The constructal law of design and evolution in nature

Adrian Bejan1,* and Sylvie Lorente2

1Department of Mechanical Engineering and Materials Science, Duke University, Box 90300, Durham, NC 27708-0300, USA
2Laboratory of Materials and Durability of Constructions, University of Toulouse, INSA, 135 avenue de Rangueil, 31077 Toulouse, France

Constructal theory is the view that (i) the generation of images of design (pattern, rhythm) in nature is a phenomenon of physics and (ii) this phenomenon is covered by a principle (the constructal law): ‘for a finite-size flow system to persist in time (to live) it must evolve such that it provides greater and greater access to the currents that flow through it’. This law is about the necessity of design to occur, and about the time direction of the phenomenon: the tape of the design evolution ‘movie’ runs such that existing configurations are replaced by globally easier flowing configurations. The constructal law has two useful sides: the prediction of natural phenomena and the strategic engineering of novel architectures, based on the constructal law, i.e. not by mimicking nature. We show that the emergence of scaling laws in inanimate (geophysical) flow systems is the same phenomenon as the emergence of allometric laws in animate (biological) flow systems. Examples are lung design, animal locomotion, vegetation, river basins, turbulent flow structure, self-lubrication and natural multi-scale porous media. This article outlines the place of the constructal law as a self-standing law in physics, which covers all the ad hoc (and contradictory) statements of optimality such as minimum entropy generation, maximum entropy generation, minimum flow resistance, maximum flow resistance, minimum time, minimum weight, uniform maximum stresses and characteristic organ sizes. Nature is configured to flow and move as a conglomerate of ‘engine and brake’ designs.

Keywords: constructal; thermodynamics; animate design; inanimate design; technology evolution; human and machine species

Why are lungs and river basins ‘vascular’?
Why do animals and rivers have ‘scaling laws’?
Why is there ‘technology evolution’?

1. THE CONSTRUCTAL LAW

In this review, we draw attention to the constructal law, which places the occurrence of design and pattern in nature on the basis of a law of physics. Constructal theory is the view that the generation of flow configuration is a universal phenomenon of all physics (e.g. figure 1), which is covered by a law of physics (the constructal law):

For a finite-size flow system to persist in time (to live) it must evolve such that it provides greater and greater access to the currents that flow through it.

(Bejan 1996, p. 815)

The universality of the design phenomenon and the constructal law is highlighted by the three questions formulated in the motto to this paper. We answer these seemingly unrelated questions by invoking one principle: the constructal law. Although lungs and animals are animate systems (bio), and river basins are considered to be inanimate (geo, not bio), they both display the same flow architecture—the dendrite. Man-made flow systems (technologies) evolve in the same way, and are predictable based on the same principle. All these classes of natural designs, the animate, inanimate and engineered (the extensions of the humans), are described by amazingly compact formulae (scaling laws), which make them easy to anticipate in spite of their great diversity.

In science, the origin (the genesis) of the configuration of flow systems has been overlooked. Design has been taken for granted—at best, it has been attributed to chance, inspiration, talent and art. In thermodynamics, our education is based on sketches of streams that flow into and out of black boxes, sketches that bear no relation to reality, to the position that each stream occupies in space and in time. The constructal law runs counter to this attitude.

Geometry or drawing is not a figure that always existed and now is available to look at, or worse to look through. The natural drawing always changes. It evolves, because it (the drawing) is the persistent movement, struggle, contortion and mechanism by
which the flow system achieves global flow access more and more easily. When the flow stops, the natural drawing becomes a flow ‘fossil’ (e.g. dry river bed, snowflake, animal skeleton, abandoned technology and the pyramids of Egypt).

In particular, the pyramids are fossils of the area-to-point flow of stones that led to the edifice. Bejan & Pépin (2006) showed that if this flow is configured such that the edifice is constructed in the easiest way, i.e. with less and less expenditure of useful energy, then the shape of the pyramid (the angle at the base) is unique, size independent and dictated by the technology of the era. The prediction is that the pyramid construction must proceed layer by layer such that the pyramid is geometrically similar to itself during its growth (like an onion).

The constructal law has been used to account for many features of ‘design’ in nature, from the tree-shaped flows of lungs and river basins, to the scaling laws of animal design, river basin design and social dynamics: for reviews, see Bejan (2000), Poirier (2003), Reis (2006) and Bejan & Lorente (2006). This body of work has two parts: (i) the use of the constructal law to predict and explain the occurrence of natural flow configurations, inanimate and animate, and (ii) the application of the constructal law as a physics principle in engineering, as a philosophy of design as science. The present article focuses on part (i). The progress on part (ii) is reviewed in a new book (Bejan & Lorente 2008).

Constructal theory is attracting many researchers and educators from many fields and in a new direction: to use the constructal law for better science and for better organization of the movement and connecting of people, goods and information (Bejan & Lorente 2005, 2006; Bejan 2006; Bejan & Merkx 2007). This direction is constructal design, and with it we seek not only better flowing configurations but also better (faster, cheaper, direct, reliable) strategies for generating the geometry that is missing from the black-box sketches.

2. THE CONCEPT AND LAW OF ‘DESIGN GENERATION’

Contemporary thermodynamics (e.g. Bejan 2006) accounts for the behaviour of systems as ‘black boxes’, without configuration. From this extreme generality stems the power and permanence of thermodynamics in physics. The first law accounts for the conservation of energy. The second law accounts for irreversibility, or the generation of entropy.

Consider, for example, one class of thermodynamic systems: closed systems operating in steady state. The power plant is one of the simplest and oldest examples of this kind. The heat input from the temperature reservoir $T_H$ is $Q_H$, the work output is $W$, and the heat rejected to the ambient temperature $T_L$ is $Q_L$. For the steady-state system considered here, the first law is the statement of energy conservation in the system,

$$Q_H = W + Q_L. \tag{2.1}$$

The second law is the statement that the system generates entropy,

$$\frac{Q_L}{T_L} - \frac{Q_H}{T_H} \geq 0, \tag{2.2}$$

where $Q_L/T_L$ is the entropy transfer out of the system, and $Q_H/T_H$ is the entropy transfer into the system. The entropy of the system ($S$) is an extensive thermodynamic property of the system (an inventory) that is constant because of the steady state. The quantity on the left-hand side is often named entropy generation ($S_{gen}$), or entropy production. The quantity $S_{gen}$ is not a thermodynamic property of the system.

It is important to note that the second law does not proclaim the ‘minimization’ or ‘maximization’ of entropy generation. It simply states that the sign of $S_{gen}$ cannot be negative. This analytical statement is an alternative to the common-sense statement that,
by themselves, all streams have the tendency to flow from high to low, i.e. one way.

There are many other important statements in thermodynamics, which have been derived by combining the first law with the second law. Examples are the formula for the Carnot limit of heat engines, the energy minimum and entropy maximum associated with evolution towards equilibrium (no flow), thermodynamic stability, the Gouy–Stodola theorem, exergy (availability) accounting, etc. These derived statements are theorems, not laws. They are about black boxes.

Here is what is new in thermodynamics, and what deserves increased scrutiny. There are many ad hoc statements of optimality (min, max) and end objective that have been used in science. They continue to be used because they work in a predictive sense, ad hoc, here and there. They are contradictory (mutually exclusive), and in many cases they speak of apples and oranges. Here are the most common examples:

(i) Entropy generation minimization and the pursuit of greater efficiency are used commonly in biology and engineering (Bejan 2006).
(ii) Entropy generation maximization (sic), not minimization, is being invoked in geophysics (Paltridge 1975).
(iii) Maximization of ‘fitness’ and ‘adaptability’ (robustness, resilience) are used in biology (see the reviews in the volume edited by Hoppeler & Weibel (2005)).
(iv) Minimization of flow resistance is invoked in river mechanics, physiology and engineering.
(v) Maximization (sic) of flow resistance is used regularly in physiology and engineering, e.g. the maximization of resistance to loss of body heat through the fur of an animal, or through the insulation of power plants and cryogenic plants, the minimization of leaks through the walls of ducts, etc. (Bejan 2006).
(vi) Minimization of travel time is used in urban traffic (Bejan & Merks 2007).
(vii) Minimization of effort and cost is the core idea of social dynamics.
(viii) Maximization of profit and utility is used in economics.
(ix) Maximization of territory is used for rationalizing the spreading of living species, deltas in the desert or empires.
(x) Uniform distribution of maximal stresses is used as an ‘axiom’ in rationalizing the design of animal bones and botanical trees.
(xi) Maximization of growth rate of flow disturbances (deformations) is invoked in the study of hydrodynamic instability and credited as the ‘origin’ of turbulence (e.g. Bejan 2000).

The list goes on. Obviously, something is working, and it is not the first law or the second law.

In 1996, the constructal law was formulated and proposed to expand thermodynamics in a fundamental way. First was the proposal to recognize that there is a universal phenomenon not covered by the first law and the second law. That phenomenon is the generation of configuration, or the generation of ‘design’ in nature. All thermodynamic systems in nature are flow systems (i.e. live, non-equilibrium systems), and they all have configuration. If they do not have it, then they acquire it, in time. The generation of configuration is ubiquitous, like other phenomena covered by other ‘laws’ in physics. Biological systems are configured. Geophysical systems are configured. Engineering and societal systems are configured. The configuration phenomenon unites the animate with the inanimate. All the other phenomena of physics (i.e. of ‘everything’) have this unifying power. Falling rocks, like falling animals, have weight, conserve energy, generate entropy, etc.

Second was the statement that this universal phenomenon should be covered by the constructal law. This law accounts for a natural tendency in time (from existing flow configurations, to easier flowing configurations). This tendency is distinct from the natural tendency summarized as the second law.

In this article, we present several examples of how the constructal law accounts for observations of many kinds (namely (i), (ii), . . . , (xi)). In fact, two kinds are sufficient for illustrating the universal reach of the constructal law, for example, the scaling rules in river basin design, and the scaling rules in animal design. One is inanimate (geo) and the other is animate (bio). Quite often, the geo is rationalized in terms of statement (ii), and the bio in terms of statement (i). Before the constructal law, these two kinds of natural design were thought to be unrelated, foreign to each other, like the disciplines and faculties of physics and biology in the university organization. We will elucidate this conflict in figure 4.

3. RIVER BASIN DESIGN

River basins evolve in time, and generate their area-to-point dendritic architectures in accordance with the constructal law. Figure 2 shows three ‘movies’ of the phenomenon. The first is the evolution of a river basin on the floor of the laboratory, under a steady and uniform artificial rain (Parker 1977). The second is a numerical simulation based on the same scenario (steady uniform rain) and a Darcy-flow model of soil erosion in which grains are removed if the local seepage velocity is high enough (Bejan 2000). The removed grains create subregions (the channels) with markedly higher permeability.

Because in the second movie the erosion property of the seepage medium (the velocity threshold needed to remove one grain) is uniform, the dendritic flow structure that emerges is symmetric. The third movie was generated with the same soil erosion mechanism, except that the erosion threshold had values distributed randomly over the square basin. The resulting dendritic pattern is asymmetric; however, its random appearance is due to the unknown (random) geological properties of the domain, and not to a random tendency of generating configuration.

The principle that generates all these configurations is deterministic. The second and third movies are reproduced in every frame when the erosion simulation process is repeated on the same basis, with the same geological erosion characteristics.
River basins of all sizes are united by firmly established empirical rules (Rosa et al. 2004), e.g. the rules of Horton, Melton and Hack at the bottom of table 1. We showed analytically that when the flow is turbulent in the fully rough regime, the flow architecture should be based on multi-scale constructs in which each mother channel is supplied by four daughter channels. Quadrupling is the constructal path to survival (to persisting in time), not dichotomy, not assemblies of eight flowing elements, etc.

Figure 2. The emergence and evolution of river basins of all sizes illustrate the natural phenomenon summarized by the constructal law. In time, existing patterns (black channels on white areas with seepage) are replaced by patterns that flow more easily. (a) Artificial river basin evolution generated under steady uniform rain on a 15 × 9 m area in a laboratory at Colorado State University (Parker 1977). The time runs from left to right. (b) Numerical simulation of seepage in a porous medium with spatially uniform erosion properties (Bejan 2000). The dislodged grains create channels with greater permeability, shown in black. (c) Seemingly random dendritic channels generated by seepage in a porous medium with spatially random erosion properties (Bejan 2000). Here, the geological properties of the flow domain are random, unknown and unpredictable. The randomness and diversity of natural river basins is due to this. The principle (the tendency) that generates the evolving flow configurations is deterministic, and not random. (b,c) The number of dislodged grains (n) is (i) 200, (ii) 400, (iii) 600 and (iv) 800.

Empirically, the average number of tributaries (daughter channels, $R_B$) is between 3 and 5. The average ratio of stream lengths (mother/longest daughter, $R_L$) is between 1.5 and 3.5. The upper part of the table shows the theoretical features read off the constructal designs (Bejan et al. 2006), where $i$ indicates the construction level, i.e. the number of times that quadrupling was used in order to construct a river basin of size $4^i A_0$ from elemental areas of size $A_0 = L_0^2$, where $L_0$ is the fixed length scale of the smallest hill slope.

The constructal law rules are $R_B = 4$ and $R_L = 2$. These are in very good agreement with the measured
scaling laws of river basins all over the world. The theoretical work showed that the constructal law anticipates all the scaling laws of river basins, not only Horton’s. For example, Melton’s correlation states that $F_s/D_{w0}$ is a constant approximately equal to 0.7. 

Hack’s correlation indicates that the length of the mainstream $L_{Mi}$ is proportional to $A^{1/3}_i$, where $A_i$ is the basin area and $b$ is an exponent between 0.5 and 0.56. In the last column of table 1, we used $b = 0.5$ to show in dimensionless terms that the constructal architecture also anticipates the trend correlated by Hack.

### 4. ANIMAL DESIGN

Scaling relations also dominate animal design, in fact they define what is meant by ‘animal design’. There is a significant literature on scaling and allometry in biology, and it was covered most recently in the references collected in the volume of review articles edited by Hoppeler & Weibel (2005), and in the books by Schmidt-Nielsen (1984), Calder (1984), Weibel (2000) and Ahlborn (2004). The best known empirical formulae relate body function to body mass ($M$). Examples of animal scaling that have been predicted based on the constructal law are as follows (Bejan 1997, 2000, 2006):

(i) Metabolic rate proportional to $M^{3/4}$.
(ii) Frequency (breathing, heart beating) proportional to $M^{-1/4}$.
(iii) Hair diameter proportional to body length scale raised to power 1/2.
(iv) Respiratory surface area proportional to $M^{7/8}$.

In addition, it was shown that the constructal law accounts for all types of animal locomotion (flying, running, swimming, speed limits in athletics), which are united by four scaling relations (Bejan & Marden 2006a, b; Charles & Bejan 2009):

(v) Speeds ($V$) proportional to $M^{1/6}$.
(vi) Frequencies (stride, flapping, fishtailing) proportional to $M^{-1/6}$.
(vii) Forces proportional to $M$.
(viii) Work requirement proportional to $M$ times the distance travelled.

The constructal scaling laws of animal locomotion were derived by viewing the ‘flow’ of animal mass in the same thermodynamic frame as the flow of water mass in the river basin. Both flow phenomena are driven by and destroy exergy: food exergy in animal movement, and gravitational exergy (potential energy) in water flow. Both phenomena achieve greater flow access through the generation of ‘design’, in time and in space.

The prediction of all measurements of animal locomotion came from analysing and optimizing the following two-step scenario. In order to advance horizontally, the animal body must ‘lose’ in two ways, vertically, by overcoming gravity ($W_2$), and horizontally, by overcoming drag ($W_1$). The total loss per unit travel is minimal when the speed, frequency and force are related to body mass in proportion with $M^{1/6}, M^{-1/6}$ and $M$. The animal data show that this is true of all flyers, runners, swimmers and record setting athletes. Furthermore, and contrary to prevailing dogma (namely that a neutrally buoyant swimmer has nothing to do with gravity), in swimming the penalty of fighting gravity ($W_2$) is as important as the penalty of fighting drag. This is why the maximum body force of a swimmer is twice its body weight ($F \approx 2 \, gM$), and why the useful energy spent per distance swum is proportional to the body weight ($Mg$) (Bejan & Marden 2006a, b; Charles & Bejan 2009).

If swimming were to have nothing to do with gravity, the empirical scaling laws of swimming would be different than the constructal formulae, which contain $g$ as a factor.

Figure 3 illustrates this scaling by using all the known flyers (insects, birds, airplanes). The straight line represents the constructal speed of flying. Two conclusions follow from the agreement demonstrated by the data in figure 3:

- All the animal mass that flows has evolved into the same $W_1 + W_2$ rhythm in order to travel longer distances with less food and fuel.
- There is no difference between the natural flying mass (birds, insects) and the engineered (humans in their airplanes). The latter are flying carpets similar to but larger than the flocks of geese.

The humans in their airplanes are one kind of human and machine species that evolves. This species evolves in our lifetime (Bejan 2000 (p. 311), 2006 (p. 789)). Technology is the extension of the
5. ORGAN SIZE

One unexpected discovery made with the constructal law is the prediction that animal organs and vehicle components should have characteristic, finite sizes (Bejan & Lorente 2002). This follows from the scaling rule (viii), which says that the useful energy (the work) that an animal or vehicle must consume to travel a specified distance $L$ is of the order $2 MgL$ (Bejan 2000; Bejan & Marden 2006a, b).

Consider the need to use a flow component (lung, heart, heat exchanger) on an animal or vehicle that moves. The size of this component is unknown. The component destroys less exergy when it is larger. For example, ducts with larger cross sections require less pumping power (exergy destroyed by fluid friction), and heat exchangers with larger heat transfer surfaces destroy less exergy as well. The first lesson from classical thermodynamics is that bigger is better. This thought comes in conflict with the constructal theory (rule (viii)), according to which the vehicle or animal must destroy an amount of exergy (fuel and food) that is proportional to the mass of the component. The cost of carrying the component on board is proportional to the mass of the component.

Important is that these two trends compete such that the total fuel and food consumption is minimum when the flow component has a certain finite size. The best organ is not the one recommended by the thermodynamics of the organ alone: that would be infinitely large. The best lung for the bird is not the one that has huge air passages. The best lung is compact, svelte and imperfect: it destroys exergy by fluid friction. This gives the impression that ‘nature makes mistakes’, but to think this way is to miss the big picture. The imperfect component makes the whole animal or vehicle the least imperfect that it can be.

6. THE ‘ENGINE & BRAKE’ DESIGN OF NATURE

The legacy of living flow systems on Earth is the same as that of inanimate flow systems. They move mass, from here to there, by destroying exergy that originates from the Sun. Animals do this clearly, based on rule (viii): the exergy used equals the moved weight ($Mg$) times the horizontal displacement ($L$). The same holds for our vehicles, on land, in air and in water. The spent fuel is proportional to the weight of the vehicle times the distance travelled.

Animal design has been morphed and perfected over millions of years. Vehicle design is evolving right now, on design tables and in factories. With reference to the engine model at the start of §2, the vehicle motor...
burns fuel in proportion to the high-temperature heating rate $Q_H$. This stream of heat is converted partially into mechanical power delivered to the wheels ($W$), while the remainder ($Q_L$ or $Q_H - W$) is dissipated into the environment. Not mentioned in thermodynamics (until now) is the fact that $W$ is itself dissipated into the ambient, because of the movement of vehicle weight ($Mg$) over the horizontal distance $L$.

To summarize, all the high-temperature heating that comes from burning fuel ($Q_H$ or the exergy associated with $Q_H$ and the high temperature of combustion; cf. Bejan 2006) is dissipated into the environment. The need for higher efficiencies in power generation (greater $W/Q_H$) is the same as the need to have more $W$ i.e. the need to move more weight over larger distances on the surface of the Earth, which is the natural phenomenon (tendency) summarized in the constructal law.

At the end of the day, when all the fuel has been burned, and all the food has been eaten, this is what animate flow systems have achieved. They have moved mass on the surface of the Earth (they have ‘mixed’ the Earth’s crust) more than in the absence of animate flow systems.

The moving animal or vehicle is equivalent to an engine connected to a brake (figure 4), first proposed by Bejan & Paynter (1976) and Bejan (1982, 2006). The power generated by muscles and motors is ultimately and necessarily dissipated by rubbing against the environment. There is no taker for the $W$ produced by the animal and vehicle. This is why the GNP of a country should be roughly proportional to the amount of fuel burned in that country (Bejan 2009).

Figure 4 also holds for the whole Earth, as a closed (note: closed does not mean isolated) thermodynamic system. Earth, with its solar heat input, heat rejection
and wheels of atmospheric and oceanic circulation, is a heat engine without shaft. Its maximized (but not ideal) mechanical power output cannot be delivered to an extraterrestrial system. Instead, the earth engine must dissipate through air, water and solid (rocks) friction and other irreversibilities (e.g. heat leaks) all the mechanical power that it produces. It does so by ‘spinning in its brake’ as fast as is necessary (and from this follow the winds and the ocean currents, which proceed along easiest routes, and flow at finite, characteristic speeds). Because the flowing Earth is a constructal heat engine, its flow configuration has evolved towards paths that generate less irreversibility. The earth engine would produce more and more power under its finite-size constraints, but then this means that it would dissipate more and more of it at the interfaces between engine and environment. Following Paltridge (1975), a principle of maximum dissipation is being invoked in geophysics: this agrees with what goes on in the brake, and is covered by the constructal law.

In engineering and biology (power plants, animals), the engines have shafts, rods, legs and wings that deliver the mechanical power to external entities that use the power (e.g. vehicles and animal bodies needing propulsion). Because the engines of engineering and biology are constructal, they morph in time towards easier flowing configurations. They evolve towards producing more mechanical power (under finiteness constraints), which, for them, means a time evolution towards less dissipation, or greater efficiency.

In sum, the constructal law shows that there is no conflict between minimum entropy generation (biology, engineering) and maximum entropy generation (geophysics), why each statement seems to work in its limited ad hoc domain, and why the constructal law is general (Reis & Bejan 2006).

Outside the engineering or biology engine, we see that all the mechanical power is destroyed through friction and other irreversibility mechanisms (e.g. transportation and manufacturing for man, animal locomotion and body heat loss to ambient). The engine and its immediate environment (the ‘brake’), as one thermodynamic system, are analogous to the entire Earth (figure 4). The flowing Earth (with all its engine + brake components, rivers, fish, turbulent eddies, etc.) accomplishes as much as any other flow architecture, animate or inanimate: it mixes the Earth’s crust most effectively—more effectively than in the absence of constructal phenomena of generation of flow configuration.

The movement of animals—the flow of animal mass on Earth—is analogous to other moving and mixing designs such as the turbulent eddies in rivers, oceans and the atmosphere. It is not an exaggeration to regard animals as self-driven packs of water, i.e. motorized vehicles of water mass, which spread and mix the Earth like the eddies in the ocean and the atmosphere. Irrefutable evidence in support of this analogy is how all these packs of biological matter have morphed and spread over larger areas, depths and altitudes, in this remarkable sequence in time: fish in water, walking fish and other animals on land, flying animals in the atmosphere, human and machine species in the air, and human and machine species in outer space and not in the opposite time direction. The balanced and intertwined flows that generate our engineering, economics and social organization are no different from the natural flow architectures of biology (animal design) and geophysics (river basins, global circulation).

One of the reviewers of the original version of the manuscript commented that the assumption of maximum dissipation is tacitly adopted in the sketch of the engine + brake design of nature concept put forth in figure 4. This is false. The engine shown in figure 4 can be ‘any’ engine, optimized or not, ideal (Carnot) or not ideal. Figure 4 is most general, i.e. as general as thermodynamics itself. Figure 4 shows that if one pursues less dissipation in the bio, geo and engineered flow architectures, then that direction of design generation is equivalent to pursuing more dissipation in the ‘brake’ that connects the constructal architecture to its environment. In the limit, minimum dissipation in the former leads the human mind on to the same design-generation path as maximum dissipation in the latter. The drawing of figure 4 is about the equivalence of the seemingly antagonistic ad hoc statements (i) and (ii) ($\S$2), and not about optimality. To repeat, the constructal law statement is about the direction of configuration evolution in time, not about min, max, opt and end design.

The reviewer’s comment is also wrong because of the history of figure 4. It is a thermodynamics drawing, first reported by Bejan & Paynter (1976) to show that the system traversed by the flow of heat from a high temperature to a low temperature is equivalent to a heat engine that dissipates its work output entirely in a brake. The heat engine is any engine. This drawing has been appearing in thermodynamics ever since (e.g. Bejan 1982, 1997, 2006), where it serves a key pedagogical role: the thermodynamic irreversibility of the heat transfer from hot to cold is made visible (palpable) when presented as dissipation of work in the brake. In the lower part of figure 4, we show that this 1976 drawing unifies the statements (i) and (ii) ($\S$2) under the constructal law, and this means that all of nature is configured to flow and move as a conglomerate of engine and brake designs.

7. VEGETATION DESIGN: THE FLOW OF WATER AND STRESSES

Trees are flow architectures that emerge during a complex evolutionary process. The generation of the tree architecture is driven by many competing demands. The tree must catch sunlight, absorb $\text{CO}_2$ and put water into the atmosphere, while competing for all these flows with its neighbours. The tree must survive droughts and resist pests. It must adapt, morph and grow towards the open space. The tree must be self-healing to survive strong winds, ice accumulation on branches and animal damage. It must have the ability to bulk up in places where stresses are higher. It must be able to distribute its stresses as uniformly as possible, so that all its fibres work hard towards the continued survival of the mechanical structure.

On the background of this complexity in demands and functionality, two flows stand out. The tree must facilitate the flow of water and stresses (i.e. it must
be strong mechanically). The demand to pass water is made abundantly clear by the geographical correlation between the presence of trees and the rate of rainfall (figure 1). This correlation is meant 'broadly speaking’, because the broad view is the key to discovering the pattern, in spite of the obvious and overwhelming diversity (random effects, local geographical variation, winds, atmospheric humidity). The demand to pass water is also made clear by the dendritic architecture, which according to the constructal law facilitates flow access between one point and a finite-size volume (§1). The demand to be strong mechanically is made clear by features such as the tapered trunks and limbs with round cross section, and other design-like features identified in this article. These features of design in solid structures facilitate the flow of stresses (Bejan & Lorente 2008; Bejan et al. 2008). The avoiding of strangulations in the flow of stresses is synonymous with the natural phenomenon of generation of ‘mechanical strength’ (cf. statement (x), §2).

According to constructal theory, ‘plants (vegetation) are completely analogous to snowflakes. They occur and survive in order to facilitate ground–air mass transfer’ (Bejan 2006, p. 770). Trees and snowflakes puzzle the mind because they are solid structures, not fluid. They are dendritic like the lungs and the river basins, but what flows through the solid is not obvious. In the snowflake, the flow is heat transfer (by thermal diffusion) from the solidification fronts to the surrounding volume, which is filled with subcooled liquid. The dendritic structure of the snowflake makes it easier for the heat currents to flow from small areas to the entire volume (Bejan 1997). Likewise, the flow access facilitated by the tree structure is the flow of water, from the ground into the atmosphere. Trees of all sizes are extended surfaces through which the ground water finds greater access to the atmosphere, so that the water completes its circuit in nature. The generation of dendritic architectures in trees is necessary (as a flow-access mechanism), in the same sense that the generation of easier-flowing river basins and atmospheric and oceanic currents (the climate) is necessary (cf. figure 1).

We relied solely on the constructal law in order to discover all the main features of plants, from root and canopy to forest (Bejan & Lorente 2008; Bejan et al. 2008). We took an integrative approach to trees as live flow systems that evolve as components of the larger whole (the environment). We treated the plant and the forest as physical flow architectures that evolve together towards greater mechanical strength against the wind, and greater access for the water flowing through the plant. Theoretical features derived from the constructal law are the tapered shape of the root and longitudinally uniform diameter and density of internal flow tubes, the near-conical shape of tree trunks and branches, the proportionality between tree length and wood mass raised to the power 1/3, the proportionality between total water mass flow rate and tree length, the proportionality between the tree flow conductance and the tree length scale raised to a power between 1 and 2, the existence of forest floor plans that facilitate ground–air flow access, the proportionality between the length scale of the tree and its rank raised to a power between −1 and −1/2 and the inverse proportionality between the tree size and number of trees of the same size. The constructal law also predicted that there must exist a characteristic ratio of leaf volume divided by total tree volume, trees of the same size must have a larger wood volume fraction in windy climates, and larger trees must pack more wood per unit of tree volume than smaller trees.

Comparisons with the empirical correlations and formulae based on ad hoc models support the constructal vegetation design, which also predicts and unifies classical empirical ‘rules’ such as Leonardo da Vinci’s rule, Huber’s rule (the proportionality between leaf specific conductivity and the specific conductivity of the stem), Zipf’s distribution and the Fibonacci sequence. These empirical rules are no longer statistical ‘coincidences’ because one of the predicted features of the tree trunk and canopy is the proportionality between the vertical spacing between two sequential branches along the trunk and the distance from the branching point to the tip of the trunk. For example, from this proportionality follows Leonardo da Vinci’s rule, which is the observed conservation of wood cross-sectional area during branching (the lower trunk cross section equals the upper trunk cross section plus the cross section of the branch).

The predicted proportionality between total water flow rate and tree length scale (height or canopy diameter) allowed us to cover an entire forest floor with a non-uniform distribution of canopy sizes such that the area transports the largest stream of water into the atmosphere. This ‘constructal forest’ of multi-scale trees reveals a ‘Zipf’ distribution of tree sizes versus tree rank, and tree sizes versus numbers of trees of the same size. Note that the Zipf rank versus frequency distribution for word usage is well known, and that it was derived from the constructal law for the distribution of city sizes (Bejan et al. 2006). The constructal hierarchy of trees in the forest adds itself to the list of constructal Zipf distributions of multi-scale flow architectures on Earth. In sum, the Zipf-distribution phenomenon is a manifestation of the constructal law in the generation of designs that facilitate the flow of water in nature (figure 1), and the distribution of city sizes and numbers on a continent (Bejan 2006, §13.4).

8. TWO FLOW MECHANISMS (STREAMS AND DIFFUSION) ARE BETTER THAN ONE

A constructual design feature that unites many of the natural flow configurations is that there is more than one mechanism at the disposal of the currents that must flow. The tree-shaped flows of nature are an icon of this duality: organized flow (streams along channels) coexists with disorganized flow (diffusion) across the spaces between channels. There is a balance between the two flow mechanisms, and between large channels and small channels (e.g. §3).

In all the river basins on Earth, the time scale of seepage down the hill slope is the same as the residence time

Phl. Trans. R. Soc. B (2010)
in the entire basin (Bejan 2006). This balance between dissimilar flow mechanisms makes the river basin flow in the same way as the lung, where there is a balance between flow resistance (or the residence time) along the airways and across the alveoli. The river basin and lung designs facilitate the flow access between one point and an infinity of points (area, volume).

The turbulent eddy is born out of the same diffusion–convection balance (Bejan 1982). Consider the flow of momentum by shear from a semi-infinite fluid of uniform speed $V$ to the rest of the fluid, which is stationary. The shear plane becomes enveloped in a layer of thickness $D$ that increases in time. If momentum transport across $D$ is by viscous diffusion (laminar flow), then $D$ grows as $(vt)^{1/2}$, where $v$ is the kinematic viscosity. If momentum is transported by streams (rolls), then $D$ acquires the scale $Vt$. Greater flow access means that laminar shear is preferred at short times, and eddies at longer times. The transition between the two is marked by the generation of configuration: the first eddy. The intersection of the $D(t)$ curves for the laminar and eddy shear leads to the prediction that transition occurs when the local Reynolds number $VDt/v$ is of order $10^2$, where $V$ and $D$ are the peripheral speed and size of the first eddy. This agrees with all observations of transition in jets, wakes, plumes, shear layers and boundary layers in forced and natural convection (Bejan 2000, 2006). This is the first time that the seed of turbulence (the first eddy, i.e. turbulence itself) is predicted from principle, i.e. not assumed (postulated) as ‘disturbances’ and ‘fluctuations’.

Dendritic solidification (e.g. snowflakes) has the same origin. In a motionless fluid medium at a temperature slightly below the solidification temperature, latent heat is released at the solidification site and flows into the subcooled medium (Bejan 1997; Ciobanas et al. 2006). How is this flow of heat facilitated in nature? It is by relying on thermal diffusion at short times and growing needles at longer times. Thermal diffusion spreads as a spherical wave of radius $(\alpha t)^{1/2}$, where $\alpha$ is the thermal diffusivity of the medium. Needles grow at constant speed parallel to themselves. This remarkable positioning of diffusive interstices relative to penetrating needles is the ‘design’, the white on which the black lines are drawn.

The same scenario rules the formation of needles and dendritic aggregates of dust particles in air and filter surfaces (Reis et al. 2006). Stony corals, bacterial colonies and plant roots favour a small-size ball shape in the beginning of their growth (Miguel 2006). Later, corals opt for larger structures with branches that invade the water flow. Nature selects diffusion at short times and ion convection at longer times (Lorente 2007). The collaboration of the two mechanisms towards greater access is shown in general terms in figure 5.

### 9. THE ORIGIN OF MULTI-SCALE POROUS MEDIA

Natural porous flow structures also exhibit multiple scales and non-uniform distribution of length scales through the available space. Such multi-scale flow structures are anticipated based on the constructal law (Lorente & Bejan 2006). We showed this by exploring the flow properties of the dendritic flow architecture proposed in figure 6. The idea is to connect two parallel lines (or two parallel planes) with trees that alternate with upside down trees. The resulting dendritic pattern connects the bottom boundary of the flow domain with the top boundary.

This alternating sequence of point-to-line trees constitutes a vasculature between the two parallel boundaries of the porous body. The fluid flows in the same direction through all the trees, e.g. upward
in figure 6. The flow access between the points of one line and the points of a parallel line can be viewed as a sequence of point-to-line flow access structures. The building block with which figure 6 is constructed was proposed by Lorente et al. (2002), where it was based on optimally shaped rectangular areas, each area allocated to one channel. We questioned how effective the tree structure of figure 6 is relative to a reference architecture: an array of \(N\) equidistant parallel tubes perpendicular to the two lines, each tube with diameter \(D\). This reference structure carries the same total flow rate in the same total tube volume and over the same area. The structure has one degree of freedom, the tube diameter \(D\), or the number of parallel tubes. The pressure drop along the reference structure \(\Delta P_{\text{ref}} = \frac{1}{\eta}\) is derivable analytically. When the \(d\) spacing is the same as the spacing between parallel tubes (figure 6), the two global flow resistances form the ratio

\[
\frac{\Delta P_{\text{trees}}}{\Delta P_{\text{ref}}} \approx \frac{14}{2\pi}.
\]

When the number of branching levels is 4 or larger, the tree-shaped architecture offers greater access to the flow that permeates through the porous structure. The superiority of the tree design increases fast as \(n\) increases: when \(n = 7\), the ratio \(\Delta P/\Delta P_{\text{ref}}\) is as low as 1/10.

When the available flow scales are sufficiently small \((d)\), the flow architecture should have trees, and not parallel channels, i.e. not single scale. From this follows the prediction of the multi-scale and non-uniform character of natural porous media: large numbers of small pores and few large pores (known as ‘pipes’ in hill slope hydrology). This prediction is crucial because it means that the apparent ‘diversity’ and ‘randomness’ are consistent with and predictable from the constructal law.

The observer who looks from the outside at the porous medium (e.g. from the top of figure 6) sees a few large pores surrounded by many small pores. From this viewing position, the porous medium appears to have only two main scales, which is why natural porous media are also described as ‘bidisperse’.

The structural properties uncovered based on the constructal law are qualitatively the same as those of natural porous materials (e.g. the soil of the hill slope in a river drainage basin, where two scales dominate: fine porous soil with seepage, and larger pores (pipes) embedded in the fine structure). The constructal law also predicts that ‘vascular’ designs must occur in nature, and that they must be stepwise more complex as they become larger (Kim et al. 2008).
10. ‘ON FLOW AND FORM’: THE DESIGN PHENOMON UNITES THE ANIMATE WITH THE INANIMATE

The three questions posed as the motto of this paper stress the unifying character of the constructal law. The natural phenomenon of configuration generation and the constructal law does not recognize the presumed walls between geophysical flow systems (river basins, lightning, turbulence, atmospheric and oceanic circulation), biological flow systems (animals, vegetation), social dynamics and technology.

The tree-shaped flows are examples of how the emerging and evolving drawing unites the animate with the inanimate. Other examples are in figure 1: the flow of water in the form of animate and inanimate flow systems with configuration. The phenomenon and principle of configuration generation and evolution is what the constructal law brings to thermodynamics.

There are many other flow phenomena that are not included in figure 1. One is the generation of veins of flowing molten rock in the Earth’s crust (Carrigan & Eichelberger 1990) where the viscosity of the isothermal flowing liquid (e.g. lava) varies across the vein. The composition of the molten rock varies with radius, such that the less viscous liquid occupies the periphery of the cross section, and lubricates the flow of the more viscous liquid that occupies the core of the vein. This phenomenon is due to the variation of chemical composition in molten rock, and not to the variation of temperature over the cross section (i.e. not to the inverse relation between viscosity and temperature). This was demonstrated in the laboratory (Carrigan 1994), and it shows that self-lubrication is a manifestation of the natural tendency summarized in the constructal law. Self-lubrication is a constructal phenomenon, like the self-shaping for drag reduction, which is visible in the round cross sections of all the ‘veins’ that we observe (earth worms, blood vessels, subterranean rivers now visited as caves).

The occurrence of the same configuration in animate and inanimate flow systems is not a new observation. The similarity between dendritic crystals and vegetation was noted by Isaac Newton in his attempt to develop a physical theory that accounts for all known natural phenomena (Newman 2006). Newton believed that metals and minerals ‘vegetated’ (i.e. ‘grew’) within the Earth, the subterranean mineral veins corresponding to the branches of terrestrial trees.

The view that growth leads to form is the prevailing paradigm, made popular by D’Arcy Thompson’s book On growth and form. With the constructal law, we discover an entirely different view that may be called On flow and form. It is flow, and not growth, that generates form.

This work was supported by a grant from the National Science Foundation.

REFERENCES


Bejan, A. 2000 Shape and structure, from engineering to nature. Cambridge, MA: Cambridge University Press.


Schmidt-Nielsen, K. 1984 Scaling: why is animal size so important. Cambridge, MA: Cambridge University Press.
