with approximately the same mass exponent ($M^{-0.17}$) as swimming fish (Drucker and Jensen, 1996). Velocity of running animals (Pennycuik, 1975; Iriarte-Diaz, 2002) scales with approximately the same mass exponent ($M^{0.17}$) as the observed and theoretically predicted speed of flying birds (Pennycuik, 1968; Tucker, 1973; Lighthill, 1974; Greenewalt, 1975; Bejan, 2000). Force output of the musculoskeletal motors of runners, swimmers and fliers conforms with surprisingly little variation to a universal mass specific value of about 60 N kg^{-1} (Marden and Allen,

2002). What explains these consistent features of animal design?

In the absence of a theory that unifies design features across different forms of locomotion, biologists have concentrated on potentially common constraints. For example, Drucker and Jensen (1996) hypothesized that the scaling of muscle shortening velocity for maximal power output during oscillatory contraction ($M^{-0.17}$; Anderson and Johnston, 1992) might explain the common scaling of stride frequency. Numerous authors have hypothesized that scale effects in locomotion are caused by constraints related to biomechanical safety factors and the need to avoid premature structural failure (McMahon, 1973, 1975; Biewener and Taylor, 1986; Biewener, 2005; Marden, 2005), or to maintain dynamic similarity (Alexander and Jayes, 1983; Alexander, 2003).

Here we take the different approach of starting not with constraints but with general and presumably universal design goals that can be used to deduce principles for optimized locomotion systems. Our approach is approximate (order of magnitude accuracy) and is not intended to account for all forms of biological variation. Rather it predicts central

tendencies. Furthermore, such a theory is not mutually exclusive of other hypotheses such as common constraints, because constraints have evolved within a design framework, i.e. perhaps theory can provide explanations for the nature of constraints.

The theory presented here follows from the more general constructal theory of the generation of flow structure in nature (Bejan, 1997, 2000, 2005). According to the constructal law, in order for a flow system to persist (to survive) it must morph over time (evolve) in such a way that it accomplishes the most based on the amount of power or fuel consumed. The latest reviews show that the constructal law accounts for spatial and temporal flow self-optimization and self-organization in animate and inanimate natural flow systems (Bejan, 1997, 2000; Poirier, 2003; Bejan and Lorente, 2004). Examples include river basins, lung design, turbulent structure, vascularization, snowflakes and mud cracks.

How can constructal theory be applied to the streams of mass flow called running, flying and swimming? In the same way that it has been applied to design features of inanimate flow systems such as the morphing of river basins and atmospheric circulation (Bejan, 1997, 2000), by examining how locomotion systems can minimize thermodynamic imperfections (friction, flow resistances) together, such that at the global level the animal moves the greatest distance while destroying minimum useful energy (or food, or 'exergy' in contemporary thermodynamics; Bejan, 1997). We show that this theoretical approach delivers in surprisingly simple and direct fashion the body-mass scaling relations for running, flying and swimming – the complete relations, the slopes and the intercepts, not just the exponents of the body mass.

There is a long and productive history of optimality models in analyses of animal locomotion (e.g. Tucker, 1973; Alexander, 1996, 2003; Ruina et al., 2005), so much so that the term maximum range speed is part of the common vocabulary of the field and instantly brings to mind a U-shaped curve of cost vs speed on which there is one speed that maximizes the ratio of distance travelled to energy expended. Our approach is similar in that it predicts maximum range speeds, but differs from previous efforts by simultaneously predicting stride/stroke frequencies and net force output, while being general across different forms of locomotion. Like other optimality models, our theory does not maintain that animals must act or be designed in the predicted fashion, only that over large size ranges and diverse taxa predictable central tendencies should emerge. Ecological factors will often favour species that move in ways other than that which optimizes distance per cost, for example where energy is abundant and the risk of being captured by active predators is high. Evolutionary history and the chance nature of mutation can also restrict the range of trait variation that has been available for selection. These and other factors should act primarily to increase the variation around predicted central tendencies.

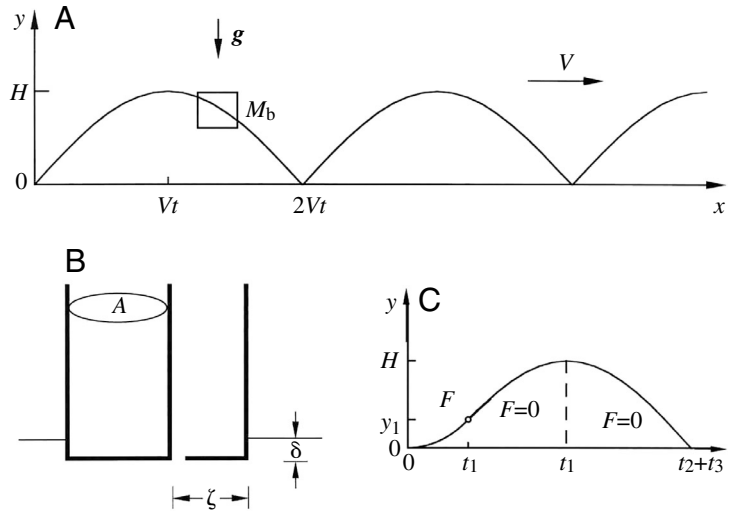


Fig. 1. (A) The periodic trajectory of a running animal, (B) model of foot contact with soft ground, and (C) history of vertical movement during one cycle, where $F=0$ refers to the time after t_1 when the foot no longer exerts a force on the ground.

Running

Consider the cyclical motion of running, Fig. 1A. In order to maintain a constant horizontal speed V , the animal must perform work to account for its two mechanisms of work destruction. One is the vertical loss W_1 : the destruction of the gravitational potential energy accumulated at the peak of each jump, $W_1=M_b g H$, where M_b is body mass and H is vertical height deviation during the jump, which is destroyed during each landing (for simplicity, we neglect elastic storage during landing). The other is the horizontal loss, W_2 , which occurs because of friction against the air, the ground, or internal friction. Internal friction is diffuse and not readily described by theory; for that reason we present a simplified theory in which all work to overcome friction is external. The constructal law calls for the minimization of the total destruction of work per distance traveled L :

$$\frac{W}{L} = \frac{W_1}{L} + \frac{W_2}{L}. \quad (1)$$

Throughout this paper we use the method of scale analysis (Bejan, 2004), which consists of solving the appropriate conservation equations as algebraic equations, with the additional simplification that dimensionless factors of order 1 are neglected. A first illustration of this method is in estimating the vertical loss term W_1/L , where $L=Vt$, and t is the time scale of frictionless fall from the height of the run (H), namely $t \sim (Hg^{-1})^{1/2}$. Since the types of animal motion that we are considering are cyclical, with motion of body parts along a roughly circular or oblong path and re-establishment of starting positions at the beginning of each cycle, it follows that height deviations, in this case the height of the run H , scales with the body length scale $L_b=(M_b/\rho_b)^{1/3}$, where ρ_b is the body density. In conclusion, $H \sim L_b$ and the vertical loss

term becomes $W_1/L = M_b g H / V t = M_b g H / (H g^{-1})^{1/2}$, which yields:

$$\frac{W_1}{L} \sim \frac{M_b g^{3/2} L_b^{1/2}}{V}. \quad (2)$$

The horizontal loss W_2/L depends on what friction effect dominates the horizontal drag. Here we consider three different drag models, and show that because they are of similar dimension, the choice of friction model does not affect the predicted optimal speed significantly.

Assume first that the drag is dominated by air friction. The air drag F_D is on the order of:

$$F_D \sim \rho_a V^2 L_b^2 C_D, \quad (3)$$

where C_D is a numerical factor of order 1 (neglected, as we would do for any value of the drag coefficient within an order of magnitude of 1) and ρ_a is the air density. The horizontal loss of useful energy is $W_2 \sim F_D L$, which means that W_2/L is replaced by F_D in Eqn 1. The total loss per unit length traveled is:

$$\frac{W}{L} \sim \left(\frac{M_b g^{3/2} L_b^{1/2}}{V} \right) + \rho_a V^2 L_b^2. \quad (4)$$

This total loss function can be minimized with respect to V by solving the equation $d(W/L)/dV=0$, which yields the scaling equation $M_b g^{3/2} L_b^{1/2} / V_{\text{opt}}^2 \sim \rho_a V_{\text{opt}}^2 L_b^2$. The result is the optimal speed:

$$V_{\text{opt}} \sim \left(\frac{\rho_b}{\rho_a} \right)^{1/3} g^{1/2} \rho_b^{-1/6} M_b^{1/6}, \quad (5)$$

where $(\rho_b/\rho_a)^{1/3} \cong 10$, because $\rho_b \cong 10^3 \text{ kg m}^{-3}$ and $\rho_a \cong 1 \text{ kg m}^{-3}$. A compilation of velocity data (Fig. 2A) for animals running over a variety of terrains shows that the speeds and their trend are anticipated well by Eqn 5. Note that the same optimal speed as in Eqn 5 is obtained if one sets equal (in an order of magnitude sense) the two terms appearing on the right side of Eqn 4. In this way, we see that to run at optimal speed is to strike a balance between the vertical loss (the first term) and the horizontal loss (the second term). Optimal running means optimal distribution of losses (imperfections) during locomotion. In this regard it is noteworthy that human runners recover approximately 50% of external kinetic energy and gravitational potential energy stored in elastic tissues (Ker et al., 1987), which means that about 50% is lost to friction, both internally in the tissues and externally to the environment. Thus, our simplified theory that considers only external friction yields an estimate of friction loss that is close to what occurs in actual elastic systems.

Consider next the case of friction dominated by contact with the ground. If the surface is flat and hard, we may use a Coulomb friction model and write that the horizontal loss is $W_2 \sim \mu F s$, where the coefficient of friction μ is a number of order 1, F is the normal force during the foot contact time t_c , and s is the foot sliding distance, $s = V t_c$. The contact time scale is dictated by the impact that the body experiences in the vertical direction, such that when the body makes contact

with the ground it is decelerated from its free fall velocity $(gH)^{1/2}$ to zero. Writing Newton's second law of motion, $F \sim M_b (gH)^{1/2} / t$, and using $H \sim L_b$, we find $W_2 \sim \mu F s \sim \mu [M_b (gH)^{1/2} / t] V t \sim \mu V M_b (g L_b)^{1/2}$, such that Eqn 1 becomes:

$$\frac{W}{L} \sim \left(\frac{M_b g^{3/2} L_b^{1/2}}{V} \right) + \mu M_b g. \quad (6)$$

The horizontal loss term $W_2/L \sim \mu M_b g$ could have been evaluated more directly by recognizing $M_b g$ as the vertical force exerted by the animal body on the ground, $\mu M_b g$ as the horizontal friction force, and $\mu M_b g L$ as the work destroyed by ground friction along the travel distance L .

The monotonic function of V obtained for the total loss in Eqn 6 indicates that there is a lower bound for the optimal running speed. Minimum work per distance is achieved when V exceeds the scale:

$$V \sim \mu^{-1} g^{1/2} \rho_b^{-1/6} M_b^{1/6}, \quad (7)$$

which is essentially the same as the scale predicted in Eqn 5 and tested against animal data (Fig. 2A).

Consider finally the model of a highly deformable ground surface such as sand, mud or snow of density ρ (Fig. 1B). A 'high enough' speed means that the accelerated terrain material does not have time to interact by friction with and entrain its neighbouring terrain material. This model is analogous to the behaviour of a pool of fluid that is hit by a blunt body at a sufficiently high speed. In summary, we assume that the sand behaves as an inviscid liquid when it is suddenly impacted by a blunt body (the foot), Fig. 1B.

The foot contact surface is A . The foot hits the ground with the vertical Galilean velocity $V_y \sim (g L_b)^{1/2}$. By analogy with drag in high-Reynolds flow, the vertical force felt by the foot during impact scales as $F_y \sim \rho V_y^2 A C_D$, where C_D is a constant of order 1. The work done by the foot to deform the sand vertically is $F_y \delta$, where δ is the depth of the sand indentation. This work is the same as W_1 , hence $F_y \delta \sim M_b g L_b$.

The foot also moves horizontally to the distance ξ through the sand. The horizontal drag force is $F_x \sim \rho V^2 A^{1/2} \delta C_D$, where $C_D \sim 1$ and $A^{1/2} \delta$ is the frontal area of the foot as it slides horizontally through the sand. The work destroyed by horizontal deformation of the sand is $W_2 \sim F_x \xi$, where $\xi = V t_c$, and the time of contact with the ground is $t_c \sim \delta / V_y$. Putting these formulae together, we find $W_2 \sim V^3 M_b^2 / \rho A^{3/2} (g L_b)^{1/2}$, and Eqn 1 becomes:

$$\frac{W}{L} \sim \left(\frac{M_b g^{3/2} L_b^{1/2}}{V} \right) + \left(\frac{V^2 M_b^2}{\rho A^{3/2} L_b} \right). \quad (8)$$

Vertical loss Horizontal loss

The optimal running speed for minimal work per unit of length traveled is:

$$V_{\text{opt}} \sim (g L_b)^{1/2} \left(\frac{\rho A^{3/2}}{M_b} \right)^{1/3}. \quad (9)$$

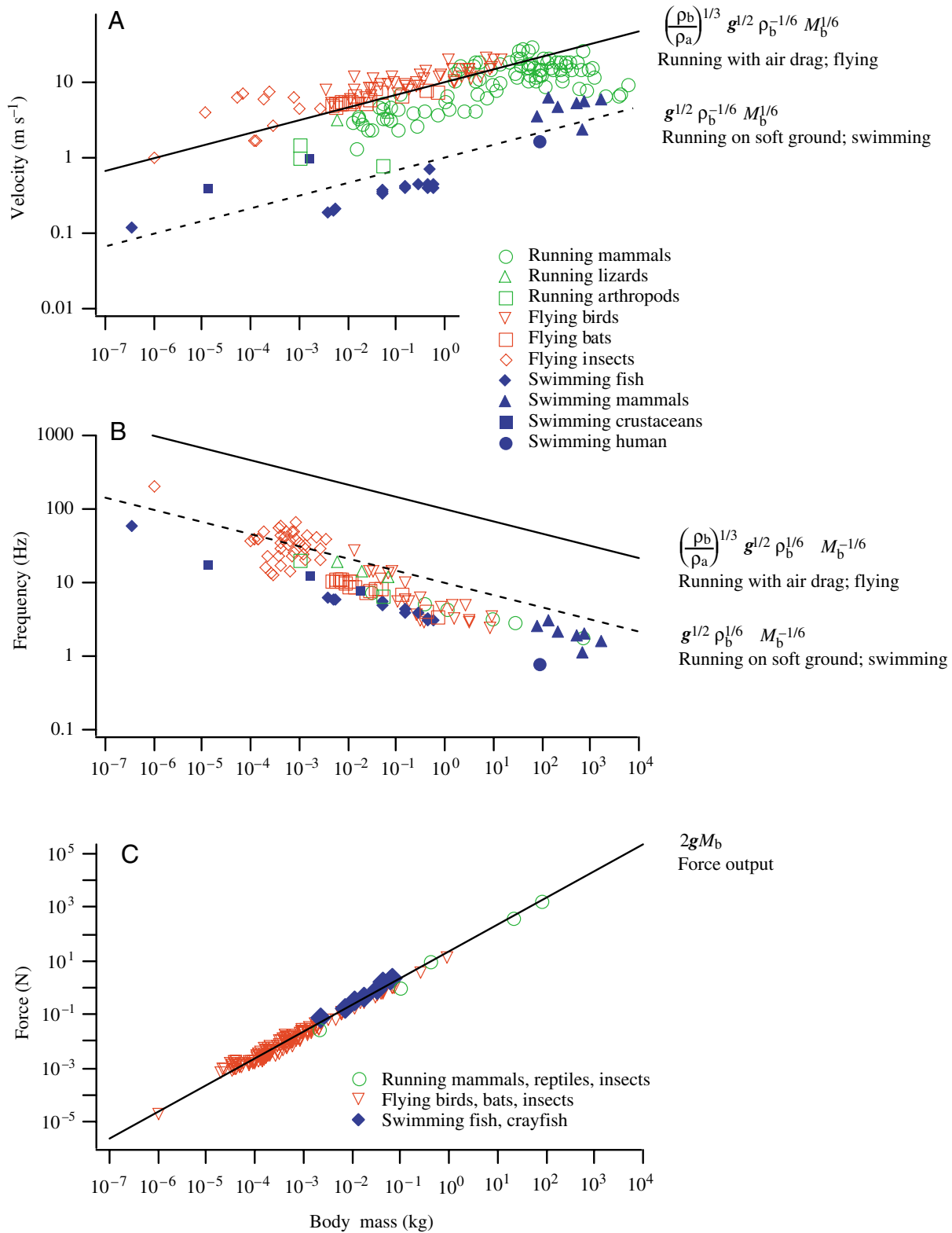


Fig. 2. Comparison of theoretical predictions with the speeds, stroke frequencies, and force outputs of a wide variety of animals (Iriarte-Diaz, 2002; Drucker and Jensen, 1996; Pennycuik, 1975; Heglund et al., 1974; Marden and Allen, 2002; Taylor et al., 2003; Rohr and Fish, 2004; Bartholomew and Casey, 1978; Marsh, 1988; Wakeling and Ellington, 1997; Marden et al., 1997; Tennekes, 1997; Kiceniuk and Jones, 1977; May, 1995; Arnott et al., 1998; Peake and Farrell, 2004; Muller and Leeuwen, 2004; Bartholomew et al., 1985; Blickhan and Full, 1987; Full and Tu, 1991). The theoretical predictions are based on scale analysis, which neglects factors of order 1 and therefore should be accurate in an order of magnitude sense.

The simplest reading of this result makes use of the rough approximation that animal bodies are geometrically similar (especially when compared over large size ranges). In this case $(\rho_b A^{3/2}/M_b)^{1/3}$ is a factor of order 1 that does not depend on M_b .

Eqn 9 shows that the optimal running speed or deformable ground is:

$$V_{\text{opt}} \sim g^{1/2} \rho_b^{-1/6} M_b^{1/6}. \quad (10)$$

This formula, as does Eqn 5, agrees in both trend and magnitude with the numerous speed data compiled by empirical studies (Fig. 2A; note that Eqn 5 and 10 are plotted on the figure, i.e. the lines on the plot are predicted from theory and are not regression fits).

The corresponding stride frequency scale is $t_{\text{opt}} \sim V_{\text{opt}}/L_b$, which yields:

$$f_{\text{opt}}^{-1} \sim g^{1/2} \rho_b^{1/6} M_b^{-1/6}. \quad (11)$$

This agrees well with the observed proportionality between stride frequency and body mass raised to -0.14 in the most rigorously defined study of scale effects on stride frequency of runners (Heglund et al., 1974). Quantitatively, the predicted frequency is approximately $10 \text{ s}^{-1} M_b^{-1/6}$, when M_b is expressed in kg (Fig. 2B).

In sum, the effect of the horizontal friction model is felt through a factor that is a dimensionless constant in the range 1–10, namely $(\rho_b/\rho_a)^{1/3}$ for air drag, μ^{-1} for hard ground, and $(\rho_b A^{3/2}/M_b)^{1/3}$ for deformable ground. The optimal running speed is a remarkably robust result, always on the order of $g^{1/2} \rho_b^{-1/6} M_b^{1/6}$. We propose that it is this robustness that accounts for the allometric law connecting V with $M_b^{1/6}$ and frequency with $M_b^{-1/6}$ in animals of so many different sizes and habitats (Fig. 2A,B).

The analysis is summarized by the observation that we have taken into account all the forces that the ground places on the leg and which dissipate through friction all the work done by the animal. We had to do this fully, without bias, without postulating that the forces are aligned with the leg or some other direction. The ground forces have one resultant, with two components, horizontal and vertical. The work dissipated by the horizontal component (W_2) was estimated in three ways in the preceding analysis. The work dissipated by the vertical ‘friction’ forces (W_1) is known exactly: it is the kinetic energy stored in the body at the peak of its cycloid-shaped trajectory. We did not have to model the friction process on the vertical because we know its total effect: W_1 . This feature alone cuts through a lot of would-be modelling, which is not relevant to the minimization of what counts, namely the total dissipation per cycle (W_1+W_2).

Additional support for this running theory comes from the calculation of the vertical force F that propels M_b to the height H during each cycle (Fig. 1C). The force F acts during a short time t_1 , when the leg makes contact with the ground, and the movement of M_b upward is governed by Newton’s second law of motion:

$$M_b \frac{d^2 y}{dt^2} = F - M_b g. \quad (12)$$

Integrating this from the initial conditions,

$$y = 0 \quad \text{and} \quad \frac{dy}{dt} = 0 \quad \text{at} \quad t = 0, \quad (13)$$

we find that the body altitude and vertical speed at the time $t=t_1$ are:

$$y_1 = \left(\frac{F}{M_b} - g \right) \frac{t_1^2}{2}, \quad (14)$$

$$\left(\frac{dy}{dt} \right)_{y_1} = \left(\frac{F}{M_b} - g \right) t_1. \quad (15)$$

When t exceeds t_1 , the body continues to move upward, reaching $y=H$ at $t=t_2$, where $dy/dt=0$. Integrating Eqn 12 with $F=0$, and satisfying the continuity conditions,

$$y = y_1 \quad \text{and} \quad \frac{dy}{dt} = \left(\frac{dy}{dt} \right)_{y_1}, \quad (16)$$

we find that

$$\frac{F}{M_b} t_1 = g t_2. \quad (17)$$

The body falls to the ground during the time

$$t_3 = \left(\frac{2H}{g} \right)^{1/2}. \quad (18)$$

Next, we note that $t_1 < t_2$, so that in an order of magnitude sense $t_2 \approx t_3$. Eliminating t_1 and t_2 from the results derived above, we obtain:

$$\left(\frac{F}{M_b g} \right)^2 = \frac{H}{y_1} \left(\frac{F}{M_b g} - 1 \right). \quad (19)$$

Because y_1 scales with H , and $H > y_1$, the final result is:

$$F \cong \frac{H}{y_1} g M_b. \quad (20)$$

In conclusion, the force produced by the leg while running at optimal speed is a multiple (of dimension 1) of the body weight. Below we show that the same theoretical force characterizes flying and swimming. Fig. 2C shows that these predictions are supported by the large volume of data on the maximal force produced by animal motors over sizes ranging from small insects to large mammals (Marden and Allen, 2002).

Flying

The proportionality between speed and $M_b^{1/6}$ has been examined previously using constructal theory, to predict speeds of animal and machine flight (Bejan, 2000). That the

same proportionality rules optimal running is not a coincidence; rather it is an illustration of the fact that a universal principle is involved. Running requires least food when during each cycle a certain amount of work is destroyed by vertical impact, and a certain amount by horizontal friction. The same balancing act is responsible for optimal flight: W_1 is the work ($M_b g H$) required to lift the body that had fallen to the vertical distance H during the cyclic time interval $t \sim (H/g)^{1/2}$. During the same period, the work spent on overcoming drag is $W_2 \sim F_D L$, where $L \sim Vt$, $F_D \sim \rho_a V^2 L_b^2 C_D$ and $C_D \sim 1$. Cycles in which the vertical and horizontal losses (W_1 , W_2) alternate in order to maintain cruising at constant altitude are sketched in Fig. 3A. The total work spent per distance traveled is:

$$\frac{W_1 + W_2}{L} \sim \frac{M_b g H}{V(H/g)^{1/2}} + \rho_a V^2 L_b^2. \quad (21)$$

The altitude increment (H) achieved during each stroke of the wing is dictated by the wing length scale, which is the length scale of the flying body, $H \sim L_b$. From Eqn 21 we learn that the spent work is minimal when (Fig. 2A):

$$V_{\text{opt}} \sim \left(\frac{\rho_b}{\rho_a} \right)^{1/3} g^{1/2} \rho_b^{-1/6} M_b^{1/6}. \quad (22)$$

The wing flapping frequency that corresponds to this optimal flying speed is $t^{-1} \sim V_{\text{opt}}/L_b$, or:

$$t_{\text{opt}}^{-1} \sim \left(\frac{\rho_b}{\rho_a} \right)^{1/3} g^{1/2} \rho_b^{1/6} M_b^{-1/6}, \quad (23)$$

which is the formula shown in Fig. 2B. The correspondence between observed wing beat frequencies and this prediction (Fig. 2B) is not as satisfying as that for velocities or stride frequencies of runners or swimmers, but nonetheless the agreement is generally within one order of magnitude. The dimensionless frequency is the Strouhal number:

$$\text{St} = \frac{t^{-1} L_b}{V_{\text{opt}}}, \quad (24)$$

which for optimal flight becomes a constant: $\text{St} \sim (\rho_a/\rho_b)^{1/3} \sim 10^{-1}$. This agrees with the large volume on St data on animal flight (Taylor et al., 2003).

Swimming

Swimming exhibits the same body-mass scaling as running and flying, not because of a coincidence, but because swimming is thermodynamically analogous to running and flying. Swimming is another example of optimal distribution of imperfections in time, or the optimization of intermittency (cf. chapter 10 in Bejan, 2000). The analogy with flying is shown in Fig. 3.

The new aspect of the present analysis of swimming is the vertical loss, $W_1 \sim M_b g L_b$. This work is spent by the fish in order to lift above itself the body of water (M_w , the same as the fish

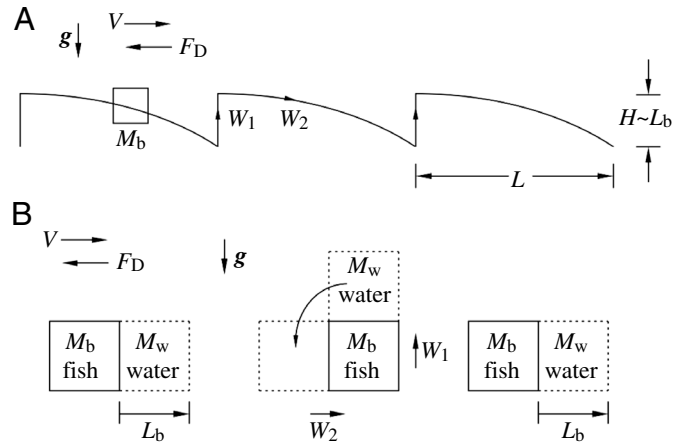


Fig. 3. (A) The periodic trajectory of a flying animal, and (B) the cyclical progress of a swimming animal.

mass because the fish and the water have nearly equal density) that it displaces during one cycle. The theory is that the product $M_w L_b$ represents W_1/g , not that during each cycle the fish lifts the mass M_w to the height L_b . The duration of the cycle, $t \sim (L_b/g)^{1/2}$, is the time in which the lifted water mass falls, to occupy the space just vacated by the fish. During this time, the fish (M_b) and its water-body partner (M_w) can be thought of as a ‘big eddy’ that will dissipate W_1 in time and space, in the wake. The fish mass M_b is as much a part of the eddy as the water mass M_w .

During the same time interval, the fish also overcomes drag by performing the work $W_2 \sim F_D L$, where $L \sim Vt \sim \rho_b V^2 L_b^2 C_D$ and $C_D \sim 1$. The fish body density ρ_b is the same as the water density. In sum, the total work spent per unit travel is:

$$\left(\frac{W_1 + W_2}{L} \right) \sim \left(\frac{M_b g L_b}{V(L_b/g)^{1/2}} \right) + \rho_b V^2 L_b^2. \quad (25)$$

The optimal swimming speed is:

$$V_{\text{opt}} \sim g^{1/2} \rho_b^{-1/6} M_b^{1/6}, \quad (26)$$

which agrees well with empirical data for animal swimming speeds (Fig. 2A). The optimal undulating frequency of the body is:

$$t_{\text{opt}}^{-1} \sim V_{\text{opt}}/L_b \sim g^{1/2} \rho_b^{1/6} M_b^{-1/6}, \quad (27)$$

(Fig. 2B), with the corresponding Strouhal number, which agrees in an order of magnitude sense with all known observations.

The net force output for travel at the speed that minimizes work per distance traveled is $2gM_b$. The force $2gM_b$ plotted in Fig. 2C is the order of magnitude of the average force exerted by the fish, which is remarkably close to the maximum force indicated by the empirical data in Fig. 2C. The average force scale $2gM_b$ also holds for flying (eqn 9.49 in Bejan, 2000, p. 239), and is comparable with the maximum force estimated in this article for running ($H/y_1 g M_b$). This is why in Fig. 2C the line $F=2gM_b$ is compared with the force data for all forms of

animal locomotion. [Note that this differs from the $\sim 6gM_b$ figure for motor force output (Marden and Allen, 2002; Marden, 2005) because in the present case M_b refers to total body mass rather than motor mass, and animal motors average about 20–75% of body mass].

To put swimming in the same theory with flying and running (Fig. 2) may seem counterintuitive, because fish are neutrally buoyant and birds are not. This ‘intuition’ has delayed the emergence of a theory that unifies swimming with the rest of locomotion. In reality, there are gravitational effects in swimming just as in flying and running. Water in front of a moving body can only be displaced upward, because water is incompressible and the lake bottom and sides are rigid. Said another way, the only conservative mechanical system (the only spring) in which the fish can store (temporarily) its stroke work W_1 is the gravitational spring of the water surface that requires a work input of size $W_1 \sim M_b g L_b$.

Elevation of the water surface has been demonstrated and used in the field of naval warfare, where certain radar systems are able to detect a moving submarine by the change in the surface water height (termed the Bernoulli hump) as it passes (unpublished US Naval Academy lecture; www.fas.org/man/dod-101/navy/docs/es310/asw_sys/asw_sys.htm) and is also evident in the data from recent studies that have examined water movement patterns around swimming fish. A two-dimensional study of water movement around the body of swimming mullet (Müller et al., 1997) shows positive pressure in front of and around the head of the fish (Fig. 4C), and suction on alternating sides of vortices that form along the fish’s posterior and in its wake. Regardless of depth, this pressure around the head must raise the water surface (at a very low angle except when the body is near the surface) over a large area centered near the anterior end of the fish, and some of this raised water subsequently falls into the vortices of the wake. A three-dimensional study of the wake of fish with a homocercal (symmetrically lobed) tail (Nauen and Lauder, 2002) found that there is a measurable downward force in these wake vortices, amounting to about 10% of the thrust force, and that there is a downward force on the head, which we interpret as the reaction force to the elevated water surface (Fig. 4B).

Previous analyses of wave effects on swimming (Hertel, 1966; Webb, 1975; Webb et al., 1991; Videler, 1993; Hughes, 2004) have focused on swimming in shallow water, where there are additional drag costs from the formation of surface waves. Surface waves (Fig. 4D) are horizontal motion of water away from the high pressure and elevated water surface height above a swimming fish (the Bernoulli hump; our W_1). As depth increases, fish must still displace water upward, but the effect of that submerged wave on the water surface becomes less and less perceptible because the area over which the free surface rises in order to accommodate W_1 is very large, and net horizontal motion of the water surface and associated frictional costs become negligible. Thus, even though the cost of surface waves decreases with increasing depth, the scales of water movement in the vicinity of the fish are dictated by the fish size

(mass, length), not by the depth under the free surface. The vertical work W_1 is non-negligible near the surface (Fig. 4D), where it is the cause of horizontal surface waves that impose additional swimming costs. What has not been appreciated previously is that this vertical work is non-negligible and is fundamental to the physics of swimming at all depths.

The swimming cycle sketched in Fig. 3B is identical to that of a shallow-water gravitational wave of depth L_b , wavelength $2L_b$ and horizontal speed $V \sim (gL_b)^{1/2}$, which is also the speed of a hydraulic jump. This is not a coincidence. The fact that this wave speed is the same as the optimized swimming speed ($g^{1/2} \rho_b^{-1/6} M_b^{1/6}$) and the observed swimming speeds of fish (Drucker and Jensen, 1996) and marine mammals (Rohr and Fish, 2004) (Fig. 2A) provides additional support for the theory that, fundamentally, swimming is an optimized intermittent movement in the gravitational field, like flying and running.

To summarize, in order to advance horizontally by one body length, the fish lifts the equivalent of a body of water of the same size as its body, to a height equivalent to the body length. What the fish does (tail flapping) is felt by the hard bottom of the lake. The hard crust of the earth supports *all* the flappers and hoppers, regardless of the medium in which the particular animal moves.

Comparison of model predictions against empirical data

So far we have only visually compared the empirical data against predictions from the theory (Fig. 2). In order to examine the fit between data and theory more rigorously, we show in Table 1 the mass scaling exponents estimated from regression slopes of \log_{10} transformed mass vs velocity, frequency, and net force output of runners, fliers and swimmers. This table shows also the mean \log_{10} difference between empirical data and predicted values (Fig. 2). Together, these comparisons address how well the data fit the theory in terms of both scaling exponents and magnitude. All of our predicted mass scaling exponents are based on the assumption of geometric similarity ($L = [M/\rho_b]^{1/3}$), but small and statistically significant deviations from this assumption tend to be the rule rather than the exception throughout the literature on animal scaling. Wingspan of birds shows a particularly large divergence from geometrical similarity, scaling as $M_b^{0.39}$ (Rayner 1987). Because larger fliers have relatively longer wings, they should have a more negative mass scaling of wingbeat frequency than predicted by our model. The observed scaling exponents in Table 1 fall near the predicted values of 0.167 and -0.167 for velocities and frequencies, with the exception of wing beat frequency of fliers (-0.26), which differs from our model in the correct direction given the unusual scaling of avian wing dimensions (see Appendix 1). In general, empirical scaling slopes vary because of the identity of the species sampled, taxonomic differences in the scaling of body dimensions, variation between studies in animal behaviour and methodology, and statistical assumptions (Martin et al., 2005). Thus, although we do not find statistical support across the board for the exact

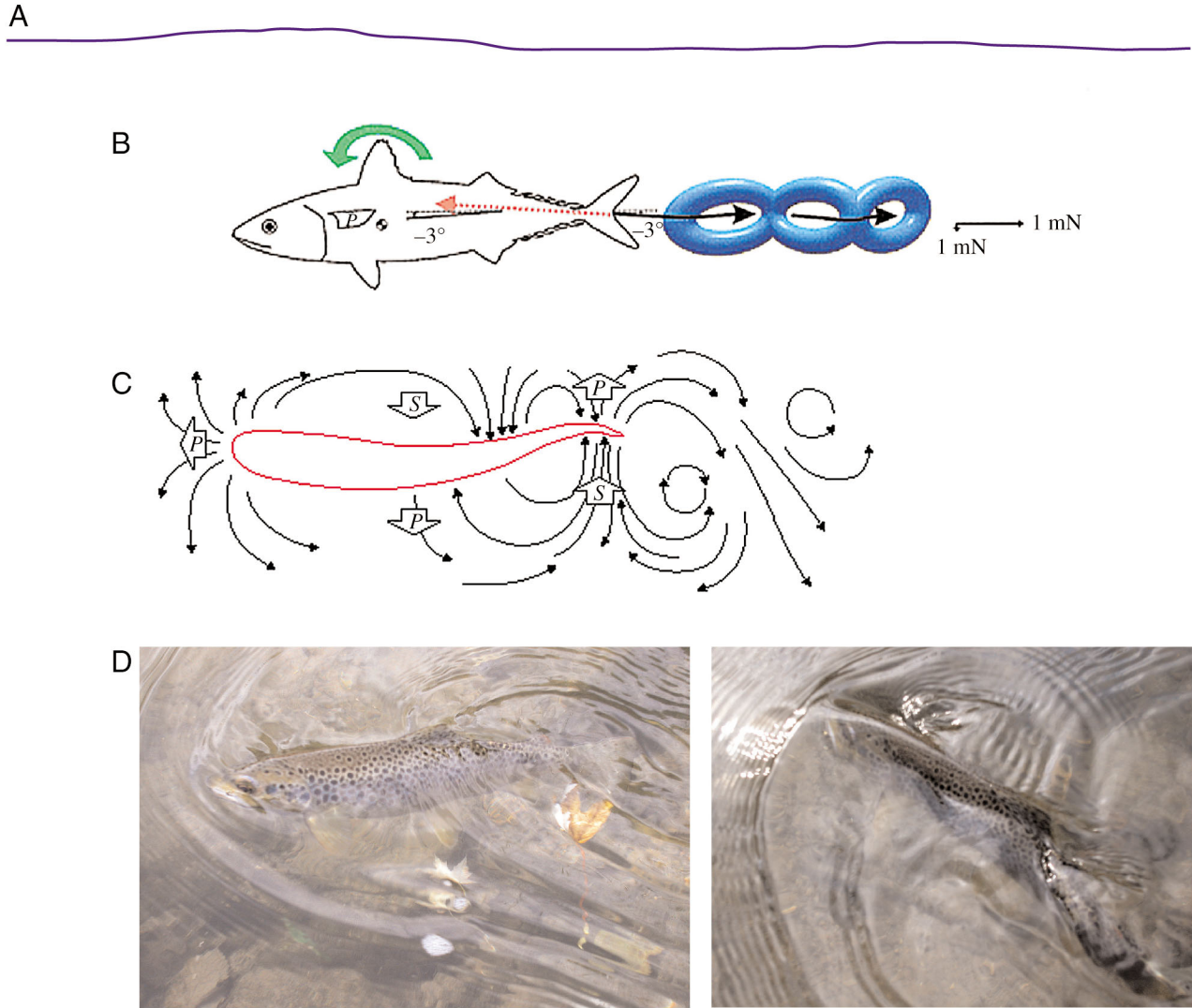


Fig. 4. The blue line (A) shows schematically and in exaggerated form the slight elevation of the water surface over a large area (which requires total work= MgL_b per body length travelled) that is centered over the anterior of a swimming body. (B,C) Reproductions of summary diagrams from two studies (Nauen and Lauder, 2002; Muller et al., 1997) of hydrodynamics of fish swimming. (B) The downward angle (-3°) of the central jet through the wake vortices formed by the tail of a mackerel, and the downward force (shown as a torque by the green arrow in this diagram) on the head. Both of these features are consistent with water lifted at the anterior end that falls at the posterior end. (C) A two-dimensional representation of water pressure and motion around a swimming mullet, with high pressure (P) at the anterior end and low pressure (S) on alternating sides of the vortices formed along the posterior. (D) Photos of a trout swimming slowly forward into a gentle current while held near the surface by a fine fishing line (not visible). Elevation of the water surface occurs over an arc at the anterior of the fish, with apparent patterns of flow and pressure quite similar to those depicted in C. Vertical displacement of the water surface during swimming at greater depth is spread over a larger area and harder to detect, but is fundamentally the same.

scaling slopes predicted from the theory, in all cases the scaling exponents derived from theory fit the data about as well as could be expected. Regarding magnitude, it is not possible to statistically examine the fit between dimensional predictions and data, other than the expectation that the data should fall within a range that spans about 0.1 to 10-fold of the predictions. As shown in Fig. 2 and Table 1, that is what we find for speeds, frequencies and force outputs. Many biologists will dismiss as meaningless a theory that has plus or minus one order of magnitude accuracy, but keep in mind

that this theory uses only density, gravity and mass, without any fitting constants, to make these predictions.

Concluding remarks

A new theory predicts, explains and organizes a body of knowledge that was growing empirically. This we have done by bringing the cruising speeds, frequencies, and force outputs of running, flying and swimming under one theory. This theory falls within the growing field of constructal theory, which has

Table 1. Comparisons of empirical data against the model predictions

Model	Observed mass scaling exponent* (<i>N</i>)	<i>P</i> value [†]	Mean log ₁₀ difference between data and dimensional analysis prediction [‡]
Running velocity	0.17±0.014 (98)	NS	
Friction=air			-0.19±0.22
Friction=ground			0.80±0.22
Flying velocity	0.13±0.010 (72)	<0.05	0.133±0.14
Swimming velocity	0.19±0.023 (27)	NS	-0.06±0.28
Running stride frequency	-0.19±0.020 (11)	NS	
Friction=air			-1.26±0.11
Friction=ground			-0.26±0.11
Flying wing beat frequency	-0.26±0.015 (77)	<0.001	-1.18±0.23
Swimming stroke frequency	-0.14±0.01 (27)	<0.05	-0.14±0.15
Running force	1.12±0.08 (4)	NS	-0.30±0.23
Flying force	0.93±0.011 (166)	<0.001	0.026±0.14
Swimming force	1.03±0.07 (8)	NS	0.16±0.09

*Estimated from the slopes of log₁₀ transformed least-squares regressions; mean ± s.e.m.

[†]*P* values are given for *t*-test of scaling exponent vs 0.167 or -0.167. NS, not significant.

[‡]Mean ± s.d.

Data sources are indicated in the legend of Fig. 2.

been used previously to account for form and design of inanimate flow structures (Bejan, 2000). Animal locomotion is no different than other flows, animate and inanimate: they all develop (morph, evolve) architecture in space and time (self-organization, self-optimization), so that they optimize the flow of material. In the past it made sense to describe the flapping frequencies of swimmers and flyers in terms of the Strouhal number (*St*). It made sense because such animals generate eddies, and because *St* is part of the language of turbulent fluid mechanics. After this unifying theory of locomotion, one can also talk about the Strouhal number of runners, $St = r^{-1} L_b / V_{opt}$, which in view of the first part of the analysis turns out to be a constant in the 0.1–10 range, just as for swimmers and flyers. The *St* constant is an optimization result of the theory, and it belongs to all flow systems with optimized intermittency, animate and inanimate.

All animals, regardless of their habitat (land, sea, air) mix air and water much more efficiently than in the absence of flow structure. Constructal theory has already predicted the emergence of turbulence, by showing that an eddy of length scale L_b , peripheral speed V and kinematic viscosity ν transports momentum across its body faster than laminar shear flow when the Reynolds number $L_b V / \nu$ exceeds approximately 30 (cf. chapter 7 in Bejan, 2000). This agrees very well with the zoology literature, which shows that undulating swimming and flapping flight (i.e. locomotion with eddies of size L_b) is possible only if $L_b V / \nu$ is greater than approximately 30 (Childress and Dudley, 2004).

And so we conclude with a promising link that this simple physics theory reveals: the generation of optimal distribution of imperfection (optimal intermittency) in running, swimming and flying is governed by the same principle as the generation of turbulent flow structure. The eddy and the animal that

produces it are the optimized ‘construct’ that travels through the medium the easiest, i.e. with least expenditure of useful energy per distance traveled.

Appendix

Bodies with two length scales

The scale analysis presented in this paper is based on the simplest geometrical model for a body that runs, flies or swims: the body geometry is represented by one length scale, $L_b \sim (M_b / \rho_b)^{1/3}$. We covered a large territory with this simple first step, and we can do more if we adopt a slightly more complex model. One reason for trying this next step is that some of the scatter (discrepancies) between the present formulae and speed and frequency data can be attributed to changes in body shape as body mass increases. Another reason is to show how the present theoretical approach can be used in future studies of more complicated living systems and processes.

Consider the analysis shown for flying in Eqn 21–23, but instead of the one-length body description (L_b), recognize that the geometry of a large bird with its wings spread out (flapping or gliding) is better captured by two length scales: the wing span L_b , which is horizontal, and the body or wing thickness, Y_b , which is vertical. By making this change, we are saying that the flying bird looks more like a flying saucer (volume $L_b^2 Y_b$) than a sphere (volume L_b^3). The body mass scale is:

$$M_b \sim \rho_b L_b^3 \lambda, \quad (\text{A1})$$

where λ is the geometric shape ratio of the two-scale body:

$$\lambda = \frac{Y_b}{L_b} < 1. \quad (\text{A2})$$

Next, we redo the analysis leading to Eqn 21. As before, for the W_1/L term we use $H \sim L_b$. To estimate F_D , we use $L_b Y_b$ in place of L_b^2 , so that the second term on the right side of Eqn 21 reads $\rho_a V^2 L_b Y_b$. In place of Eqn 22 and 23 we find:

$$V_{\text{opt}} \sim \left(\frac{\rho_b}{\rho_a} \right)^{1/3} g^{1/2} \rho_b^{-1/6} \left(\frac{M_b}{\lambda} \right)^{1/6}, \quad (\text{A3})$$

$$t_{\text{opt}}^{-1} \sim \left(\frac{\rho_b}{\rho_a} \right)^{1/3} g^{1/2} \rho_b^{1/6} \left(\frac{M_b}{\lambda} \right)^{-1/6}. \quad (\text{A4})$$

Flying bodies have smaller shape ratios (λ) when they are larger. This is true of insects, birds and airplanes alike. The fluid mechanics reasons for why this trend must exist deserve a study of their own. Here, we accept empirically the notion that λ decreases as M increases, and write that in a narrow enough range of large M values, λ behaves as:

$$\lambda \sim (\text{constant}) M^{-a}, \quad (\text{A5})$$

where $a > 0$. This means that the group (M/λ) , which appears in Eqns A3 and A4, behaves as M^m , where $m > 1$ because $m = 1 + a$. In conclusion, the M effect in Eqn A3 and A4 is:

$$V_{\text{opt}} \sim (\text{constant}) M^{m/6}, \quad (\text{A.6})$$

$$t_{\text{opt}}^{-1} \sim (\text{constant}) M^{-m/6}, \quad (\text{A.7})$$

meaning that the modulus of the exponent of M should become larger than 1/6 when M increases. This conclusion agrees with V_{opt} data for birds (e.g. fig. 9.13 in Bejan, 2000), which shows that the exponent of M in Eqn A6 becomes greater than 1/6 as M increases. The same conclusion is in agreement with observations that wing beat frequencies are more closely proportional to $M_b^{-0.26}$ rather than $M_b^{-0.17}$ (Table 1; Rayner, 1987).

In conclusion, the accuracy of the theoretical approach presented in this paper can be improved by basing it on more realistic (multi-scale) body geometries.

List of symbols and abbreviations:

a	the exponent describing how λ scales with body mass
$A^{1/2}$	frontal length scale of a foot
C_D	drag coefficient
F	force
F_D	drag force in air
F_X	drag force on foot sliding on the ground
F_y	force normal to a horizontal surface
g	gravitational acceleration
H	vertical height deviation during a cycle of locomotion
L	horizontal distance traveled during a cycle of locomotion
L_b	a characteristic length of an animal
m	a mass scaling exponent that equals 1+a
M	mass

M_b	body mass
M_w	water mass
s	foot sliding distance
St	Strouhal number
t	time
t_c	duration of foot contact with the ground during one cycle
t_{opt}^{-1}	optimal cycle frequency
V	velocity
V_{opt}	optimal velocity; maximizes distance per total work expended
V_y	vertical velocity
W_1	work expended in the vertical plane
W_2	work expended in the horizontal plane
y	vertical distance above the ground
Y_b	characteristic body length along the dorso-ventral axis
μ	coefficient of friction
δ	depth of ground indentation during foot impact
ξ	horizontal sliding distance by a foot impacting the ground
λ	ratio of characteristic lengths that describes body shape
ν	kinematic viscosity
ρ_a	air density
ρ_b	body density

We thank the following colleagues for sharing data: J. Iriarte-Diaz, F. Fish, J. Rohr, G. Taylor. This work was supported by National Science Foundation grant CTS-0001269 to A.B. and IBN-0091040 and EF-0412651 to J.H.M. We thank the organizers of the 2004 Ascona Conference (published in *J. Exp. Biol.* **208**, part 9, 2005): we are very grateful for having been invited, because the ideas developed in this paper were formed in Ascona.

References

- Alexander, R. McN. (1996). *Optima for Animals* (revised edition). Princeton: Princeton University Press.
- Alexander, R. McN. (2003). *Principles of Animal Locomotion*. Princeton: Princeton University Press.
- Alexander, R. McN. and Jayes, A. S. (1983). A dynamic similarity hypothesis for the gaits of quadrupedal mammals. *J. Zool.* **201**, 135-152.
- Alexander, R. McN. and Maloiy, G. M. O. (1984). Stride lengths and stride frequencies of primates. *J. Zool.* **202**, 577-582.
- Anderson, M. E. and Johnston, I. A. (1992). Scaling of power output in fast muscle fibers of the Atlantic cod during cyclical contractions. *J. Exp. Biol.* **170**, 143-154.
- Arnott, S. A., Neil, D. M. and Ansell, A. D. (1998). Tail-flip mechanism and size-dependent kinematics of escape swimming in the brown shrimp *Cangon crangon*. *J. Exp. Biol.* **201**, 1771-1784.
- Bartholomew, G. A. and Casey, T. M. (1978). Oxygen consumption of moths during rest, pre-flight warm-up, and flight in relation to body size and wing morphology. *J. Exp. Biol.* **76**, 11-25.
- Bartholomew, G. A., Lighton, J. R. B. and Louw, G. N. (1985). Energetics of locomotion and patterns of respiration in tenebrionid beetles from the Namib desert. *J. Comp. Physiol.* **155**, 155-162.
- Bejan, A. (1997). *Advanced Engineering Thermodynamics*, 2nd edn. New York: Wiley.
- Bejan, A. (2000). *Shape and Structure, from Engineering to Nature*. Cambridge: Cambridge University Press.

- Bejan, A.** (2004). *Convection Heat Transfer*, 3rd edn. New York: Wiley Press.
- Bejan, A.** (2005). The constructal law of organization in nature: tree-shaped flows and body size. *J. Exp. Biol.* **208**, 1677-1686.
- Bejan, A. and Lorente, S.** (2004). The constructal law and the thermodynamics of flow systems with configuration. *Int. J. Heat Mass Transfer* **47**, 3203-3214.
- Biewener, A. A.** (2005). Biomechanical consequences of scaling. *J. Exp. Biol.* **208**, 1665-1676.
- Biewener, A. A. and Taylor, C. R.** (1986). Bone strain: a determinant of gait and speed? *J. Exp. Biol.* **123**, 383-400.
- Blickhan, R. and Full, R. J.** (1987). Locomotion energetics of the ghost crab. II. Mechanics of the center of mass during walking and running. *J. Exp. Biol.* **130**, 155-174.
- Childress, S. and Dudley, R.** (2004). Transition from ciliary to flapping mode in swimming mollusc: flapping flight as a bifurcation in Re_{ω} . *J. Fluid Mech.* **498**, 257-288.
- Drucker, E. G. and Jensen, J. S.** (1996). Pectoral fin locomotion in striped surfperch. II. Scaling swimming kinematics and performance at gait transition. *J. Exp. Biol.* **199**, 2243-2252.
- Full, R. J. and Tu, M. S.** (1991). Mechanics of rapid running insects: two-, four- and six-legged locomotion. *J. Exp. Biol.* **156**, 215-231.
- Gatesy, S. M. and Biewener, A. A.** (1991). Bipedal locomotion: effects of speed, size and limb posture in birds and humans. *J. Zool.* **224**, 127-147.
- Greenewalt, C. H.** (1975). The flight of birds. The significant dimensions, their departure from the requirements of geometrical similarity, and the effect on flight aerodynamics of that departure. *Trans. Am. Phil. Soc.* **65**, 1-67.
- Heglund, N. C. and Taylor, C. R.** (1988). Speed, stride frequency and energy cost per stride: how do they change with body size and gait? *J. Exp. Biol.* **138**, 301-318.
- Heglund, N. C., Taylor, C. R. and McMahon, T. A.** (1974). Scaling stride frequency and gait to animal size: mice to horses. *Science* **186**, 1112-1113.
- Hertel, H.** (1966). *Structure Form and Movement*. New York: Reinhold.
- Hughes, N. F.** (2004). The wave-drag hypothesis: an explanation for size-based lateral segregation during the upstream migration of salmonids. *Can. J. Fish. Aquat. Sci.* **61**, 103-109.
- Iriarte-Diaz, J.** (2002). Differential scaling of locomotor performance in small and large terrestrial mammals. *J. Exp. Biol.* **205**, 2897-2908.
- Ker, R. F., Bennett, M. B., Bibby, S. R., Kester, R. C., and Alexander, R. M.** (1987). The spring in the arch of the human foot. *Nature* **325**, 147-149.
- Kiceniuk, J. W. and Jones, D. R.** (1977). The oxygen transport system in trout (*Salmo gairdneri*) during sustained exercise. *J. Exp. Biol.* **69**, 247-260.
- Lighthill, J.** (1974). Aerodynamic aspects of animal flight. *Fifth Fluid Science Lecture*, p. 30. London: British Hydrodynamics Research Association.
- Marden, J. H.** (2005). Scaling of maximum net force output by motors used for locomotion. *J. Exp. Biol.* **208**, 1653-1664.
- Marden, J. H. and Allen, L. R.** (2002). Molecules, muscles, and machines: universal characteristics of motors. *Proc. Natl. Acad. Sci. USA* **99**, 4161-4166.
- Marden, J. H., Wolf, M. R. and Weber, K. E.** (1997). Aerial performance of *Drosophila melanogaster* from populations selected for upwind flight ability. *J. Exp. Biol.* **200**, 2747-2755.
- Marsh, R. L.** (1988). Ontogenesis of contractile properties of skeletal muscle and sprint performance in the lizard *Dipsosaurus dorsalis*. *J. Exp. Biol.* **137**, 119-139.
- Martin, R. D., Genoud, M. and Hemelrijk, C. K.** (2005). Problems of allometric scaling analysis: examples from mammalian reproductive biology. *J. Exp. Biol.* **208**, 1731-1747.
- May, M. L.** (1995). Dependence of flight behavior and heat production on air temperature in the green darner dragonfly, *Anax junius* (Odonata: Aeshnidae). *J. Exp. Biol.* **198**, 2385-2392.
- McMahon, T. A.** (1973). Size and shape in biology. *Science* **179**, 1201-1204.
- McMahon, T. A.** (1975). Using body size to understand the structural design of animals: quadrupedal locomotion. *J. Appl. Physiol.* **39**, 619-627.
- Müller, U. K. and van Leeuwen, J. L.** (2004). Swimming of larval zebrafish: ontogeny of body waves and implications for locomotory development. *J. Exp. Biol.* **207**, 853-868.
- Müller, U. K., Van Den Heuvel, B. L. E., Stamhuis, E. J. and Videler, J. J.** (1997). Fish footprints: Morphology and energetics of the wake behind a continuously swimming mullet (*Chelon labrosus* Risso) *J. Exp. Biol.* **200**, 2893-2906.
- Nauen, J. C. and Lauder, G. V.** (2002). Hydrodynamics of caudal fin locomotion by chub mackerel, *Scomber japonicus* (Scombridae). *J. Exp. Biol.* **205**, 1709-1172.
- Peake, S. J. and Farrell, A. P.** (2004). Locomotory behaviour and post-exercise physiology in relation to swimming speed, gait transition and metabolism in free-swimming smallmouth bass (*Micropterus dolomieu*). *J. Exp. Biol.* **207**, 1563-1575.
- Pennycuik, C. J.** (1968). The mechanics of bird migration. *Ibis* **111**, 525-556.
- Pennycuik, C. J.** (1975). On the running of the Gnu (*Connochaetes taurinus*) and other animals. *J. Exp. Biol.* **63**, 775-799.
- Poirier, H.** (2003). A theory explains the intelligence of nature. *Sci. Vie* **1034**, 44-63.
- Rayner, J. M. V.** (1987). Form and function in avian flight. *Curr. Ornithol.* **5**, 1-66.
- Rohr, J. J. and Fish, F. E.** (2004). Strouhal numbers and optimization of swimming by odontocete cetaceans. *J. Exp. Biol.* **207**, 1633-1642.
- Ruina, A., Bertram, J. E. and 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.
- Taylor, G. K., Nudds, R. L. and Thomas, A. L. R.** (2003). Flying and swimming animals cruise at a Strouhal number tuned for high power efficiency. *Nature* **425**, 707-711.
- Tennekes, H.** (1997). *The Simple Science of Flight: From Insects to Jumbo Jets*. Cambridge (MA): MIT Press.
- Tucker, V.** (1973). Bird metabolism during flight: evaluation of a theory. *J. Exp. Biol.* **58**, 689-709.
- Videler, J. J.** (1993). *Fish Swimming*. London: Chapman and Hall.
- Wakeling, J. M.** (1997). Dragonfly flight. III. Lift and power requirements. *J. Exp. Biol.* **200**, 583-600.
- Wakeling, J. and Ellington, C.** (1997). Dragonfly flight. II. Velocities, accelerations and kinematics of flapping flight. *J. Exp. Biol.* **200**, 557-582.
- Webb, P. W.** (1975). Hydrodynamics and energetics of fish propulsion. *Bull. Fish. Res. Bd Can.* **190**, 1-158.
- Webb, P. W., Sims, D. and Schultz, W. W.** (1991). The effects of an air/water surface on the fast-start performance of rainbow trout (*Oncorhynchus mykiss*). *J. Exp. Biol.* **155**, 219-226.