



**HAL**  
open science

# Statistical Topology of Radial Networks: A Case Study of Tree Leaves

Rak-Kyeong Seong, Carolyn Salafia, Dimitri Vvedensky

► **To cite this version:**

Rak-Kyeong Seong, Carolyn Salafia, Dimitri Vvedensky. Statistical Topology of Radial Networks: A Case Study of Tree Leaves. Philosophical Magazine, 2011, pp.1. 10.1080/14786435.2011.614965 . hal-00744823

**HAL Id: hal-00744823**

**<https://hal.science/hal-00744823>**

Submitted on 24 Oct 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**Statistical Topology of Radial Networks: A Case Study of Tree Leaves**

Journal:	<i>Philosophical Magazine &amp; Philosophical Magazine Letters</i>
Manuscript ID:	TPHM-11-May-0204.R1
Journal Selection:	Philosophical Magazine
Date Submitted by the Author:	08-Aug-2011
Complete List of Authors:	Seong, Rak-Kyeong; Imperial College London, Physics Salafia, Carolyn; Placental Analytics LLC Vvedensky, Dimitri; Imperial College London, Physics
Keywords:	topological analysis, statistical physics, statistical mechanics
Keywords (user supplied):	radial networks, triangulation, biological networks

SCHOLARONE™  
Manuscripts

*Philosophical Magazine*

Vol. 00, No. 00, 00 Month 2010, 1–14

## RESEARCH ARTICLE

## Statistical Topology of Radial Networks: A Case Study of Tree Leaves

Rak-Kyeong Seong,<sup>\*a</sup> Carolyn M. Salafia,<sup>†b</sup> and Dimitri D. Vvedensky<sup>‡a</sup><sup>a</sup>*The Blackett Laboratory, Imperial College London, London SW7 2AZ, United Kingdom*<sup>b</sup>*Placental Analytics LLC, 93 Colonial Avenue, Larchmont, New York 10538, USA**(Received 00 Month 200x; final version received 00 Month 200x)*

We describe a statistical method for characterizing the topological properties of radial networks based on the distribution of their ‘energy’ states, which are determined from the structural triangulation of the network. The partition function obtained from these energy states is used to calculate thermodynamic functions that embody the statistical properties of the network. The entropy, in particular, is a measure of the distribution of triangulated areas in the network, with a larger entropy corresponding to a higher symmetry in the branching structure. By varying the distribution parameter, which corresponds to an inverse ‘temperature’ in the statistical thermodynamic interpretation, we are able to vary the weight of the different generations of the network. This analysis identifies similar networks at their maturation state – the state when the system stops growing – as well as enabling the development of a network to be investigated. The latter feature is especially important for biological systems, where the details of the expansion of the network are not typically available. We illustrate our methodology with a model for the optimal transport of nutrients within tree leaves and show that statistical thermodynamic functions are capable of discriminating between various types of such radial networks. We conclude with a discussion about applications to the vasculature of the human placenta, which is our main motivation in developing this approach.

**Keywords:** radial networks, triangulation, topological invariants, partition function, statistical thermodynamics

### 1. Introduction

A network consists of edges connecting objects called vertices or nodes. A modern example of a network is the internet, where the vertices are routers and computers, and the edges are physical or wireless links. Networks describe many important functions and processes, among the best-known of which are the distribution or flow, for example, of information on the internet, power on a grid, people and freight on road, rail, and air networks, and nutrients within biological networks. Interest in networks has expanded rapidly in recent years [1–3]. In particular, there has been substantial interest in modelling venation patterns [4, 5], optimizing supply networks [6, 7], and determining final system shapes and sizes based on a simulated growth mechanisms [8–10].

The study of networks dates back to the celebrated Königsberg bridge problem, which asks if the seven bridges in the city of Königsberg can be traversed in a single journey without doubling back and ending in the same place as it began. By abstracting this problem into one of edges and nodes, Euler [11] answered this question in the negative and thereby laid the foundation for what has become

\* Email: rak-kyeong.seong05@imperial.ac.uk

† Email: carolyn.salafia@gmail.com

‡ Email: d.vvedensky@imperial.ac.uk

1 known as graph theory. Further developments were not immediate. Some 100 years  
2 after Euler's paper, the theory of trees was initiated by Kirchhoff [12] for appli-  
3 cations to electrical networks and by Cayley [13] for enumerating the isomers of  
4 saturated hydrocarbons. The modern incarnation of graph theory is due in large  
5 part to König, who collated his own work and that of other mathematicians in the  
6 first book on the subject [14].

7 An expansion of the foregoing work on *regular* graphs was the introduction of  
8 *random* graphs by Erdős and Rényi [15, 16] and by Gilbert [17] in the late 1950s.  
9 Random graphs have no particular distributions of vertices and edges and their  
10 organization principles are not readily identified. The Erdős-Rényi construction  
11 starts with  $N$  nodes that are connected pairwise with probability  $p$ , creating a  
12 graph with approximately  $\frac{1}{2}pN(N-1)$  randomly distributed edges. This model  
13 had been the starting point for discussions about networks for many years, but the  
14 growing interest in complex systems prompted a re-evaluation of this paradigm  
15 [1, 18]. Many of the features of real networks that have attracted interest concern  
16 the ways in which such networks *differ* from random graphs. Complex systems are  
17 expected to have organizing principles based upon the competition between ex-  
18 tending the network and the 'cost' of this extension. At some level, this should be  
19 encoded in the topology and geometry of the network. Indeed, the geometry of real  
20 networks is extremely varied, with the different structures reflecting the different  
21 functions that networks perform. Thus, the development of theoretical principles  
22 and measures to capture and quantify the underlying organizing principles of net-  
23 works is of paramount importance.

24 There are several levels of description of physical and biological networks. In the  
25 'coarse-grained' approach that is the domain of statistical mechanics, the statistical  
26 properties of the 'skeletal' network are of primary interest. These include global  
27 properties, such as the fractal dimension and other scaling properties [19, 20],  
28 to more local quantities such as the degree distribution, where the degree of a  
29 node is the number of edges emanating from that node. At a more microscopic  
30 level, the main interest focuses on the function of the network, such as the flow  
31 of nutrients, to characterize, in the case of vascular networks [21], the flow rate  
32 and the length and diameter of vessel branches. **The problem of the maximal flow  
33 through networks given the capacity of the edges was solved long ago by Ford and  
34 Fulkerson [22], though the formulation of this problem goes back much further  
35 [23, 24]. Our interest is not in maximal flow *per se*, but in how flow affects the  
36 expansion of a network based on some criteria, such as optimization based on a  
37 cost function.** Here, the coarse-grained and macroscopic approaches come together  
38 because the spreading of the network is driven by the fulfilment of a function, such  
39 as the delivery of nutrients, but limited by the overall cost incurred by extending  
40 the network.

41 The rules and parameters governing the supply and demand of a system are ex-  
42 pected to be invariant over time scales beyond its lifespan. Hence, the same rules  
43 and parameters govern the creation and extinction states of a system. What hap-  
44 pens in between is affected by external factors. An important question is whether  
45 systems can be grouped according to their shape and venation at a given state in  
46 their development. Going a step further, one may ask whether interpretations of  
47 life history can be made by analyzing and comparing a state in the lifespan of a  
48 system to a comparable state of another system. We expect that the growth history  
49 of a system is encoded in the distribution of energy in its venation, i.e. the work  
50 done in building up the supply system. Due to external factors and the existing  
51 structure of the system, some parts of a network may have required more energy  
52 to build up than others.

In this paper, we introduce a method for studying the **statistical topology** of networks based on the distribution of their ‘energy’ states, as represented by the structural triangulation of the network. The partition function obtained from these states is used to calculate thermodynamic functions in the usual manner as derivatives of the Helmholtz free energy. The ‘temperature’-dependence of these functions provides a way of assigning weights to different generations of the network. The statistical thermodynamic functions identify similar networks at their maturation state – the state when the system stops growing – as well as enabling the development of a network to be investigated. The latter feature is especially important for biological networks, where the details of the expansion of the network are not always available. We illustrate our methodology by analyzing a model for **optimal transport in tree leaves** to demonstrate that our approach allows networks to be examined in a systematic manner and, most significantly, **allows similar networks to be identified and grouped together**.

The organization of this paper is as follows. The leaf network and the optimal transport model for leaf growth [10] are formulated in Secs. 2 and 3, respectively. This model accounts for the variations of leaf sizes and shapes seen in nature and provides a useful test case for our approach because the characteristics of these networks can be varied systematically. The triangulation of radial networks and their statistical analysis is the subject of Sec. 3. Applications to tree leaves are discussed in Sec. 4, where the model in Sec. 2 is used to produce leaves of different sizes and shapes. Our analysis shows that the set of leaves produced can be partitioned into distinct classes. We summarize our results in Sec. 5 and discuss the application to biological networks and, in particular, to the human placenta, which is the motivation for developing this approach.

## 2. The Supply Network of Tree Leaves

### 2.1. Coordinate System of the Leaf

A leaf consists of living cells that **may be regarded as lying** on a plane. Plant cells differ from other cells in that they can be approximated by squares or rectangles. Accordingly, we place cells on a lattice with spacing  $h$  such that a particular cell in the leaf is labelled by coordinates  $(m, n)$ , where  $m$  and  $n$  are integers. The actual location of the biological cell in relation to the entire leaf is given by the cell coordinate  $\mathbf{x}_{m,n} = (mh, nh)$ . The lattice  $\Lambda = \{(m, n)\}$  is the set of all possible lattice labels, and the grid  $\Gamma_h = \{(mh, nh)\}$  is the set of all possible cell positions. The leaf itself is a finite selection of grid positions,  $\Omega_h = \{\mathbf{x}_{m_1, n_1}, \dots, \mathbf{x}_{m_N, n_N}\}$ , for a total of  $N$  living cells. We always start the growth of the leaf from an infinitesimally small seed identified as the origin  $\mathcal{O} = (0, 0)$ .

### 2.2. The Supply Network

The leaf cells require water and minerals for the production of nutrients from photosynthesis. The nutrients need to be transported back to the root and stem of the plant, where they are stored. The supply network  $G$  consists of a subset of lattice vertices  $\{(m, n)\}$  in  $\Lambda$ , edge vectors  $\mathbf{e} = ((m_p, n_p), (m_c, n_c))$  that point from a **parent vertex**  $u = (m_p, n_p)$  to a **child vertex**  $v = (m_c, n_c)$ , and positive integer **weights**  $w$  that are assigned to the edges. The weight of an edge counts the child vertices  $v$  that follow after the edge. The supply network  $G$  of the leaf lattice is defined as the vertex subset  $V(G) = \{(m, n)\}$ , edges  $E(G)$ , and weight function  $w: G = \{V(G), E(G), w\}$ .

For leaf cells, the parent vertex is the first object supplied with nutrients prior to the child vertices. We interpret edges  $E(G)$  as tubes connecting two leaf cells, and weights  $w$  on edges as a measure of the amount of supply flowing through that particular element of the transport system. This enables the lattice supply network  $G$  to be mapped directly to the **venation** of the leaf  $\Lambda_G$ . For the following discussion, we set  $h = 1$ , which allows us to use the terms ‘lattice’ and ‘leaf’ interchangeably.

A key property of the supply network is that each vertex has only one parent vertex. This is in line with the biological observation that cell division is a constructive process. Thus, introducing the more compact edge notation  $e_v = (p(v), v)$  with  $p(v)$  being the parent of  $v$ , there cannot be two edges  $e_1 = (p(v_1), v_1)$  and  $e_2 = (p(v_2), v_2)$  such that  $v_1 = v_2$ , i.e. two edges cannot end at the same vertex. This has the natural consequence that a single supply network can never be closed through loops. The origin  $\mathcal{O}$  has, by definition, no parent, so we **have**  $e_{\mathcal{O}} = 0$ . Let us define the set of *ancestors*  $V_A(u) = \{p^k(u)\}$  with positive integer  $k$ , and *descendants*  $V_D(u) = \{p^k(u)\}$ , with negative integer  $k$ , of a given vertex  $v$  in  $V(G)$ , such that the origin  $\mathcal{O}$  has no ancestors and **its** descendants form the entire vertex set of the leaf:  $V_D(\mathcal{O}) \equiv V(G)$ . We can now give an explicit recursive form of the weight function  $w(e)$ , of a given edge  $e = (u, v)$ ,

$$w(e) = w((u, v)) = w((v, p^{-1}(v))) + 1, \quad (1)$$

which conforms to our definition of the weight function as being the size of  $V_D(v)$  for a given edge  $e = (u, v)$ .

### 3. Optimization and the Cost Function

#### 3.1. Measuring Cost

Supplying a leaf with water and minerals has an associated cost. Minimizing this cost determines the shape of the leaf. A primary source of cost arising in the supply network is due to the flow direction of nutrients. Transport cost is minimized when nutrients are moved in the same direction and maximized when moved in opposing directions. For two consecutive edges  $e_1 = (u, v)$  and  $e_2 = (v, p^{-1}(v))$ , we define the instantaneous cost **parametrized by the quantity**  $\beta_i$  as

$$H_{\beta_i}(e_1, e_2) = \begin{cases} |e_1 \cdot e_2|^{-\beta_i}, & \text{if } e_1 \cdot e_2 > 0; \\ g, & \text{otherwise,} \end{cases} \quad (2)$$

where  $e_1 \cdot e_2 = (u - v) \cdot (v - p^{-1}(v))$  **with**  $u$  and  $v$  being 2-vectors, as defined in **Sec. 2.2**. By setting the function  $g \rightarrow \infty$ , we eliminate the possibility of opposite flows in the transport network, i.e.  $e_1 \cdot e_2 \leq 0$ . For all other flows,  $\beta_i$  scales the extent to which the flow direction contributes to the total cost of sustaining the supply network. We note also that, by definition, the instantaneous cost is **expressed in terms of** two *consecutive* edge vectors,  $e_1$  and  $e_2$ , such that correlations between edge directions separated by many generations are limited and contribute minimally to the total cost of the supply network.

In addition to the instantaneous cost, there is also the cost due to the amount of supply flow through a network edge. More matter flowing through a section of a transport system leads to a larger cross-sectional area of that edge, hence requiring more energy to sustain the pressure in the descendant supply network. The ‘thickness’ of a given edge is proportional to the edge’s weight and, accordingly,

the cost of maintaining an edge  $e$  with a given weight  $w(e)$ , called the  $\alpha_f$ -cost, is defined as

$$H_{\alpha_f}(e) = (w(e))^{\alpha_f} \ell(e), \quad (3)$$

where  $\ell(e) = \ell(u, v) = |v - u|$  is the lattice length of the edge. We note that  $H_{\alpha_f}(e)$  is proportional to the volume of fluid present at the edge at a given time.

There is an important characteristic difference between the  $\alpha_f$ - and  $\beta_i$ -costs. Whereas the  $\alpha_f$ -cost of a single edge  $e = (u, v)$  is independent of the costs associated with its ancestors, the *total*  $\beta_i$ -cost depends on the *instantaneous*  $\beta_i$ -costs  $H_{\beta_i}(e)$  associated with these ancestors due to the fact that the fluid arriving at  $e$  had to experience the curvature of the transport system connecting this edge to the origin. Accordingly, the total  $\beta_i$ -cost at a given edge  $e = (u, v)$  can be given in a recursive form as

$$m_{\beta_i}(e) = m_{\beta_i}((u, v)) = m_{\beta_i}((p(u), u)) H_{\beta_i}(e, (p(u), u)). \quad (4)$$

We can consider the number  $D$  of edges in the transport system as the dimension of a vector space. Then, we define the two cost vectors,

$$\begin{aligned} \mathbf{H}_{\alpha_f} &= (H_{\alpha_f}(e_1), H_{\alpha_f}(e_2), \dots, H_{\alpha_f}(e_D)), \\ \mathbf{m}_{\beta_i} &= (m_{\beta_i}(e_1), m_{\beta_i}(e_2), \dots, m_{\beta_i}(e_D)), \end{aligned} \quad (5)$$

with components  $H_n^{(\alpha_f)} = H_{\alpha_f}(e_n)$  and  $m_n^{(\beta_i)} = m_{\beta_i}(e_n)$ , where  $n = 1, \dots, D$ . The total cost of sustaining the entire transport system  $G$  can then be defined as the sum over the total cost of sustaining each edge:

$$F(G) = \sum_{n=1}^D H_n^{(\alpha_f)} m_n^{(\beta_i)}. \quad (6)$$

### 3.2. Growth, Supply and Optimizing Cost

Growth is defined by the addition of an edge such that the cost is always less than or equal to the supply available. For the leaf, supply is defined as the energy available for the pumping process, and energy is supplied mainly by the process of photosynthesis. Accordingly, the supply should be proportional to the area of a leaf cell – the larger the area, the more light is absorbed for photosynthesis. We define the lattice supply function  $S_{\varepsilon_p}$  as

$$S_{\varepsilon_p}(G) = Dk(h)\varepsilon_p, \quad (7)$$

where  $\varepsilon_p$  is the efficiency coefficient of the photosynthesis process, and  $k(h)$  is an area function which for the square lattice is  $k(h) = h^2$ .

Between each growth step  $G_i \mapsto G_j$  we add a new vertex  $(m_b, n_b)$  to  $V(G_i)$ , where  $b = (m_b, n_b)$  is identified as a boundary element of the transport system. For a given growth step  $G_i \mapsto G_j$ , there is the possibility of minimizing the cost  $F(G_j)$  to  $\min(F(G_j)) \leq F(G_j)$  by connecting  $b$  to the rest of the transport network appropriately. The **desirable** situation for continued growth is when the condition

$$\min(F(G_j)) < S_{\varepsilon_p}(G_j) < F(G_j) \quad (8)$$

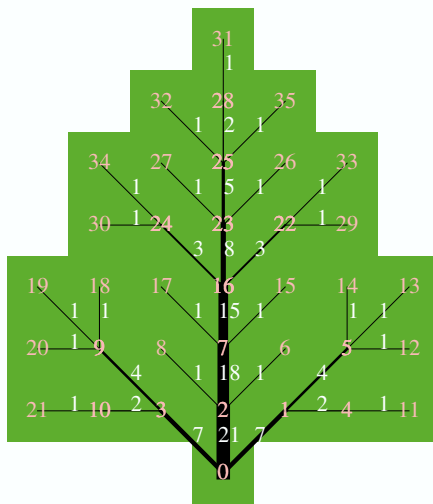


Figure 1. A leaf produced with parameters  $\alpha_f = 0.6$ ,  $\beta_i = 0.5$  and  $\varepsilon_p = 3.0$ . Numbers in white indicate edge weights  $w$  and numbers in pink indicate the historical order of the vertices. The thicknesses of the edges indicate the relative amount of material flow.

is **maintained** at each growth step. The growth history

$$G_1 \mapsto G_2 \mapsto \dots \mapsto G_{N_H-1} \mapsto G_{N_H} \quad (9)$$

has a fixed length  $N_H$  where the history length is determined when the condition

$$S_{\varepsilon_p}(G^{N_H+1}) < \min(F(G_{N_H+1})) \leq F(G_{N_H+1}) \quad (10)$$

is satisfied for the first time. We note that  $N_H = N_H(\alpha_f, \beta_i, \varepsilon_p)$  is a function of the cost and supply parameters, leading to the conclusion that the shape of the final leaf is solely determined by the cost and supply parameters  $\alpha_f, \beta_i$  and  $\varepsilon_p$ . By introducing alterations to the system (both cost and supply) as outlined in the following section, it is possible to vary the final shape of the leaf.

An example of a leaf generated by the model described above is shown in Fig. 1 with the parameters  $\alpha_f = 0.6$ ,  $\beta_i = 0.5$  and  $\varepsilon_p = 3.0$ . Numbers in **white** indicate edge weights  $w$  and numbers in **pink** indicate the historical order of the vertex. Several large-scale features of this leaf are immediately apparent. This leaf is symmetric and finite in extent, as are all leaves that are produced under the model described above. Asymmetric leaves can be obtained by introducing defects into the lattice either before or during growth, e.g. no-growth zones. The finite extent of the leaf is, of course, a natural result of the optimal transport algorithm. An examination of the historical order of the vertices shows that, while the mature leaf is symmetric, the leaf *during* growth is decidedly asymmetric. This is likely an artefact of our model – the important aspect for our analysis is the number of edges emerging from the vertices of the mature leaf.

#### 4. Statistical Thermodynamics and Weighted Topology

##### 4.1. Triangulation

Topological invariants are used to characterize shapes and structures of geometrical objects. A prominent example is the Euler characteristic  $\chi$  [25], a measure of the *topological* structure of a space that is invariant under certain deformations. For



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

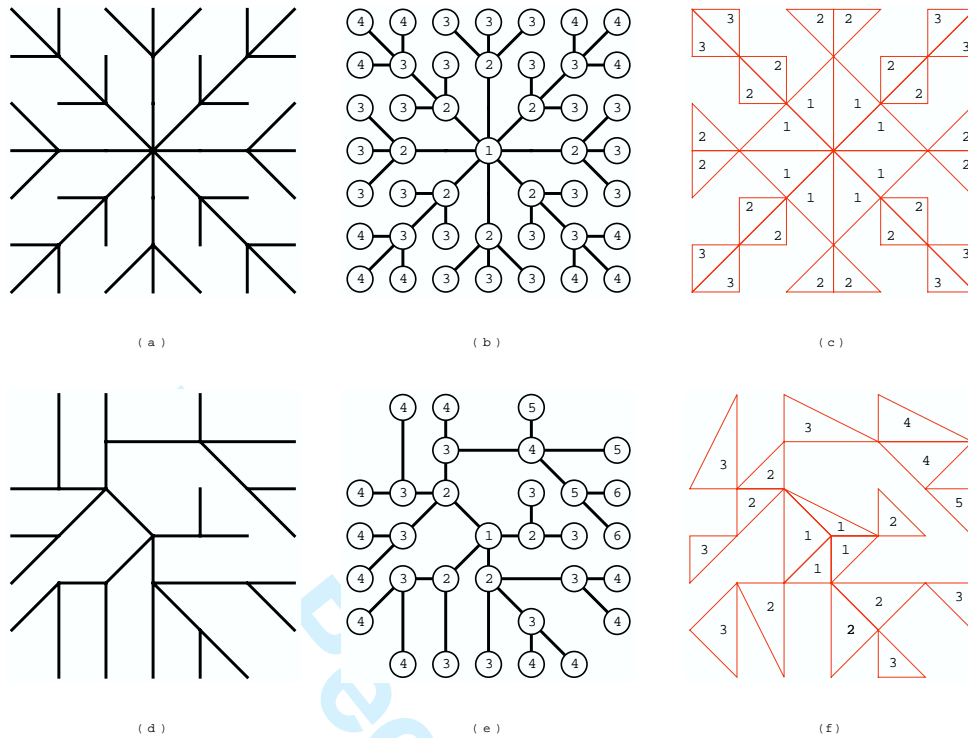


Figure 2. Two radial networks (a,d) shown with the generation number of their vertices indicated with circles centered at each vertex (b,e) and the corresponding triangulations, with the generation of each triangle indicated by the enclosed number (c,f).

any convex polytope, for example,  $\chi = V - E + F = 2$  where  $V$ ,  $E$  and  $F$  are the numbers of vertices, edges and faces, respectively, of the polytope.

We focus now on planar radial networks, such as the tree leaf model described in Secs. 2 and 3. Recall that edges  $e = (v_i, v_j)$  connect vertices  $v_i = (x_i, y_i)$  and  $v_j = (x_j, y_j)$ , with one vertex designated as the origin  $v_0 = (0, 0)$ . The origin is the first generation vertex,  $g(v_0) = 1$ , and an edge  $e = (v_i, v_j)$  always connects an  $n$ th generation vertex,  $g(v_i) = n$ , to an  $(n + 1)$ th generation vertex,  $g(v_j) = n + 1$ . We define a triangular element  $\Delta$  as a set of three vertices,  $\Delta = \{v_i, v_j, v_k\}$ , such that the associated vertex generations are  $g(v_i) = n$  and  $g(v_j) = g(v_k) = n + 1$  and  $p(v_j) = p(v_k) = v_i$ . A triangular element is assigned the generation  $g(\Delta) = g(v_i)$ . It is important to avoid an overlap of area between two elements with a common vertex. Let the vertex  $u$  with  $g(u) = n$  be the parent of the  $c$  vertices  $v_m$  with  $g(v_m) = n + 1$ , where  $m = 1, \dots, c$  with  $c + 1 \equiv 1$  labelling the vertices  $v_m$  in a counterclockwise direction around the parent vertex  $u$ . Such an ordering is unique if no two vertices  $v_m$  lie on a ray from  $u$  (if multiple vertices  $v_m$  lie on a ray from  $u$ , the furthest away from  $u$  is taken). With these conventions,  $\Delta_m = \{u, v_m, v_{m+1}\}$ , given that the counterclockwise angle from  $v_m$  to  $v_{m+1}$  is less than  $180^\circ$ . The resulting fan  $\{\Delta_m\}_u$ , therefore, does not contain overlapping triangular elements. We call the set of all triangular elements in a network the ‘triangulation’  $\mathcal{M}$ . Since  $\mathcal{M}$  consists of fans associated with vertices in the network, it is unique for a given network.

Two radial networks and their triangulations are shown in Fig. 2. Each triangular element  $\Delta_m$  has an associated generation  $g(\Delta_m)$  and an associated area  $A(\Delta_m)$ . While the structure of a triangulation is unchanged by deformations of the network, the areas of the elements allow the original and transformed networks to be distinguished. Motivated by this property, we use the generation and area associated

with the triangular elements as a measure of the energy distribution in a venation. In effect, we have a weighted topological description of the network based on its triangulation.

For increasing generation number  $g$ , the system expends correspondingly more energy by creating a vertex  $v$  and element  $\Delta$  associated with  $g(v)$ . The system must move material from the supply source at the origin to the position at generation  $g$ . A reasonable physical interpretation for  $g$  is, therefore, as an energy level of the system. If the energy unit of the system is  $E_0$ , the energy  $E_g$  is the work required to produce a unit at generation  $g$ :  $E_g = gE_0$ . Building on this point of view, a natural interpretation of the area  $A(\Delta)$  of an element  $\Delta$  is a measure of the material needed to assemble this element at generation  $g(\Delta)$ . In general, if the material needed for an element of the system is  $N$ , then  $N(\Delta) = \rho A(\Delta)$ , where  $\rho$  is a proportionality constant.

Taking the system of tree leaves, which is embedded in an integer lattice with lattice spacing  $h = 1$  as discussed above, the smallest triangular element  $\Delta_{\min}$  has an area  $A(\Delta_{\min}) = \frac{1}{2}$ . In order to deal only with integer  $N$ , we set the proportionality constant to  $\rho = 2$ . This choice is not a requirement of our model, but it allows us to interpret  $N(\Delta)$  as the number of ‘particles’ with energy  $E_{g(\Delta)}$  required to create the triangle element  $\Delta$ . In general,  $\rho$  should be kept constant for analysis within the chosen system.

#### 4.2. Statistical Thermodynamics from Triangulation

We are now able to define the partition function of the triangulation  $\mathcal{M}$  as

$$Z(\mathcal{M}) = \sum_{\Delta \in \mathcal{M}} N(\Delta) t^{g(\Delta)}, \quad (11)$$

where  $t = e^{-\beta E_0}$  and  $\beta$  is a parameter that controls the thermodynamic weight given to successive generations. The range of  $\beta$  is  $0 < \beta < \infty$ , so the range of  $t$  is  $0 < t < 1$ . The factor  $t^{g(\Delta)}$  originates from the fact that particles in the same generation are assigned to the same energy level  $E_{g(\Delta)} = g(\Delta)E_0$ . Hence, the partition function is a polynomial whose order is the number of generations in  $\mathcal{M}$ . The partition functions for the networks in Fig. 2 are

$$\begin{aligned} Z(\mathcal{M}_c) &= 16t + 16t^2 + 8t^3, \\ Z(\mathcal{M}_f) &= 5t + 9t^2 + 9t^3 + 4t^4 + t^5, \end{aligned} \quad (12)$$

where  $\mathcal{M}_c$  and  $\mathcal{M}_f$  are the triangulations in Fig. 2(c) and 2(f), respectively. Note that the triangular elements are distributed among a greater number of generations in the network in Fig. 2(c) than that in Fig. 2(a).

The partition function is the sum over the microstates the system can occupy. These are identified in the network as ‘particles’ that have an associated history. To create a particle in element  $\Delta$ , the network had to occupy a specific state where work  $E_{g(\Delta)}$  was done to create this particle. Accordingly, when we sum over all network system particles in (11), we are indeed summing over all microstates in its history with associated energy  $E_{g(\Delta)}$ . Because energy levels depend on the generation  $g$  of the triangle elements which, in turn, depend on the structure of the venation, the partition function in (11) can be considered not only as a statistical measure of the network, but also as a measure of the network tree structure. Therefore, we expect two systems with similar partition functions to display similar venation structures. Interestingly enough, this does not exclude the possibility of

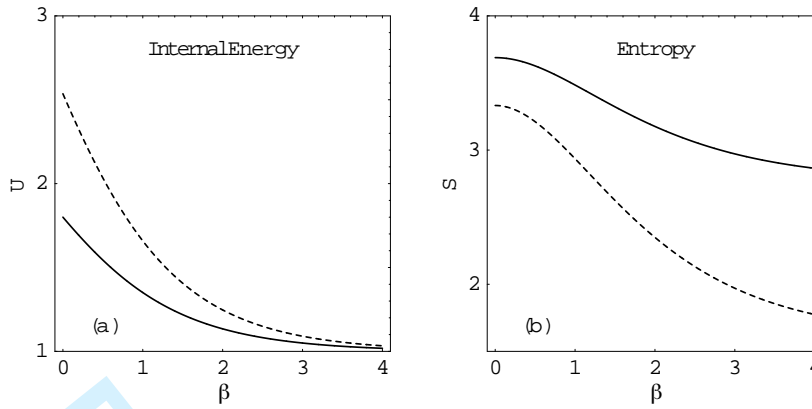


Figure 3. (a) Internal energy and (b) entropy for the networks in Fig. 2(a) (solid curves) and Fig. 2(d) (broken curves), determined from their triangulations.

having systems with ostensibly different partition functions that can show similar behavior over restricted ranges of  $\beta$ .

### 4.3. Statistical Topology

The partition function enables us to calculate thermodynamic variables for the network. The internal energy is

$$U = -\frac{\partial \ln Z}{\partial \beta}. \quad (13)$$

Figure 3(a) shows that  $U(\mathcal{M}_f) > U(\mathcal{M}_c)$  for all  $\beta$ . In the first few generations (large  $\beta$ ), the internal energies of the two networks are similar, but diverge as higher generations are included (decreasing  $\beta$ ). A physical rationale for this behavior is suggested by using the definition of the partition function (11) in the definition of the internal energy to obtain

$$U = E_0 \frac{\sum_{\Delta \in \mathcal{M}} N(\Delta) g(\Delta) t^{g(\Delta)}}{\sum_{\Delta \in \mathcal{M}} N(\Delta) t^{g(\Delta)}} \equiv \langle E_0 g(\Delta) \rangle, \quad (14)$$

which is the thermodynamic average of the work required to produce a unit across all generations. As  $\beta \rightarrow \infty$  ( $t \rightarrow 0$ ), i.e. in the initial stages of development,

$$\lim_{\beta \rightarrow \infty} U = E_0, \quad (15)$$

since only the first generation contributes. For each network, we work in energy units of  $E_0$ , so all internal energy profiles approach 1 at  $\beta \rightarrow \infty$ . The other limit, where  $\beta \rightarrow 0$  ( $t \rightarrow 1$ ) is, according to (14),

$$\lim_{\beta \rightarrow 0} U = E_0 \frac{\sum_{\Delta \in \mathcal{M}} N(\Delta) g(\Delta)}{\sum_{\Delta \in \mathcal{M}} N(\Delta)}, \quad (16)$$

which is the average work required to produce a unit across all generations, with each generation weighed equally. Hence, internal energy profiles of all networks coincide initially ( $\beta \rightarrow \infty$ ), but then differences in later generations are manifested as deviations in these profiles as  $\beta \rightarrow 0$ . Thus, the structure of the asymmetric

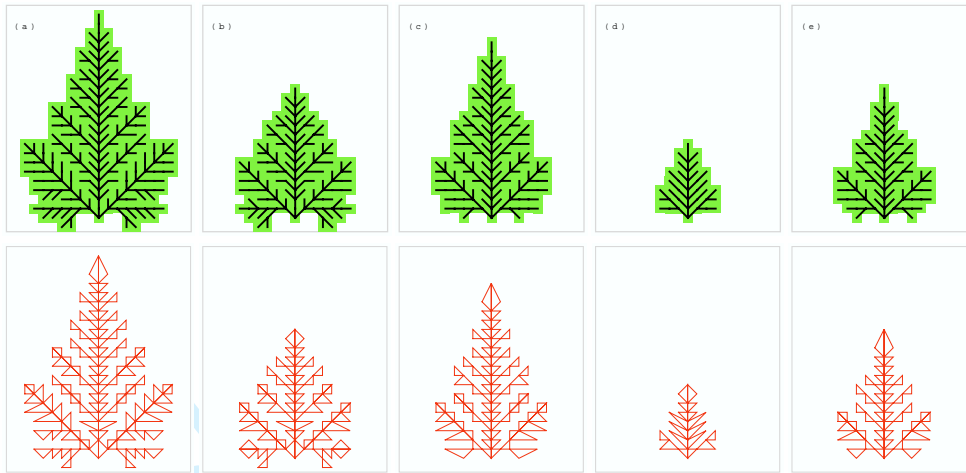


Figure 4. Simulated leaves (top panels) with parameters (a)  $\alpha_f = 0.6, \beta_i = 0.5$ , (b)  $\alpha_f = 0.65, \beta_i = 0.38$ , (c)  $\alpha_f = 0.6, \beta_i = 0.8$ , (d)  $\alpha_f = 0.75, \beta_i = 0.7$ , and (e)  $\alpha_f = 0.65, \beta_i = 0.7$ , with corresponding triangulations (bottom panels). All leaves have  $\varepsilon_p = 4.0$ .

network in Fig. 2(b) displays a higher internal energy through more elements being in higher generations of the system than in the symmetric network in Fig. 2(a).

The entropy is defined as

$$S = k_B(\ln Z + \beta U) \quad (17)$$

in the canonical ensemble, where we set Boltzmann's constant  $k_B = 1$ . As Fig. 3(b) shows,  $S(\mathcal{M}_c) > S(\mathcal{M}_f)$  for all  $\beta$ . To understand this result, we identify entropy as a measure of energy dispersal in the network, rather than as a measure of structural 'disorder'. The energy in the symmetric network is more evenly distributed between the different generations than in the asymmetric network and, hence, has the higher entropy. As  $\beta \rightarrow \infty$  ( $t \rightarrow 0$ ), the entropy approaches the logarithm of the triangulated area associated with the first generation:

$$\lim_{\beta \rightarrow \infty} S = \ln \left[ N(\Delta) \Big|_{g=1} \right]. \quad (18)$$

The other limit is  $\beta \rightarrow 0$  ( $t \rightarrow 1$ ), where the entropy approaches the logarithm of the area of the entire triangulation:

$$\lim_{\beta \rightarrow 0} S = \ln \left[ \sum_{\Delta \in \mathcal{M}} N(\Delta) \right]. \quad (19)$$

Both of these limits can be obtained directly from the partition function; the interesting situation is when the entropy curves of different networks cross. The next section will provide examples of where this occurs.

## 5. Analysis of Tree Leaves

We have simulated the growth of five tree leaves using the model of leaf growth described in the preceding section. These are shown in Fig. 4, together with their parameters in the cost and supply function. There are discernible differences in size, shape, and branching patterns. In fact, the modelling of tree leaf venation

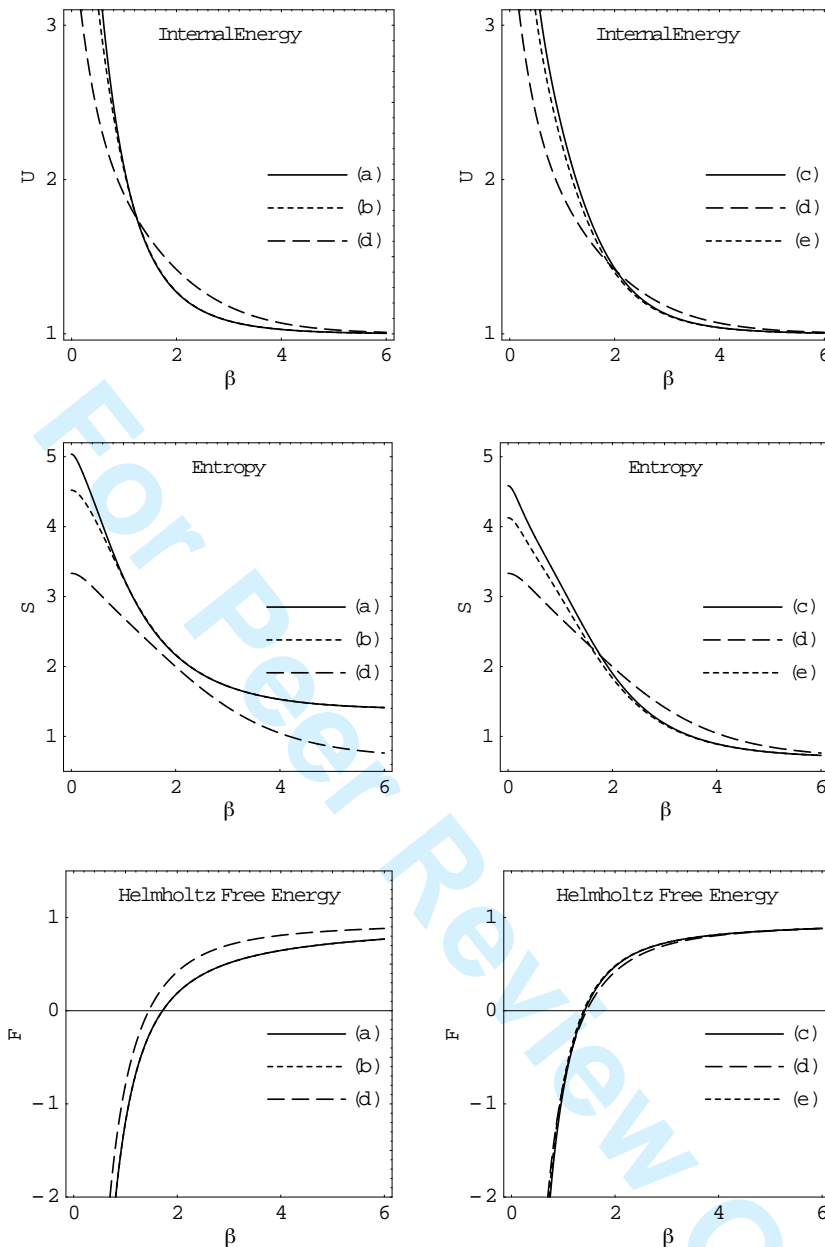


Figure 5. The internal energy  $U$ , Helmholtz free energy  $F$ , and the entropy  $S$  for each of the simulated leaf networks in Fig. 4. Large  $\beta$  corresponds to the early stages of development of each network, while small  $\beta$  corresponds to the fully developed network. The entropy shows the greatest variation in the leaves as a function of  $\beta$ , with the Helmholtz free energy showing the smallest variation. The thermodynamic relation  $F = U - \beta^{-1}S$  indicates that the information provided by the three thermodynamic functions is not independent.

patterns is an active area of research that involves several branches of science and engineering [8–10]. Our objective here is not to account for the variety of naturally occurring venation patterns *per se*, but to use a model of network generation with enough variability to enable us to demonstrate the utility of the methodology described in Sec. 4.

The thermodynamic functions calculated from the triangulations in Fig. 4 are shown in Fig. 5. The thermodynamic profiles of the tree leaves have two regimes that are of particular interest. The high-energy regime ( $\beta \rightarrow 0$ ) is associated with

the late stage of the leaf development process, and the low-energy regime ( $\beta \rightarrow \infty$ ) with the initial stage of development. The internal energy profiles show that, in the early stages of growth, a similar amount work has been done to ‘build’ all five leaves, as expected from the discussion in Sec. 4.3. However, in the later stages of development, the average amount of work for leaf ( $d$ ) becomes the lowest of the five leaves, which agrees with intuitive expectations, since this leaf is the smallest in the sample.

The entropy, in conjunction with (18) and (19) provides a complementary description of leaf development to the internal energies. In the early stages of network development ( $\beta \rightarrow \infty$ ),

$$S_a(\infty) \approx S_b(\infty) > S_c(\infty) \approx S_d(\infty) \approx S_e(\infty). \quad (20)$$

According to (18), the lower panels in Fig. 4 support this order because of the qualitative differences in their first-generation triangulations. Some of these similarities persist as the network develops, but the development of leaf ( $d$ ) is altogether different from the others. Leaves ( $a$ ) and ( $b$ ) and leaves ( $c$ ) and ( $e$ ) shows differences only in the late stages of network expansion. At  $\beta = 0$ , we have

$$S_a(0) > S_b(0) \approx S_c(0) > S_e(0) > S_d(0). \quad (21)$$

Equation (19) indicates that this ordering corresponds to the total triangulations area of the leaves, which again broadly conforms to our expectations based on the lower panels of Fig. 4.

The Helmholtz free energy  $F = U - \beta^{-1}S$  represents a weighted sum of the internal energy and entropy and so does not provide any substantial additional information about our networks. In particular, in the initial stages of leaf development ( $\beta \rightarrow \infty$ ),  $F \approx U$ , so differences between leaves are small, for reasons discussed above, while for the fully-developed leaf ( $\beta \rightarrow 0$ ),  $F \approx \beta^{-1}S$ , which shows the largest differences of this quantity. The internal energy and entropy separately provide a more discriminating interrogation of the venation network. On this basis, we are able to classify the sample leaves in Fig. 4 into three main categories:  $\{a, b\}$ ,  $\{c, e\}$ , and  $\{d\}$ . This grouping is confirmed by the low-energy limit ( $\beta \rightarrow \infty$ ) of the entropy profiles of the tree leaves, and the qualitatively different profile for leaf ( $d$ ). For our purposes the significance of this result is the ability of our method to discriminate between different types of radial networks where the microscopic rules for the formation of such networks are not available.

## 6. Summary and Further Applications

There are two main topics that have been discussed in this paper. We have described a model in which supplying a leaf with water and nutrients has a cost. Minimizing that cost determines the size and shape of the leaf. The transport cost is minimized when nutrients are moved in the same direction, with movement in opposing directions being prohibited. For all other flows, the parameter  $\beta_i$  scales the extent to which the flow direction contributes to the total cost of sustaining the supply network. Another cost is associated with supply flowing through a network edge. The ‘thickness’ of a given edge is proportional to its weight, so the cost of maintaining an edge with a particular weight leads to a second parameter,  $\alpha_f$ . A third parameter,  $\varepsilon_p$ , is an efficiency coefficient for photosynthesis.

While a model for leaf growth is interesting in its own right, both for analyzing tree leaves that occur in nature, as well as highlighting the role of optimal transport

1 in living systems, we have introduced this model as a test for analyzing radial  
2 networks. We have described a method based on a structural triangulation that  
3 associates an energy with each generation of the network. This enables a partition  
4 function to be constructed from which standard thermodynamic functions can be  
5 directly computed. The interpretation of these functions for networks provides a  
6 powerful classification scheme. The comparison of the internal energy, entropy, and  
7 Helmholtz free energy for several simulated leaves, all of which grow by the same  
8 mechanism, but with differently parametrized cost functions, illustrates how our  
9 methodology is able to discriminate between networks that show subtle structural  
10 differences. Our method differs from other thermodynamic formulations of networks  
11 [26, 27] in that individual networks are characterized through their triangulation,  
12 rather than as an average **over ensembles**.

13 The analysis method reported here can be applied to biological networks, such  
14 as the vasculature of the human placenta. There are two levels of description. At  
15 the most basic level is the description of the expansion of such a system with an  
16 accompanying cost function. A benchmark for vascular systems was proposed by  
17 the physiologist Cecil Murray [28] in 1926 as a compromise between the frictional  
18 and metabolic costs, which was expressed as a cost function. The formulation of a  
19 minimum energy hypothesis led to a scaling law,  $Q \propto d^3$ , where  $Q$  and  $d$  are the  
20 volumetric flow rate and the diameter of a blood vessel segment, respectively. This  
21 scaling law is universal for all trees whose internal flows are laminar. Subsequently,  
22 Uylings [29] argued that the exponent in Murray's scaling law can be reduced from  
23 3 to near 2.3 if the fluid flow is turbulent. Numerous studies have found support  
24 for Murray's scaling law [30, 31], but with significant scatter. Moreover, a recent  
25 report [32] on the applicability of Murray's law to botanical trees suggests that  
26 animals and plants have reached similar solutions for efficient fluid transport.

27 An alternative approach in the absence of a cost function for placental vascula-  
28 ture is the method described in Sec. 4, which based solely on statistical topology,  
29 i.e. an analysis only of the network itself. The vascular system of a placenta has  
30 a branching structure within the chorion that resembles a tree. The origin of the  
31 tree is the umbilical cord insertion in the placenta. **The vasculature and its branch-**  
32 **ing points are represented by edges and vertices, in order to apply the method of**  
33 **analysis of our work.** The characterization of placental vasculature is only now be-  
34 coming available in **sufficient numbers** to enable statistically significant conclusions  
35 to be reached. A noteworthy early effort [33], based on scanned x-ray angiograms  
36 to produce a binary image of the edges of the blood vessels, focussed largely on  
37 the fractal dimension of the vascular tree (see below), which is a global (average)  
38 measure of the branching. But we envisage an analysis where the statistical anal-  
39 ysis of a cohort will reveal systematic differences between 'normal' placentas and  
40 those with various abnormalities. Such studies are in progress and will be reported  
41 elsewhere.  
42  
43  
44  
45

#### 46 Acknowledgements

47 R.-K. S. was supported in part by Placental Analytics LLC. C. M. S. was parti-  
48 tially supported by a NARSAD Young Investigator Award and by K23 MidCareer  
49 Development Award NIMH K23MH06785.  
50  
51

#### 52 References

- 53 [1] R. ALBERT AND A.-L. BARABÁSI, *Rev. Mod. Phys.* **74** (2002), p. 47.  
54  
55  
56  
57  
58  
59  
60

- [2] S. BOCCALETTI, V. LATORA, Y. MORENO, M. CHAVEZ, AND D. U. HWANG, *Phys. Rep.* **424** (2006), p. 175.
- [3] S. N. DOROGOVTSSEV, A. GOLTSEV, AND J. F. F. MENDES, *Rev. Mod. Phys.* **80** (2008), p. 1275.
- [4] F. CORSON, *Phys. Rev. Lett.* **104** (2010), 048703.
- [5] J. CLARKE, S. BARMAN, P. REMAGNINO, K. BAILEY, D. KIRKUP, S. MAYO AND P. WILKIN, in *Advances in Visual Computing* (Springer, Berlin, 2006), pp. 427-436.
- [6] E. KATIFORI, G. J. SZÖLLÖSI, AND M. O. MAGNASCO, *Phys. Rev. Lett.* **104** (2010), 048704.
- [7] P. S. DODDS, *Phys. Rev. Lett.* **104** (2010), 048702.
- [8] A. ROTH-NEBELSICK, D. UHL, V. MOSBRUGGER, AND H. KERP, *Ann. Bot.* **87** (2001), p.553.
- [9] A. RUNIONS, M. FUHRER, B. LANE, P. FEDERL, A.-G. ROLLAND-LAGAN, AND P. PRUSINKIEWICZ, *ACM Trans. on Graphics* **24** (2005), p. 702.
- [10] Q. XIA, *ESAIM Control Optim. Calc. Var.* **13** (2007), p. 359.
- [11] L. EULER, *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, **8** (1736), p. 128. [English translation: *Sci. Am.* **189**(1), p. 66; reprinted in *Mathematics in the Modern World* (Freeman, San Francisco, 1968), p. 141.]
- [12] G. KIRCHHOFF, *Poggendorf Ann. Physik.* **72** (1847), p. 497. [English translation: *IRE Trans. Circuit Theory* **5** (1958), p. 4.]
- [13] A. CAYLEY, *Phil Mag.* **13** (1857), p. 172.
- [14] D. KÖNIG, *Theorie der Endlichen und Unendlichen Graphen* (Chelsea, New York, NY, 1950).
- [15] P. ERDŐS AND A. RÉNYI, *Publications Mathematicae* **6** (1959), p. 290.
- [16] P. ERDŐS AND A. RÉNYI, *Publ. Math. Inst. Hung. Acad. Sci.* **5** (1960), p. 17.
- [17] E. N. GILBERT, *Ann. Math. Statist.* **30** (1959), p. 1141.
- [18] M. E. J. NEWMAN, *SIAM Review* **45** (2003), p. 167.
- [19] S. C. CROSS, *J. Pathol.* **182** (1997), p. 1.
- [20] G. B. WEST, J. H. BROWN, AND B. J. ENQUIST, *Science* **284** (1999), p. 1677.
- [21] G. S. KASSAB, *Am. J. Physiol. Heart Circ. Physiol.* **290** (2006), p. H894.
- [22] L. R. FORD AND D. R. FULKERSON, *Can. J. Math.* **8** (1956), p. 399.
- [23] F. L. HITCHCOCK, *J. Math. and Phys.* **20** (1941), p. 224.
- [24] A. N. TOLSTOÏ, *Methods of finding the minimal total kilometrage in cargo transportation planning in space* (in Russian), in *Transportation Planning*, Vol. 1 (TransPress of the National Commissariat of Transportation, Moscow, 1930), p. 23.
- [25] M. NAKAHARA, *Geometry, Topology and Physics* (IOP Publishing Ltd., Bristol, UK, 2003).
- [26] G. BIANCONI, *EPL* **81** (2008), 28005.
- [27] G. BIANCONI, *Phys. Rev. E* **79** (2009) 036114.
- [28] C. D. MURRAY, *Proc. Natl. Acad. Sci. USA* **12** (1926), p. 207.
- [29] H. B. M. UYLINGS, *Bull. Math. Biol.* **399** (1977), p. 509.
- [30] T. F. SHERMAN, *J. Gen. Physiol.* **78** (1981), p. 431.
- [31] P. R. PAINTER, P. EDÉN, AND H. U. BENGTTSSON, *Theor Biol Med Model.* **3** (2006), p. 31.
- [32] K. A. MCCULLOH, J. S. SPERRY, AND F. R. ALDER, *Nature* **421** (2003), p. 939.
- [33] D. L. BERGMAN AND U. ULLBERG, *J. Theor. Biol.* **193** (1998), p. 731.