Constructal Theory of Social Dynamics

Constructal Theory of Social Dynamics

Edited by

Adrian Bejan

Department of Mechanical Engineering and Materials Science
Duke University
Durham, NC, USA

and

Gilbert W. Merkx

Department of Sociology
Duke University
Durham, NC, USA

Springer

Adrian Bejan
Department of Mechanical Engineering and Materials Science
Duke University
Durham, NC, USA

Gilbert W. Merkx
Department of Sociology
Duke University
Durham, NC, USA

Constructal Theory of Social Dynamics

Library of Congress Control Number: 2006939575

ISBN-13: 978-0-387-47680-3 e-ISBN-13: 978-0-387-47681-0

Printed on acid-free paper.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1

springer.com

Contents

Preface

Society is a “live” flow system, perhaps the most complex and puzzling we know. It is a jungle of flow systems—a vast multiscale system of systems—with organization, pattern, hierarchies, and usefulness (design). It is the most difficult to comprehend because we, the individuals who try to make sense of it, are inside the flow system. Difficult, because each of us is like an alveolus in the lung, an eddy in a turbulent river, or a leaf on a tree branch. From such a position of singularity, which is *identical* in rank to the positions of enormous numbers of individuals, it is a formidable task to see and describe the big picture—the lung, the river basin, and the forest.

Man’s great fortune has been the fact that Nature has shape, structure, configuration, pattern, rhythm, and similarity. From this stroke of luck, science was born and developed to the present day, where it is responsible for our physical and intellectual well-being. The puzzling architecture and history of society has many things in common with the architecture and evolution of other complex (but simpler) flow systems: blood vascularization, river basins and deltas, animal movement, turbulence, respiration, dendritic solidification, etc. Coincidences that occur in the billions are loud hints that a universal phenomenon is in play. Is there a single physics principle from which the phenomenon of configuration and rhythm can be *deduced* without recourse to empiricism?

In this book we show that there is such a principle, and it is based on the common observation that if a flow system (e.g., river basin, vascularized tissue, city traffic) is endowed with sufficient freedom to change its configuration, the system exhibits configurations that provide progressively better access routes for the currents that flow. Observations of this kind come in the billions, and they mean one thing: a time arrow is associated with the sequence of flow configurations that constitutes the existence—the survival—of the system. Existing drawings are replaced by easier-flowing drawings. This physics principle is the *constructal law* of the generation of configuration in Nature: “For a finite size flow system to persist in time (to survive) its configuration must evolve in such a way that it provides easier and easier access to the currents that flow through it.”

At Duke, where constructal theory began by accident in 1996 as a thermodynamics principle that unites physics with biology and engineering, we have stumbled upon another accident: scientists and sociologists view the generation of design in societies based on the same principle. Duke is a wonderful place not because of beautiful gardens and basketball, but because of *freedom*. Freedom is good for all design, from the better-flowing river basins to the faster, cheaper, and safer flowing rivers of people and goods (our society with all its live tree flows), all the way to the design called "better science."

Freedom brought the two of us together, a sociologist and an engineering scientist, and we were soon joined in this fertile discussion by our prominent colleagues Ed Tiryakian and Ken Land. Together we decided that the élan that constructal theory had generated in science is so contagious, and the theory itself so commonsense, concise, and useful, that it deserves to be discussed more broadly with colleagues throughout social sciences. We proposed this vision to the Human and Social Dynamics program of the National Science Foundation, which gave us an exploratory grant to "develop a community of scholars around the constructal theory of social dynamics."

This book is the first of its kind in this new field. It is the first account of the ideas, results, and future plans that came out of putting scientists, sociologists, and engineers together. The chapters of this book are based on the contributions made by prominent invited speakers at the First International Workshop on the Constructal Theory of Social Dynamics, which was held on 4–5 April 2006 at Duke University. We wish to thank the authors for their contributions to the workshop and to this book:

Prof. Sylvie Lorente
Prof. Heitor Reis
Prof. Antonio Miguel
Mr. Stephen Périn
Prof. Edward Tiryakian
Prof. John Staddon
Prof. Anthony Oberschall
Prof. Kenneth Land
Prof. Carter Butts
Ms. Miruna Petrescu-Prahova
Ms. Lorien Jasny
Dr. Franca Morroni
Dr. John Angle
Mr. Cyrus Amoozegar
Mr. Jean-Christophe Danaës

The constructal theory of social dynamics developed in this book surprises even us with the breadth and freshness of the territory that it covers. Major threads of this emerging theory of social organization are as follows:

- The organized multiscale distribution of living settlements. The idea is to place the community–community access in geometric terms, and to optimize it everywhere, subject to space constraints. Allocation of territory to movement (people, goods, information) is the fundamental idea.
- The occurrence of multiscale structure inside a settlement. In a city, for example, we see a compounding of scales, and each flowing thing has its

own hierarchy of scales. One example is how small streets coexist with larger (fewer) streets, and how the latter sustain a single artery. There are macroscopic features that appear in the largest cities (finger-shaped growth, beltways) that may be attributed to the same global principle of maximization of access.

- Development, and the connection between "flowing" societies, advancement, and prosperity. There is an opportunity to exploit the constructal idea of the need to be free to change the flow configuration, and connect it with the Darwinian view that the living constructs that prosper are those that possess the greatest ability to change.
- Migration patterns on the globe, in space and in time. Where and when people settle may be random individually, but the society appears to be the result of global optimization.
- Globalization, and the problematic aspects of overcoming obstacles to efficient flows, e.g., investment funds from the public and private sectors.

In sum, this book is about the tearing down of fences that are presumed to exist between the most central fields of human thought. To tear down fences means the opposite of "to destroy." It means *to construct* a far bigger tent that covers the designs (the bodies of knowledge) of historically separate fields.

Science is our knowledge of how nature works. Nature is *everything*, including engineering and society. Our knowledge is condensed in simple statements (thoughts, connections), which evolve in time by being replaced by simpler statements. We "know more" because of this evolution in time, not because brains become bigger and neurons smaller and more numerous. Our finite-size brains keep up with the steady inflow of new information through a process of simplification by replacement: in time, and stepwise, bulky catalogs of empirical information (e.g., measurements, observations, data, complex empirical models) are replaced by much simpler summarizing statements (e.g., concepts, formulas, constitutive relations, laws). A hierarchy of statements emerges along the way: it emerges naturally, because it is better than what existed before.

The simplest and most universal are the laws. The bulky and the laborious are being replaced by the compact and the fast. In time, science optimizes and organizes itself in the same way that a river basin evolves: toward configurations (links, connections, design) that provide faster access, or easier flowing.

The hierarchy that science exhibited at every stage in the history of its development is an expression of its never-ending struggle to optimize and redesign itself. Hierarchy means that measurements, ad hoc assumptions, and empirical models come in huge number, a "continuum" above which the compact statements (the laws) rise as needle-shaped peaks. Both are needed, the numerous and the singular. One class of flows (information links) sustains the other.

Civilization with all its constructs (science, religion, language, writing, etc.) is this never-ending physics of generation of new configurations, from the flow of mass, energy, and knowledge to the world migration of the special persons to whom ideas occur (the creative). Good ideas travel and persist. Better-flowing configurations replace existing configurations. Empirical facts are extremely numerous, like the hill slopes of a river basin. The laws are the extremely few big rivers, the Seine and the Danube. This book is about the big river of all "live" flow systems, including social dynamics: the constructal law.

Adrian Bejan
Gilbert W. Merkx

Duke University
December 2006

List of Contributors

Adrian Bejan
Department of Mechanical Engineering and Materials Science
Duke University
Durham, NC 27708-0300
abejan@duke.edu

Gilbert W. Merkx
Department of Sociology
Duke University
Durham, NC 27708-0006
gilbert.merkx@duke.edu

Sylvie Lorente
National Institute of Applied Sciences
Department of Civil Engineering
Laboratory of Materials and Durability
of Constructions, 31077 Toulouse, France
Lorente@insa-toulouse.fr

A. Heitor Reis
Departamento de Fisica
Universidade de Évora
Colegio Luis Verney
Rua Romao Ramalho, 59
7000-67 1 Évora Portugal
ahr@uevora.pt

Antonio F. Miguel
Departamento de Fisica
Universidade de Évora
Colegio Luis Verney
Rua Romao Ramalho, 59
7000-67 1 Évora Portugal
afm@uevora.pt

Stephen Périn
NeoSYS
7 rue du Théâtre
91300 Massy, France
stephen.perin@neosys.fr

Edward A. Tiryakian
Professor Emeritus
Department of Sociology
Duke University
Durham, NC 27708-0088
durkhm@soc.duke.edu

John E. R. Staddon
James B. Duke Professor
Duke University
Durham, NC 27708-1050
Staddon@psych.duke.edu

Anthony Oberschall
Emeritus professor
University of North Carolina at Chapel Hill
Chapel Hill, NC 27514
tonob@email.unc.edu

Kenneth G. Manton, **Kenneth C. Land**, and Eric Stallard
Department of Sociology and the Center for Demographic Studies
Duke University
Durham, NC 27708-0088
kland@soc.duke.edu

Carter T. Butts
Department of Sociology
University of California, Irvine
2145 Social Science Plaza A
Irvine, CA 92697-5100
buttsc@uci.edu

Miruna Petrescu-Prahova
Department of Sociology
University of California, Irvine
2145 Social Science Plaza A
Irvine, CA 92697-5100
mpetresc@uci.edu

Lorien Jasny
Department of Sociology
University of California, Irvine
2145 Social Science Plaza A
Irvine, CA 92697-5100
ljasny@uci.edu

Franca Morroni
DEXIA Asset Management
Head of SRT Analysis
Brussels, Belgium
fmorroni@yahoo.fr

John Angle
Economic Research Service
US Department of Agriculture
1800 M Street NW, Room N4097
Washington, DC, 20036-5831
jackangle@earthlink.com

Cyrus Amoozegar
Department of Mechanical Engineering and Materials Science
Duke University
Durham, NC 27708-0300
cyrus.amoozegar@duke.edu

Jean-Christophe Danaës
Cognitive Computer Science Department
University of Quebec at Montreal
Montreal, Quebec, Canada
denaes.jean-christophe@uqam.ca

Chapter 1
The Constructal Law in Nature and Society

Adrian Bejan

1.1. The Constructal Law

Society with all its layers and features of organization is a flow system. It is a "live" system, perhaps the most complex and puzzling we know. It is the most difficult to comprehend because we, the individuals who try to make sense of it, are inside the flow system. Each of us is like an alveolus in the lung, an eddy in a turbulent river, or a leaf on a tree branch. From such a position* of singularity, which is identical in rank to the positions of enormous numbers of individuals, it is a formidable task to see and describe the big picture—the lung, the river basin, and the forest.

Nature impresses us with shape, structure, configuration, pattern, rhythm, and similarity. This was our stroke of luck. From it, science was born and developed to the present day, where it is responsible for our physical and intellectual well-being. The puzzling architecture and history of society has many things in common with the architecture and evolution of other complex (but simpler) flow systems: blood vascularization, river basins and deltas, animal movement, respiration, dendritic solidification, etc. Coincidences that occur in the billions are loud hints that a universal phenomenon is in play. Is there a single physics principle from which the phenomenon of configuration and rhythm can be *deduced* without recourse to empiricism?

There is such a principle, and it is based on the common (universal) observation that if a flow system (e.g., river basin, blood vessel) is endowed with sufficient freedom to change its configuration, the system exhibits configurations that provide progressively better access routes for the currents that flow. Observations of this kind come in billions, and they mean one thing: a time arrow is associated with the sequence of flow configurations that constitutes the existence of the system. Existing drawings are replaced by easier-flowing drawings.

* Here, the meaning of position is geometric. The individual is a *particular* point of view in space. That point is occupied by *this* individual (his or her view of the world) and not by anybody else.

I formulated this principle in 1996 as the *constructal law* of the generation of flow configuration (Bejan 1996, 1997a–c):

> For a finite size flow system to persist in time (to survive) its configuration must evolve in such a way that it provides easier and easier access to the currents that flow through it.

This law is the basis for the *constructal theory* of organization in nature, which was first summarized in book form in Bejan (1997c). Today this body of work represents a new extension of physics: the thermodynamics of flow systems with configuration (Bejan and Lorente 2004, 2005).

To see why the constructal law is a law of physics, ask why the constructal law is different than (i.e., distinct from, or complementary to) the other laws of thermodynamics. Think of an isolated thermodynamic system that is initially in a state of internal nonuniformity (e.g., regions of higher and lower pressures or temperature, separated by internal partitions that suddenly break). The first and second laws account for billions of observations that describe a tendency in time, a time arrow: if enough time passes, the isolated system settles into a state of equilibrium (no internal flows, maximum entropy at constant energy, etc.). The first and second laws speak of a black box. They say nothing about the configurations (the drawings) of the things that flow. Classical thermodynamics was not concerned with the configurations of its nonequilibrium (flow) systems.

This tendency, this time sequence of drawings that the flow system exhibits as it evolves, is the phenomenon covered by the constructal law: not the drawings per se, but the time direction in which they morph if given freedom. No drawing in nature is “predetermined” or “destined” to be or to become a particular image. The actual evolution or lack of evolution (rigidity) of the drawing depends on many factors, which are mostly random. One cannot count on having the freedom to morph in peace (undisturbed).

Once again, the juxtaposition of the constructal law with the laws of classical thermodynamics can be useful. No isolated system in nature is predetermined or destined to end up in a state of mathematically uniform intensive properties so that all future flows are ruled out. One cannot count on the removal of all the internal constraints. One can count even less on anything being left in peace, in isolation.

As a thought, the second law proclaims the existence of a “final” state: the concept of equilibrium in an isolated system, at sufficiently long times. Similarly, the constructal law proclaims the existence of a concept: the equilibrium flow architecture, when all possibilities of increasing morphing freedom have been exhausted.

Constructal theory is now a fast-growing field with contributions from many sources, which have been reviewed on several occasions (Poirier 2003; Lewins 2003; Rosa et al. 2004; Torre 2004; Upham and Wolo 2004; Bejan and Lorente 2006; Reis 2006). The basic idea, however, is that constructal theory is the 1996 law cited at the start of this section.

The constructal law statement is general: it does not use words such as *tree, complex* versus *simple*, and *natural* versus *engineered*. How to deduce a class of flow configurations by invoking the constructal law is an entirely different (separate, subsequent) thought, which should not be confused with the constructal law. There are several (not many) classes of flow configurations, and each class can be derived from the constructal law in several ways, analytically (pencil and paper) or numerically, approximately or more accurately, blindly (via random search) or using intelligence (strategy, short cuts), etc. Classes that we have treated in detail, and by several methods, are the cross-sectional shapes of ducts, the cross-sectional shapes of rivers, internal spacings, and tree-shaped architectures (Bejan 1997c, 2000, 2006; Bejan and Lorente 2005).

Regarding trees, our group treated them not as models* (many have published and continue to publish models), but as fundamental access-maximization problems: volume to point, area to point, line to point, and the respective reverse flow directions. Important is the geometric notion that the "volume," the "area," and the "line" represent *infinities* of points. Our theoretical discovery of trees stems from the decision to connect one point (source or sink) with an *infinity* of points (volume, area, line). It is the reality of the continuum that is routinely discarded by modelers who approximate the space as a finite number of discrete points, and then cover the space with "sticks" drawings, which (of course) cover the space incompletely (and, from this, fractal geometry). Recognition of the continuum requires a study of the interstitial spaces between the tree links. The interstices can only be bathed by high-resistivity diffusion (an invisible, disorganized flow), while the tree links serve as conduits for low-resistivity organized flow (visible streams, ducts).

The two modes of flowing with thermodynamic imperfection (i.e., with resistances), the interstices and the links, must be *balanced* so that together they contribute minimum imperfection to the global flow architecture. Choke points must be balanced and distributed. The flow architecture is the graphical expression of the balance between links and their interstices. The deduced architecture (tree, duct shape, spacing, etc.) is the *optimal distribution of imperfection*. Those who model natural trees and then draw them as black lines on white paper (while not optimizing the layout of every black line on its optimally sized and allocated white patch) miss half of the drawing. The white is as important as the black.

Our discovery of tree-shaped flow architectures was based on three approaches. In Bejan (1996), the start was an analytical short cut based on several simplifying assumptions: 90° between stem and tributaries, a construction sequence in which smaller optimized constructs are retained, constant-thickness branches, etc. (e.g., Section 1.2). Months later, we published the same problem (Ledezma et al. 1997) numerically, by abandoning most of the simplifying assumptions (e.g., the construction sequence) used in the first papers. We also

* The great conceptual difference between modeling and theory is spelled out in *Physics Today*, July 2005, p. 20.

did this work in an area-point flow domain with random low-resistivity blocks embedded in a high-resistivity background (Errera and Bejan 1998), by using the language of Darcy flow (permeability, instead of conductivity and resistivity). Along the way, we found better performance and "more natural looking" trees as we progressed in time; i.e., as we endowed the flow structure with more freedom to morph.

And so I end this section with the "click" that I felt as I ended the second paper on constructal trees (for the full version, see p. 813–815 in Bejan 1997a):

The commonality of these phenomena is much too obvious to be overlooked. It was noted in the past and most recently (empirically) in fractal geometry, where it was simulated based on repeated fracturing that had to be assumed and truncated. The origin of such algorithms was left to the explanation that the broken pieces (or building blocks, from the point of view of this paper) are the fruits of a process of self-optimization and self-organization. The present paper places a purely deterministic approach behind the word "self": the search for the easiest path (least resistance) when global constraints (current, flow rate, size) are imposed.

If we limit the discussion to examples of living flow systems (lungs, circulatory systems, nervous systems, trees, roots, leaves), it is quite acceptable to end with the conclusion that such phenomena are common because they are the end result of a long running process of "natural selection". A lot has been written about natural selection and the impact that efficiency has on survival. In fact, to refer to living systems as complex power plants has become routine. The tendency of living systems to become optimized in every building block and to develop optimal associations of such building blocks has not been explained: it has been abandoned to the notion that it is imprinted in the genetic code of the organism.

If this is so, then what genetic code might be responsible for the development of equivalent structures in inanimate systems such as rivers and lightning? What genetic code is responsible for man-made networks (such as the trees in this paper)? Certainly not mine, because although highly educated, neither of my parents knew heat transfer (by the way, classical thermodynamics was not needed in this paper). Indeed, whose genetic code is responsible for the societal trees that connect us, for all the electronic circuits, telephone lines, air lines, assembly lines, alleys, streets highways and elevator shafts in multistory buildings?

There is no difference between the animate and the inanimate when it comes to the opportunity to find a more direct route subject to global constraints, for example, the opportunity of getting from here to there in an easier manner. If living systems can be viewed as engines in competition for better thermodynamic performance, then inanimate systems too can be viewed as living entities (animals!) in competition for survival.

This analogy is purely empirical: we have a very large body of case-by-case observations indicating that flow configurations (animate and inanimate) evolve and persist in time, while others do not. Now we know the particular feature (maximum flow access) that sets each surviving design apart, but we have no theoretical basis on which to expect that the design that persists in time is the one that has this particular feature. This body of empirical evidence forms the basis for a new *law of nature* that can be summarized as . . . [the constructal law, at the start of this section]. This new law brings life and time explicitly into thermodynamics and creates a bridge between physics and biology.

1.2. The Urge to Organize Is an Expression of Selfish Behavior

Why are streets usually arranged in clusters (patterns, grids) that look almost similar from block to block and from city to city? Why are streets and street patterns a mark of civilization? Indeed, *why do streets exist?* Constructal theory provided answers to these questions by addressing the following area-point access maximization problem.

Consider a finite-size geographical area A and a point M situated inside A or on its boundary (Fig. 1.1). Each member of the population living in A must travel between his or her point of residence P(x, y) and point M. The latter serves as common destination for all the people who live in A. The density of this traveling population—i.e., the rate at which people must travel to M—is fixed and described by $\dot{n}''$ (people/m^2 s). This also means that the rate at which people are streaming into M is constrained, $\dot{n} = \dot{n}''A$. Determine the optimal bouquet of paths that link the points P of area A with the common destination M such that the time of travel required by the entire population is the shortest.

The problem is how to connect a finite area (A) to a single point (M). Area A contains an infinite number of points, and every one of these points must be taken into account when optimizing the access from A to M and back. Time has shown that this problem was a lot tougher than the empirical game of connecting "many points": i.e., a finite number of points distributed over an area. The many-points problem can be solved on the computer using brute-force methods (random walk or Monte Carlo—more points on better computers), which are not theory.

The area A could be a flat piece of farmland populated uniformly, with M as its central market or harbor. The oldest solution to this problem was to unite with a straight line each point P and the common destination M. The straight-line solution was the preferred pattern as long as humans had only one mode

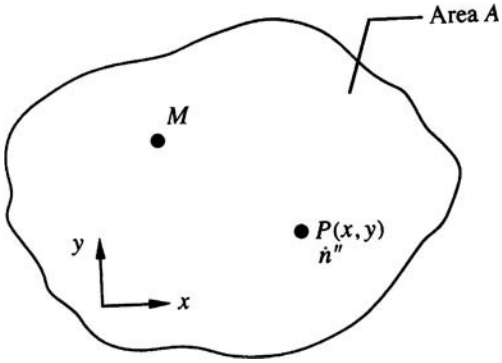

FIGURE 1.1. Finite-size area (A) covered by a uniformly distributed population ($\dot{n}''$) traveling to a common destination (M) (Bejan 1996)

of locomotion: walking, with the average speed V_0. The farmer and the hunter would walk straight to the point (farm, village, river) where the market was located.

The radial pattern disappeared naturally in areas where settlements were becoming too dense to permit straight-line access to everyone. Why the radial pattern disappeared "naturally" is the area-point access problem. Another important development was the horse-driven carriage: with it, people had two modes of locomotion, walking (V_0) and riding in a carriage with an average velocity V_1 that was significantly greater than V_0. It is as if the area A became a composite material with two conductivities, V_0 and V_1. Clearly, it would be faster for every inhabitant (P, in Fig. 1.1) to travel in straight lines to M with the speed V_1. This would be impossible, because the area A would end up being covered by beaten tracks, leaving no space for the inhabitants and their land properties.

The modern problem, then, is one of bringing the street near a small but finite-size group of inhabitants; this group would first have to walk to reach the street. The problem is one of allocating a finite length of street to each finite patch of area A_1, where $A_1 << A$. The problem is also one of connecting these street lengths in an optimal way such that the time of travel of the population is minimum.

The first analytical approach to this problem was "atomistic," from the smaller subsystem (detail) of area A to the larger subsystem, and ultimately to area A itself. The area subsystem to which a street length may be allocated cannot be smaller than the size fixed by the living conditions (e.g., property) of the people who will be using the street. This smallest area scale is labeled A_1 in Fig. 1.2. For simplicity we assume that the A_1 element is rectangular. Although A_1 is fixed, its shape or aspect ratio H_1/L_1 is not. Indeed, the first objective is to anticipate optimal form: the area shape that maximizes the access of the A_1 population to the street segment allocated to A_1.

Symmetry suggests that the best position for the street segment is along the longer of the axes of symmetry of A_1. This choice has been made in Fig. 1.2, where $L_1 > H_1$ and the street has the length L_1 and width D_1. The traveling population density $\dot{n}''$ is distributed uniformly on A_1. To get out of A_1, each person must travel from a point of residence P(x, y) to the (0,0) end of the street. The person can travel at two speeds: (1) a low speed V_0 when off the street and (2) a higher speed A_1 when on the street.

We assume that the rectangle $H_1 \times L_1$ is sufficiently slender ($L_1 > H_1$) so that the V_0 travel is approximated well by a trajectory aligned with the y axis. The time of travel between P(x, y) and (0,0) is $(x/V_1) + (y/V_0)$. The average travel time of the A_1 population is given by

$$\bar{t}_1 = \frac{1}{H_1 L_1} \int_{-H_1/2}^{H_1/2} \int_0^{L_1} \left(\frac{x}{V_1} + \frac{y}{V_0} \right) dx\, dy \tag{1.1}$$

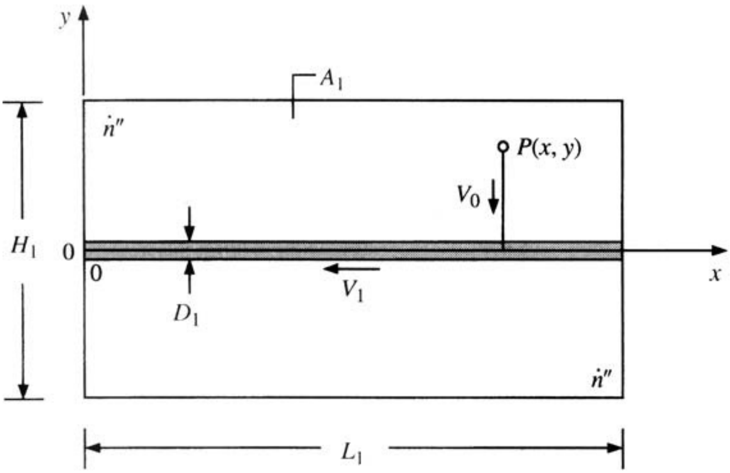

FIGURE 1.2. Smallest (innermost) elemental area, A_1, and the street segment allocated to it (Bejan 1996)

which yields

$$\bar{t}_1 = \frac{L_1}{2V_1} + \frac{H_1}{4V_0} \tag{1.2}$$

The elemental area is fixed ($A_1 = H_1 L_1$, constant); therefore, $\bar{t}_1$ can be expressed as a function of H_1, which represents the shape of A_1:

$$\bar{t}_1 = \frac{A_1}{2V_1H_1} + \frac{H_1}{4V_0} \tag{1.3}$$

The average travel time has a sharp minimum with respect to H_1. Solving $\partial\bar{t}_1/\partial H_1 = 0$, we obtain

$$H_{1,\,opt} = \left(2\,\frac{V_0}{V_1}\,A_1\right)^{1/2} \tag{1.4}$$

and subsequently,

$$L_{1,\,opt} = \left(\frac{V_1A_1}{2V_0}\right)^{1/2} \tag{1.5}$$

$$\left(\frac{H_1}{L_1}\right)_{opt} = \frac{2V_0}{V_1} < 1 \tag{1.6}$$

Equation (1.6) shows the optimal slenderness of the smallest area element A_1. This result validates the initial assumption that $H_1/L_1 < 1$; indeed, the optimal smallest rectangular area should be slender when the street velocity is sensibly greater than the lowest (walking) velocity. The rectangular area A_1 must become more slender as V_1 increases relative to V_0—i.e., as time passes and technology advances. This trend is confirmed by a comparison between the streets built in antiquity and those that are being built today. In antiquity the first streets were short, typically with two or three houses on one side. In the housing developments that are being built today, the first streets are sensibly longer, with 10 or more houses on one side. This contrast is illustrated by modern Rome (Fig. 1.3). In the center, which is the ancient city, the streets are considerably shorter than in the more recently built, peripheral areas (e.g., the upper corners in Fig. 1.3).

Important is the observation that exactly the same optimum [Eqs. (1.4–1.6)] is found by minimizing the longest travel time (t_1) instead of minimizing the area-averaged time of Eq. (1.1). The longest time is required by those who travel from one of the distant corners ($x = L_1$, $y = \pm H_1/2$) to the origin (0,0) and is given by

$$t_1 = \frac{L_1}{V_1} + \frac{H_1}{2V_0} \tag{1.7}$$

Equations (1.7) and (1.2) show that the geometric minimizations of t_1 and $\bar{t}_1$ are equivalent. It is both interesting and important that the optimization of the shape of the A_1 element is of interest to every inhabitant: *What is good for the*

FIGURE 1.3. Plan of modern Rome, showing that in the ancient city (the center) the street length scales are considerably shorter than in the newer outskirts (Bejan 1997c)

most disadvantaged person is good for every member of the community. This conclusion has profound implications in the spatial organization of all living groups, from bacterial colonies all the way to our own societies. *The urge to organize is an expression of selfish behavior*.

The time obtained by minimizing t_1 or by substituting Eqs. (1.4) and (1.5) into Eq. (1.7) is

$$t_{1,\min} = \left(\frac{2A_1}{V_0 V_1}\right)^{1/2} \tag{1.8}$$

At this minimum, the two terms that make up t_1 in Eq. (1.7) are equal. This equipartition of time principle means that the total travel time is minimum when it is divided equally between traveling along the street and traveling perpendicularly to the street. We return to this feature in Section 1.4.

In Fig. 1.2 we see the smallest loop of the traffic network that will eventually cover the given area A. The next question is how to connect the D_1 streets such that each innermost loop has access to the common destination M. One answer—the simplest, albeit approximate—is obtained by *repeating* the preceding geometric optimization several times, each time for a larger area element, until the largest scale (A) is reached. This construction is detailed in Bejan (1996, 1997c), and is explained by considering the rectangular area $A_1 = H_2L_2$ shown in Fig. 1.4. This area consists of a certain number of the smallest patches A_1. The purpose of this assembly of A_1 elements is to connect the D_1 streets so that the traveling population ($\dot{n}''A_2$) can leave A_1 in the quickest manner. We invoke symmetry as the reason for placing the new (second) street along the long axis of the A_2 rectangle. In Fig. 1.4, the stream of travelers ($\dot{n}''A_2$) leaves A_2 through the left end of the D_2 street. There exists an optimal shape H_2L_2, and a minimal global travel time.

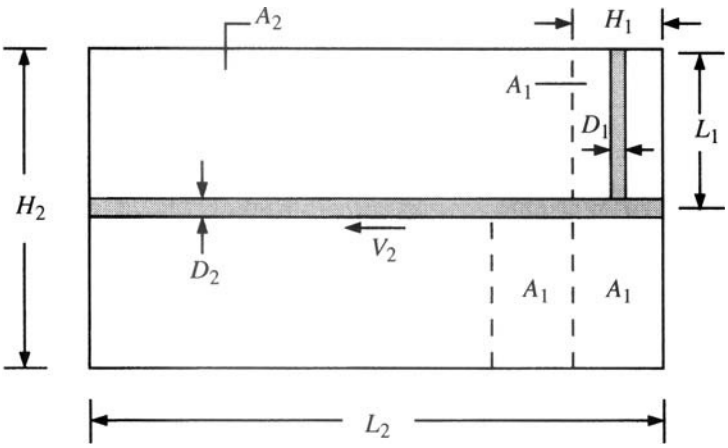

FIGURE 1.4. Area construct A_2 as an assembly of connected innermost elements A_1 (Bejan 1996)

The atomistic construction started in Figs. 1.2 and 1.4 can be continued toward larger assemblies of areas. This is not the best way to allocate streets to areas mathematically, but it is the most transparent. Its value is that it shows the emergence of a tree network (the streets) from principle (the maximization of access), not by copying from nature. In constructal theory, flow architectures such as trees are *discovered*. They are now known, observed, modeled, or copied from nature.

This sequence is shown in Fig. 1.5 (top) only for illustration, because it is unlikely to be repeated beyond the third-generation street. The reason is that as the community and the area inhabited by it grow, other common destinations (e.g., church, hospital, bank, school, train station) emerge on A in addition to the original M point (Fig. 1.1). Some of the streets that were meant to provide access to only one end of the area element must be extended all the way across the area to provide access to both ends of the street. As the destinations multiply and shift around the city, the dead ends of the streets of the first few generations disappear, and what replaces the growth pattern is a *grid* with access to both ends of each street. The multiple scales of this grid, and the self-similar structure of certain areas (neighborhoods) of the grid, however, are the fingerprints of the deterministic organization principle (the constructal law).

The area-point access problem formulated in this section was stated in two dimensions (Fig. 1.1). The corresponding problem in three dimensions is this: minimize the time of travel from all the points P of a volume V to one common destination point M, subject to the constraint that the traveling population rate is fixed. One application is the sizing and shaping of the floor plan in a multistory building, along with the selection and placement of the optimal number of elevator shafts and staircases.

The same organization theory can be extended generally to areas that are populated unevenly, or specifically to highways, railroads, telecommunications, and air routes (e.g., the organization of such connections into hubs, or centrals). A clear application of these concepts is in operations research and manufacturing, where the invention of the first auto assembly line is analogous to the appearance of the first street (Carone 2003; Carone et al. 2003; Hernandez 2001; Hernandez et al. 2003).

The atomistic construction sequence presented until now is just an approximate and simple way to illustrate how a tree of organized (channeled) flow emerges on a background covered by individual (disorganized) movement. The "exact" way to generate the tree architecture from the *same* principle is to endow the flow architecture with maximum freedom to morph (Bejan and Lorente 2004, 2005) and to use numerical simulations to morph the flow structure through all its eligible configurations. This more exact work is illustrated by relaxing the assumption (made in Figs. 1.2 and 1.4) that the paths intersect at right angles.

Assume that in Fig. 1.2 the angle between the V_0 and V_1 paths may vary. This general situation is shown in Fig. 1.6, which is set for calculating the maximum

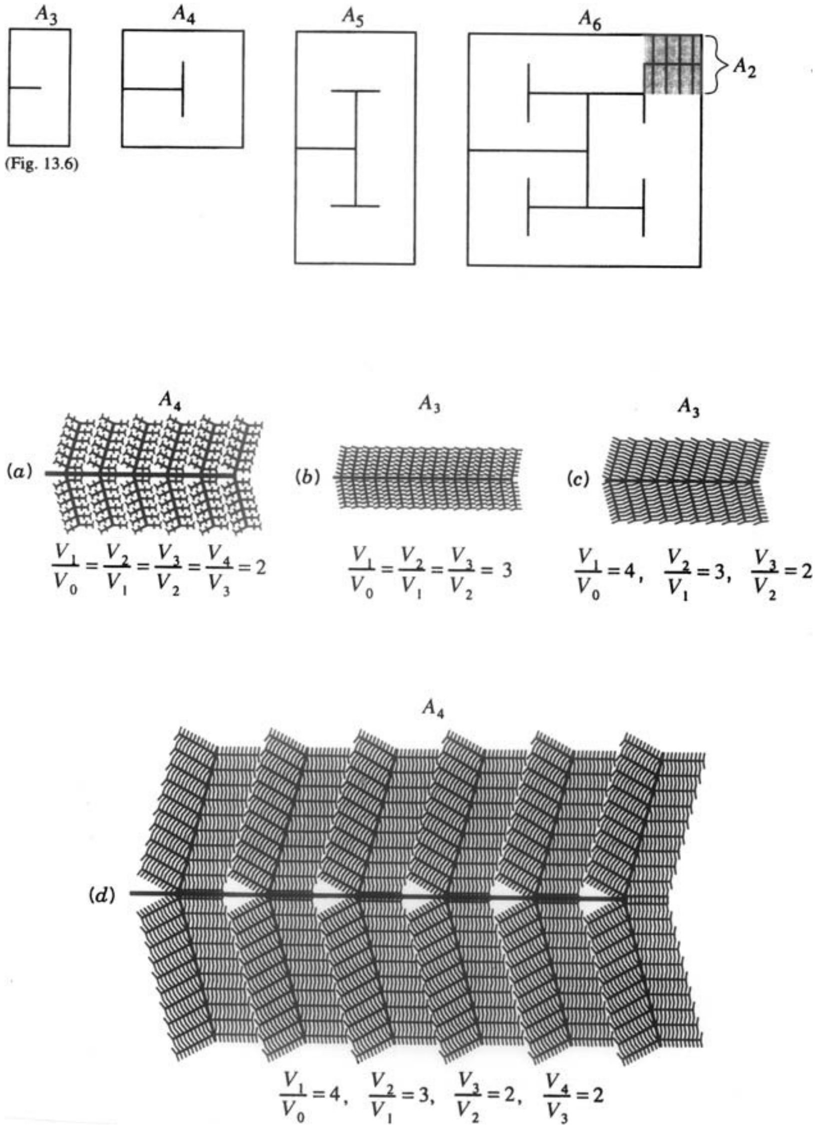

FIGURE 1.5. Top: higher-order constructs in the sequence of Figs. 1.2 and 1.4 (Bejan 1996). Bottom: urban growth patterns in which each construct was optimized for overall shape and angle of street confluence (Ledezma and Bejan 1998)

travel time between the distant corner (P) and the common destination (M). In place of Eq. (1.7), we obtain

$$t_1 = \frac{L_1}{V_1} + \frac{H_1}{2V_0}\left(\frac{1}{\cos\beta} - \frac{V_0}{V_1}\frac{\sin\beta}{\cos\beta}\right) \tag{1.9}$$

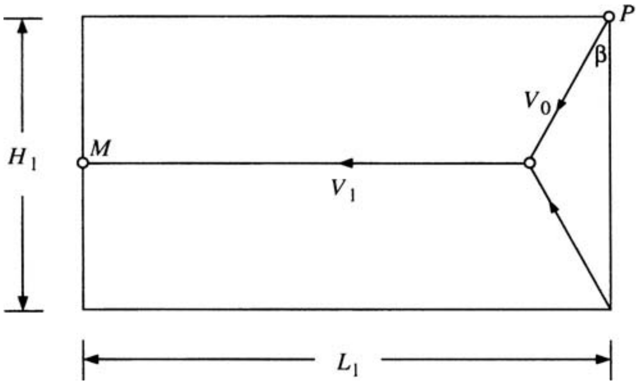

FIGURE 1.6. Smallest area (A_1) and the variable angle between the V_0 and V_1 paths

where β is the angle between the V_0 path and the H_1 side. Now the minimization of t_1 has two degrees of freedom: the geometric aspect ratio H_1/L_1 and the angle β. The optimal angle for minimum t_1

$$\beta_{opt} = \sin^{-1}\frac{V_0}{V_1} \tag{1.10}$$

confirms the statement made above Eq. (1.1) that V_0 should be perpendicular to V_1 (i.e., that $\beta = 0$) when $V_0 << V_1$. The minimization with respect to H_1/L_1 subject to $A_1 = H_1/L_1$ is the same as earlier in this section. The twice-minimized travel time is

$$t_{1,\,min} = \left(\frac{2A_1}{V_0V_1}\cos\beta_{opt}\right)^{1/2} \tag{1.11}$$

The lower part of Fig. 1.5 shows four examples of optimal urban growth, in which each area construct ($A_1, A_2, \ldots$) has been optimized in two ways: overall shape and angle between each new street and its tributaries. The assumed changes in velocity are listed under each drawing. Comparing examples (a) and (d) we see that when the velocity increase factor V_i/V_{i-1} is large the street pattern spreads fast (in few steps) over the given area, and each area assembly is slender. In the opposite limit, the spreading rate is lower, the assembly steps are more numerous, and each area assembly is less slender. These trends appear together in example (d), where the velocity increase factor decreases as the construction grows.

Comparing Eq. (1.11) with Eq. (1.8), we note that the second degree of freedom (the optimized angle β) plays only a minor role as soon as V_1 is greater than V_0. In other words, the change from V_0 to V_1 does not have to be dramatic for the $\beta = 0$ design (Fig. 1.2) to perform nearly as well as the optimal design. We reach the important conclusion that small internal variations in the organization pattern have almost no effect on the global performance of the organized system ($t_{1,min}$, in this case).

The practical aspect of this observation is that a certain degree of variability (imperfection) is to be expected in the patterns and emerge naturally. These patterns are not identical, nor are they perfectly similar; this accounts for the historic difficulty of attaching a theory to naturally organized systems. Natural patterns are quasi-similar, but only in the same sense in which no two human faces are identical. Their performance, however, is practically the same as that of the best pattern. We call these top performers *equilibrium flow structures* (Bejan and Lorente 2004). The contribution of constructal theory is that the performance and the main geometric features (mechanism, structure) of the organized system can be predicted in purely deterministic fashion.

1.3. The Distribution of Human Settlements*

Every sector of society is a conglomerate of mating flows that morph in time in order to flow more easily: people, goods, money, information, etc. The view that society is a flow system with intertwined morphing (improving) architectures was part of the original disclosure of constructal theory (Bejan 1996, 1997c). This deterministic physics principle is in sharp contrast with the empirical (descriptive, modeling) approaches that have been tried to explain social organization. Society is viewed like the photograph of a turbulent flow. Even though the existence of structure is obvious, the image is so complicated, and so much the result of individual behavior, that description is the norm, not prediction For a review, see Bretagnolle et al. (2000), who argue in favor of introducing a spatial dimension (geography) in modeling, toward the development of an evolutionary theory of settlement systems. Such a theory would provide insights for better policy in the future, and will predict the future evolution of towns, cities, and their heterogeneous distribution on land.

Society may be complicated, but pattern is not. Indeed, pattern is "pattern" because it is not complicated. If it were not simple enough for us to grasp, it would be noise, chaos, turbulence, and randomness. Strikingly clear images such as Fig. 1.7 remain unexplained: the size of a city in Europe is inversely proportional to its rank (Bretagnolle et al. 2000; Bairoch et al. 1988; Moriconi-Ebrard 1994), this throughout history. Why?

Figure 1.7 is derivable from the constructal tree-shaped structures deduced for traffic (Section 1.2), which we now review as an introduction. Consider again the minimization of travel time for traffic between an infinity of points (an area) and one point (Figs. 1.1–1.5). The construction of the tree-shaped architecture of the river basin of people starts with the smallest elemental area A_1, which is fixed by the culture of those who live on A_1. For example, A_1 is the farmland surrounding the smallest road (V_1, L_1) that leads to a single marketplace, M_1. The slow movement covers A_1 and touches every point: the slow movement attaches every single inhabitant to the traffic architecture.

* This section is based on Bejan et al. (2006).

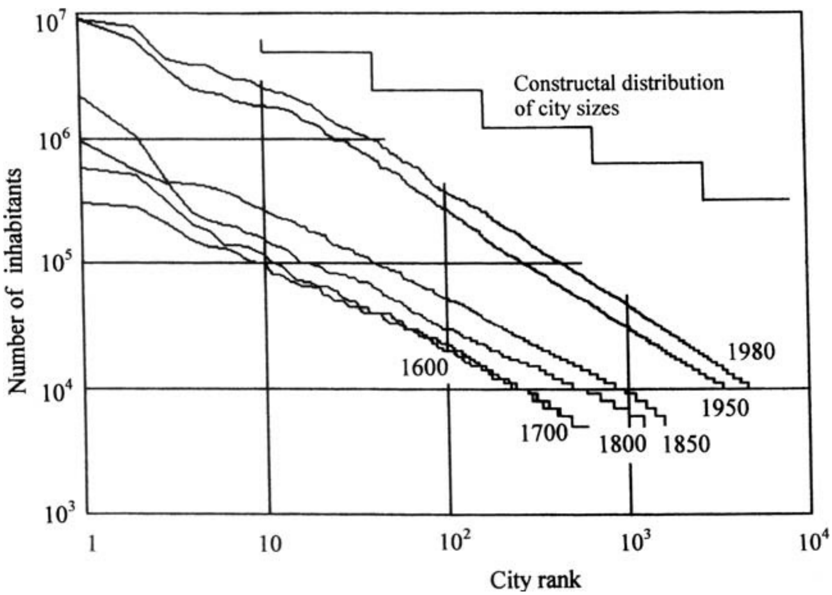

FIGURE 1.7. City size (population) versus city rank, in 1600–1980 Europe (Bejan et al. 2006)

The existence of two modes of movement implies a certain level of civilization: person living not alone (i.e., person with horse) and person with vehicle. Civilization is also the name for the coexistence of farmland (A_1) with markets (M_1). Those who live on A_1 exchange farm products with those who manufacture products and deliver services in compact places such as M_1. It is this balanced counterflow between A_1 and M_1 that justifies this key idea:

The number of those who live on A_1 must be proportional to the number N_1 of inhabitants living at M_1, and both numbers must be proportional to A_1. Both groups are sustained by the agriculture and the "environment" that A_1 provides; therefore, $N_1 = cA_1$, where c is the average number of inhabitants per unit area.

The "culture" factor c accounts for the age and history of the civilization (e.g., technology, commerce, neighbors, natural disasters, plagues, war, peace).

The next larger area that is civilized (A_2) is covered by an assembly of n_1 optimal A_1 rectangles ($A_2 = n_1 A_1$). A central road (speed V_2, such that $V_2 > V_1$) collects or distributes the traffic associated with the elements. This first construct (A_2) can be optimized to provide minimal travel time between A_2 and the new boundary point M_2.

The counterflow of goods between n_1 small markets (M_1) and the largest market (M_2) requires a proportionality between the number (N_2) of inhabitants at M_2 and the total number of inhabitants at the M_1 points. This means that $N_2 = cA_2$. We see here two directions in which hierarchy develops: areas coalesce,