PHYSICAL REVIEW E **80**, 037102 (2009)

From complex structures to complex processes: Percolation theory applied to the formation of a city

Agnieszka Bitner,[1] Robert Hołyst,[2,3] and Marcin Fiałkowski[2,*]
[1]*Department of Applied Mathematics, Agricultural University, Mickiewicza 24/28, 30-059 Krakow, Poland*
[2]*Institute of Physical Chemistry, Polish Academy of Sciences, Kasprzaka 44/52, 01-224 Warsaw, Poland*
[3]*Department of Mathematics and Natural Sciences, College of Science, Cardinal Stefan Wyszyński University, Dewajtis 5, 01-815 Warsaw, Poland*
(Received 9 April 2009; published 25 September 2009; corrected 28 September 2009)

We investigate the morphology of the spatial pattern resulting from the division of land into the parcels that is observed in the centers of the cities, by analyzing the distribution function of the parcel areas. A simple model based on a two-dimensional bond percolation is employed to mimic the process of the formation of the city. The model reproduces the empirical distribution of the parcel areas that is found to exhibit the power law with the exponent $\tau=2.0$. We argue that the city emerges from a collection of separated settlements in a process that can be described as a structural phase transition.

DOI: 10.1103/PhysRevE.80.037102 PACS number(s): 89.75.Kd, 64.60.ah, 89.65.Lm, 89.75.Da

Many structures in natural and artificial systems arise as a result of complex processes. In some cases it is possible to infer the process from the final structure. The best known examples are spatial two-dimensional (2D) processes [1]. In the present work we seek to link the spatial structure of an artificial system—a city—to the processes that led to its formation. Urban systems have been found to conform to universal scaling laws describing city growth and spatial organization. The best known is the Zipf law [2]. According to this law, the distribution of city population obeys a power law with the exponent $\tau=2$. The Zipf law has been confirmed in a number of studies [3–5]. The distribution of the fraction of cities of a given area follows the power law with $\tau=1.85$ [6,7]. There is much empirical evidence [8,9] that the large-scale urban structure displays fractal morphology. In [10] it has been shown that the spatial arrangement of cities on a plane can be successfully modeled using the classical theory of liquids. Recently, the diffusion-limited aggregation model has been employed [6,7,11,12] to simulate the growth of urban systems in a large scale and successfully reproduced both the fractal morphology and the population power laws. The scaling laws of cities have also been reproduced in discrete [13] and continuous [14] stochastic models. Recent studies [15] on human agglomeration suggest the existence of a universal mechanism governing the population growth at different scales. On a smaller scale, the urban morphology has been studied [16–20] by investigating the structure of the street network. It has been found empirically [18] and reproduced in theoretical model [20] that the distribution of the areas of the cells created by streets obeys a power law with the exponent $\tau=1.9$ [18].

For the first time, the morphology on the *smallest* scale—on the level of the land parcels—has been studied in [21]. In the work cited, the parcel patterns both in urban and rural areas have been analyzed based on the parcel size distribution function $f(A)$. It has been shown that $f(A)$ exhibits universal generic features that are unique for the (i) city core, (ii) suburbs, and (iii) rural areas. In the city core $f(A)$ displays the power-law behavior

$$f(A) \sim A^{-\tau}, \tag{1}$$

with $\tau=2$, and has a peak located at about 10^3 m^2. The core occupies only small part of the urban system and is well approximated by a circle centered in the center of the city—usually in the middle of the central business district (CBD). The radius R_{core} of the core varies from about 0.5 km to about 10 km. The core represents the oldest part of the city with the highest urban density and consists of commercial and residential zones as well as green areas. The core is surrounded by a wide ring of suburbs [22] undergoing into rural area. In suburbs, $f(A)$ follows the logarithmic-normal distribution. In the rural area $f(A)$ exhibits the power-law distribution given by Eq. (1) with $\tau\approx 1.1$.

Our studies have revealed that the properties of the land morphology are universal and robust with respect to geographical, historical, and economical conditions accompanying development of a given area. These regularities indicate that some common generic mechanism underlies the urbanization processes. The results presented in Ref. [21] suggest that the nature of these processes is reflected in the way people divide their land and can be inferred from the morphology of the fragmentation pattern. That is, the parcel pattern is the spatial fingerprint of the urbanization mechanism. Thus, the morphology of the core—the oldest part of the urban system—should contain the information about the process that led to the formation of the city. It is the purpose of this Brief Report to investigate the morphology of the city core and to propose a plausible mechanism for the city formation. Using a bond percolation model that reproduces the empirical findings, we argue that the process of the city formation can be described as the structural phase transition.

The land parcel is the basic spatial unit in a land survey system (a cadaster) and is assigned with a unique parcel number. Physically, the parcel is defined with its boundaries and forms a polygon. In our analysis, we use the geographic information system (GIS) parcel data in the Environmental

*Corresponding author. fialkows@ichf.edu.pl

Systems Research Institute (ESRI) shapefile format, where the parcels are represented by sets of vertexes defining the polygons. The spatial position of the parcel is determined by the location of its centroid calculated as the geometric center of mass of the polygon. The data contain all types of land parcels, including developed and undeveloped areas, green and industrial areas, and public utilities, except for streets and roads. In the present work, we use the GIS data from Ref. [21] and perform the analysis for 29 cities from Australia (20) and USA (8) and for one European (Polish) city. We investigate the parcel morphology of the core to answer the question whether it represents a spatially homogeneous collection of the parcels or if there exists some spatial arrangement of the parcels within the core. In order to answer this question we (1) examined $f(A)$ in concentric rings at increasing distances r from the origin (CBD) and (2) studied the dependence of the average parcel size $\langle A \rangle$ on r.

Our analysis showed that the parcels are distributed homogeneously within the core and do not exhibit any spatial arrangement. We also found that the average parcel size does not display any regular behavior as a function of the distance from CBD, but fluctuates around the value A_0—the average parcel size in the core. In some cases $\langle A \rangle$ displays a shallow minimum located roughly at the half of R_{core}. We found that A_0 was different for each city and oscillated from town to town around the value of about 10^3 m^2. Although the scaling with $\tau=2$ is universal for all the cities, both the shape and the location of the peak are specific for a given urban system. The results of the analysis for the city of Houston (Texas, USA) are shown in Fig. 1.

Our empirical findings indicate that the spatial distribution of the "centers" (new houses, businesses, or public utilities) in the core does not follow the population density law [6,7,23]. According to this law, the distribution of the centers decays exponentially with the distance from the origin. Note that the population density law was employed in Ref. [20] that proposed a mechanism to create roads and reproduced the distribution of the areas of the cells formed by the road network.

To model the city formation process, we consider an urban system as a plane tessellated into streets and parcels. The parcels are aligned along the streets and fill the cells of the street network. Thus, the system of the streets determines the primary tessellation and the resulting distribution of the areas of the parcels. In our simplified approach, we represent the street network on a 2D square grid as a collection of segments of equal length (bonds). We found that the parcels are uniformly arranged within the core. This fact indicates that the urbanization process occurs simultaneously in the whole area. Also, it provides us with the rule to construct the street network: it is assembled from randomly connected bonds. The above picture suggests that a 2D bond percolation model [24] is a suitable tool to mimic the morphology of the urban system. The street network emerges when the number of bonds increases and the system approaches the percolation threshold. Then, patches of regular "street grid" appear. This grid structure is a characteristic of the city core and can be seen in Fig. 1(a). Of course, the percolation model is not capable to reproduce the exact structure of the street network. However, it captures the basic *statistical* properties of the tessellation morphology.

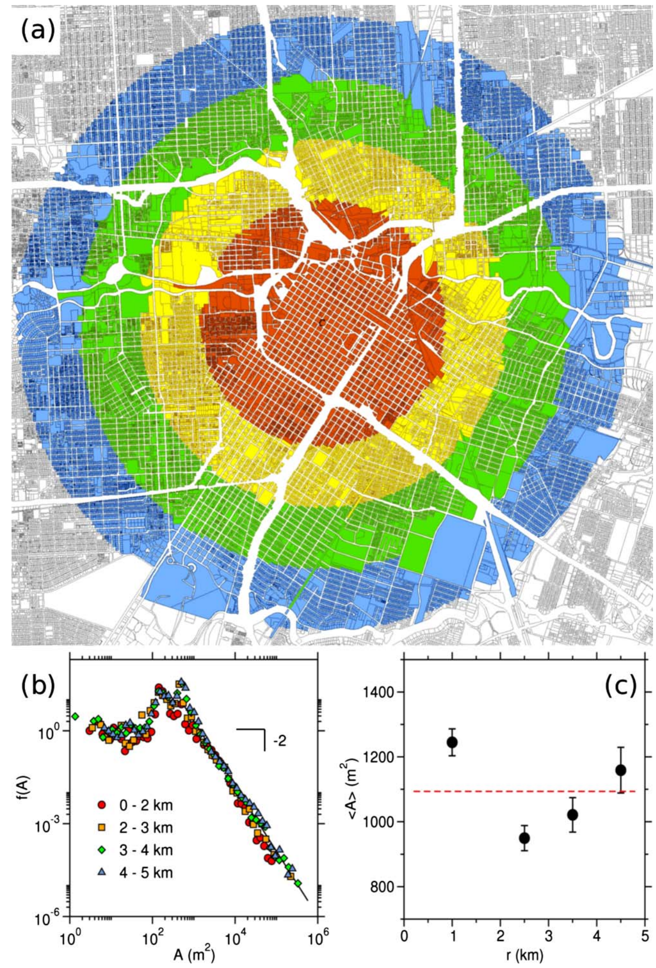

FIG. 1. (Color online) (a) The core of the city of Houston (Texas, USA). The black lines are boundaries of the land parcels. The core is a circle of the radius 5 km containing 44 852 parcels. It is divided into four rings: 0–2 (red), 2–3 (orange), 3–4 (green), and 4–5 km (blue). (b) Plots of the area distribution functions $f(A)$ for the four rings shown in (a). $f(A)$ collapse onto a single curve displaying the power-law tail with the exponent $\tau=2$. (c) The dependence of the average parcel area $\langle A \rangle$ on the distance r from the origin. The data points represent the averages calculated in the middle of each ring. The red line is the average parcel area in the core; $A_0=1083.6$ m^2.

As we found, the morphology of the city exhibits a scaling behavior for the parcel area distribution for the parcel areas $A > 10^3$ m^2 The scaling follows from the self-similar hierarchical spatial structure of the urban system [8]. The self-similarity holds in a statistical sense and reflects an important feature of the processes of the city formation—merging of neighboring urban settlements into bigger and bigger systems. In this respect, the formation of urban system resembles the geometric phase transition in the percolation model. Close to the percolation threshold, the percolating system displays both the self-similar morphology and the scaling of the cluster area distribution. The percolation theory predicts [24–26] that at the critical point the number of clusters of size A normalized per lattice follows Eq. (1) with $\tau=187/91 \approx 2.055$ (the Fisher exponent). The value of

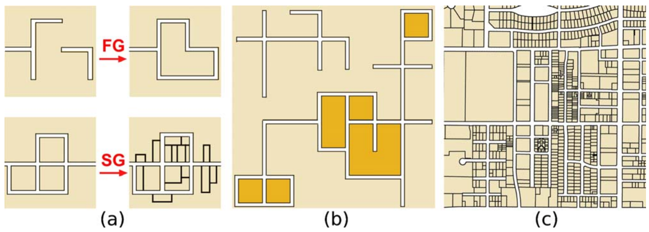

FIG. 2. (Color online) (a) Schematic representation of two generic processes that lead to the formation of the parcel: (top) The street network develops to form loops. The cells encircled by the loops are the FG parcels. (bottom) The SG parcels are cut out from the existing FG parcels. (b) An example illustrating the way of the calculation of the areas of the cells. Here, the system consists of four cells of size 1, one cell of size 2, and one cell of size 4. Neither the tree structure nor the blind protrusions contribute to the cell areas. (c) The parcel pattern observed in the core of the city of Houston. The white stripes are the streets, and the black lines are the parcels boundaries.

the exponent $\tau=2$ observed for all the cities provides strong support for applying the percolation model.

The formation of the city is accompanied by fast land development processes, in which the parcels get built up. The parcel pattern of the built-up areas changes very slowly with time. For this reason the city core displays the frozen morphology that was established at the time when the city was formed. The city emerges when the neighboring urban settlements get connected by a common street network. In our approach, the formation of the city corresponds to the percolation threshold.

We assume that the following two generic processes underlie the land fragmentation into parcels: the parcels are created (1) by the streets when they connect to form a cell (a loop) and (2) along the existing street. The second process leads either to the fragmentation of an existing cell or—if the street is not a part of the loop—to the inclusion of a portion of undeveloped land into the city. These two processes are presented schematically in Fig. 2(a). They can occur simultaneously; the second one dominates however at the late stages of the city formation and results in the parcel pattern shown in Fig. 2(c). Below, we demonstrate that the first process gives rise to the universal power law displayed by the parcel area distribution function, while the second process is specific for a given city and determines the position and the shape of the peak.

We generated the land parcel pattern in two steps. In the first step, the street network was simulated based on the percolation model. We refer to these parcels as to the first generation (FG) ones. Next, the FG parcels were further subdivided and the second generation (SG) of the parcels was created. To simulate the street patterns, we employed the bond percolation model [24] on a two-dimensional square lattice. In this model, the connection between two neighboring nodes in the lattice can be open with the probability p or closed with the probability $1-p$, which are assumed to be independent. Here, the open connection represents a street linking two points separated by the distance d—the lattice constant. Because we simulated the urban system on a finite area, no periodic boundary conditions were applied. To complete the first step of the simulation procedure, we calculated the areas A and the circumferences L for all the cells encircled by the streets. The quantities L and A were measured in the units of the lattice constant d and the area of the unit cell d^2, respectively, and were integer numbers. We analyzed the cell pattern obtained by eliminating all streets forming "blind" protrusions or tree structures, as illustrated in the example of Fig. 2(b).

To perform the next simulation step and generate the SG of the parcels, we applied the following procedure to the collection of the FG parcels: for each FG parcel A^{FG}, we determined the number n of the SG parcels to be cut out from it. To calculate n, we assumed that n is proportional to the total length L of the streets forming the circumference of this parcel, and that it is a random integer variable uniformly distributed in the interval $[0,\eta L]$. Next, the areas $A_1^{\mathrm{SG}},\ldots,A_n^{\mathrm{SG}}$ of the SG parcels were generated assuming that they are normally distributed with the mean A_{m} and the variance σ^2. If the sum of the generated areas was smaller than A^{FG}, the actual number n_{SG} of the SG parcels subtracted from the FG parcel was equal to n. Otherwise, n_{SG} was the biggest number for which the above condition was satisfied. Finally, the updated area A_0^{SG} was obtained as $A_0^{\mathrm{SG}}=A^{\mathrm{FG}}-\sum_{i=1}^{n_{\mathrm{SG}}}A_i^{\mathrm{SG}}$, and the set $\{A_{\mathrm{i}}^{\mathrm{SG}}\}$ of n_{SG} the SG areas was added to the current collection of the parcels. (The area A_0^{SG} is included in the computation of the parcel size distribution.) The quantities η, A_{m}, and σ are parameters of the model. In our approach we assume that the unit cell represents on the lattice the street block formed in the urban street pattern [see Fig. 2(c)]. This follows that the area of the unit cell undergoes a complete fragmentation into the SG parcels. The average number P of the SG parcels generated in the unit cell is equal to $1/A_{\mathrm{m}}$. On the other hand, P can be expressed in terms of the parameter η as $P=4(\eta/2)$. Thus, one gets the following relation:

$$\eta = 1/2A_{\mathrm{m}}. \tag{2}$$

We started the analysis with the investigation of the FG parcel pattern created by the streets. The simulations were carried out for the values of the probability p varying from 0.05 to 0.95 for $N=1024$. For each value of p the simulations were performed for 20 samples. We found that for $p=p_{\mathrm{c}}=0.5$ (the percolation threshold) $f(A^{\mathrm{FG}})$ follows the power law with $\tau=2.056\pm0.007$. Note that at the percolation threshold beyond the spanning cluster also smaller satellite clusters are formed. In our analysis we took into account the parcels belonging to the spanning cluster as well as to the satellite clusters, because the whole system represents the same level of urbanization measured here by the probability p. However, one expects that the statistics calculated for the spanning cluster only and that calculated for all the clusters do not differ much since the system exhibits a self-similarity near the critical point. To check the possible effect of the average cluster size on the statistics, we performed also simulations at p_{c} for N ranging from 256 to 2048 and obtained identical results.

To generate the SG parcels, we assumed that the unit cell corresponds to the street block consisting of $P=4$ parcels. This yielded $A_{\mathrm{m}}=0.25$ and, consequently, from Eq. (2), η

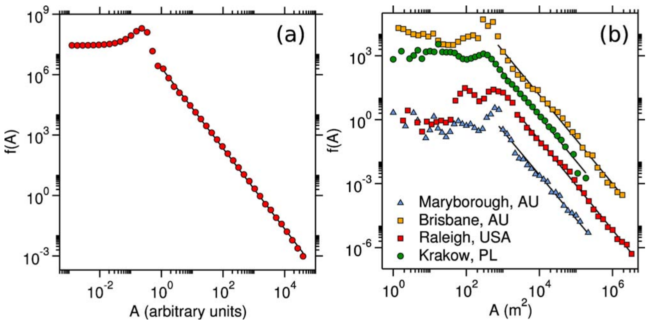

FIG. 3. (Color online) (a) The distribution of parcel areas $f(A)$ obtained in the simulations for the lattice size $N=2048$ for $p=0.5$, $A_m=0.25$, $\sigma=0.1$, and $\eta=2$. $f(A)$ displays the power-law tail with the exponent $\tau=2$. (b) The parcel size distribution obtained for four selected cities, exhibiting the power-law tail with the exponent $\tau=2$. For the sake of clarity, the distributions have been shifted mutually in vertical direction.

=2. We found that the generation of the SG parcels does not affect the power law exhibited by $f(A^{FG})$ and gives rise to the peak in $f(A)$. We also found that σ does not change the shape of $f(A)$ significantly and has an effect on the width of the peak. The result of the simulations obtained at p_c for the lattice size $N=2048$ and $\sigma=0.1$ is shown in Fig. 3(a). For comparison, $f(A)$ for four selected cities are plotted in Fig. 3(b). As seen, the shape of $f(A)$ obtained in our model reproduces—up to the scaling factor—those observed in the cities. (In Fig. 3, we refer to the areas of the SG parcels simply as A.)

To conclude, we proposed a simple model based on the percolation theory to explain the universal scaling exhibited by the land parcel pattern in the city cores. This model assumes that two processes lead to the observed land fragmentation pattern. The first process involves the formation of the street network. We argue that this process gives rise to the universality exhibited by the parcel pattern and can be modeled as the structural phase transition. In the second process the cell structure created by the streets undergoes further fragmentation. The second process is not universal and reflects small-scale "details" that are specific for a given urban system. In the present approach, the evolution of the urban system is assumed to proceed spontaneously without any external control or design and to bring the system to the point (the percolation threshold) where it changes its structure and transforms from the collection of settlements into a new structure—the city.

This work was supported by the Ministry of Science and Higher Education as a scientific project (2007-2010). This project was operated within the Foundation for Polish Science Team Programme co-financed by the European Regional Development Fund Grant No. TEAM/2008-2/2. R.H. acknowledges support from the Foundation for Polish Science.

[1] D. Weaire and N. Rivier, Contemp. Phys. **25**, 59 (1984).
[2] G. K. Zipf, *Human Behavior and the Principle of the Least Effort* (Addison-Wesley, Cambridge, MA, 1951).
[3] X. Gabaix and Y. M. Ioannides, in *Handbook of Urban and Regional Economics: Cities and Geography*, edited by V. Henderson and J. Thisse (North-Holland Publishing Company, Amsterdam, 2004), Vol. 4, pp. 2341–2378.
[4] M. E. J. Newman, Contemp. Phys. **46**, 323 (2005).
[5] A. Blank and S. Solomon, Physica A **287**, 279 (2000).
[6] H. Makse, S. Havlin, and H. E. Stanley, Nature (London) **377**, 608 (1995).
[7] H. A. Makse, J. S. Andrade, M. Batty, S. Havlin, and H. E. Stanley, Phys. Rev. E **58**, 7054 (1998).
[8] M. Batty and P. Longley, *Fractal Cities: A Geometry of Form and Function* (Academic Press, London, 1994).
[9] Y. Chen and Y. Zhou, Chaos, Solitons Fractals **35**, 85 (2008).
[10] L. Glass and W. Tobler, Nature (London) **233**, 67 (1971).
[11] D. H. Zanette and S. C. Manrubia, Phys. Rev. Lett. **79**, 523 (1997).
[12] S. C. Manrubia, D. H. Zanette, and R. V. Sole, Fractals **7**, 1 (1999).
[13] M. Marsili and Y.-C. Zhang, Phys. Rev. Lett. **80**, 2741 (1998).
[14] W. J. Reed, J. Regional Sci. **42**, 1 (2002).
[15] H. D. Rozenfeld, D. Rybski, J. S. Andrade, Jr., M. Batty, H. E. Stanley, and H. A. Makse, Proc. Natl. Acad. Sci. U.S.A. **105**, 18702 (2008).
[16] A. Cardillo, S. Scellato, V. Latora, and S. Porta, Phys. Rev. E **73**, 066107 (2006).
[17] P. Crucitti, V. Latora, and S. Porta, Phys. Rev. E **73**, 036125 (2006).
[18] S. Lammer, B. Gehlsen, and D. Helbing, Physica A **363**, 89 (2006).
[19] D. Volchenkov and P. Blanchard, Phys. Rev. E **75**, 026104 (2007).
[20] M. Barthelemy and A. Flammini, Phys. Rev. Lett. **100**, 138702 (2008).
[21] M. Fialkowski and A. Bitner, Landscape Ecol. **23**, 1013 (2008).
[22] Note that in the Hawaiian islands the urban system represents a "reverse" structure and the core is the most outer part of the system and forms a thin ring along the coast.
[23] C. J. R. Clark, J. R. Stat. Soc. Ser. A (Gen.) **114**, 490 (1951).
[24] D. Stauffer and A. Aharony, *Introduction to Percolation Theory*, 2nd ed. (Taylor and Francis, Cambridge, MA, 1994).
[25] D. C. Rapaport, J. Stat. Phys. **66**, 679 (1992).
[26] J. Cardy and R. M. Ziff, J. Stat. Phys. **110**, 1 (2003).