Emergence in Urban Growth

John Nichol
ESM 569
8 December 2007

Cities exhibit emergent properties. For example, the distribution of the number of cities with respect to population is a power law across multiple spatial, temporal, and cultural scales. Specific emergent phenomena of urban development are discussed. Current theories of urban development are described and evaluated with regard to how well they model emergent properties of urban growth.

Introduction

Cities *are* emergent phenomena. A city exists because of the strong interactions among humans and their resulting collective behavior. Surprisingly, however, cities *themselves* exhibit emergent behavior. For example, the distribution of cities with respect to population is a power across multiple spatial, temporal, and cultural scales. Physicists know this type of emergence within emergence from the study of vortex states, for example, but the study of urban growth provides a new and non-technical example of emergent behavior.

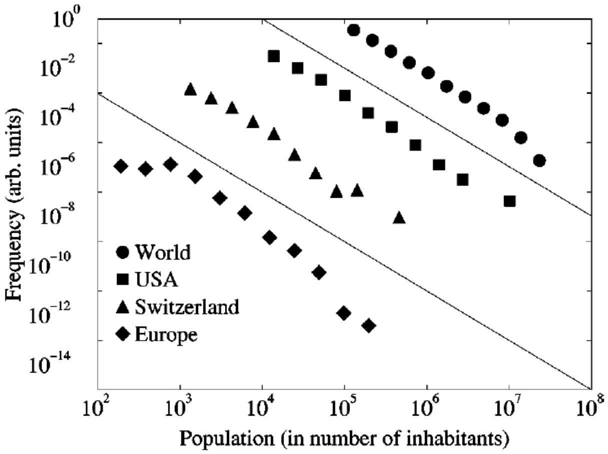
Emergence in urban growth demonstrates that different microscopic theories can give rise to the same macroscopic physics. A successful microscopic theory of urban development taking into account human interactions is a daunting many-body problem. As in superconductivity, researchers have put forth phenomenological theories first. Even though the different theories of urban growth are vastly different, they reproduce the same emergent behavior.

From an intellectual perspective, the problem of urban growth is interesting because it maps onto different areas of physics. The first models of urban growth borrowed from theories of aggregation. The aggregation model, it turns out, is the same as the problem of electric discharge and dielectric breakdown. The dielectric breakdown model yields additional insight into the mechanisms responsible for urban growth. Subsequent models have used percolation theory, stochastic models, and phenomenological differential equations. The study of urban growth may seem at first glance a mundane problem, but deeper investigation unearths interesting and subtle physical problems. In the next section I discuss in more precise terms the emergent properties of urban development. Then I describe and evaluate a few current models of urban growth.

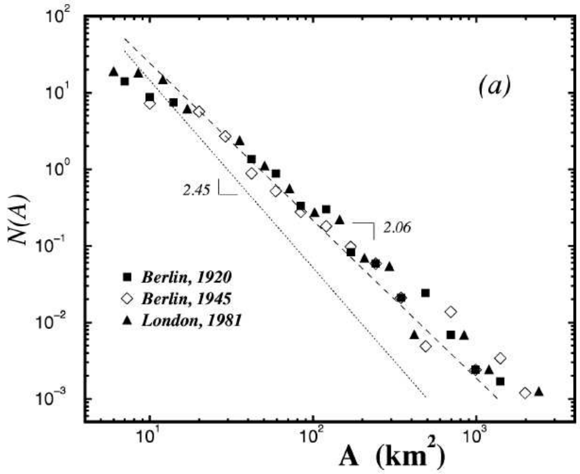
Emergence in Urban Growth

In *Human Behavior and the Principle of Least Effort*, George Kingsley Zipf argues on the basis of the economy of effort that the distribution of the number cities with respect to population ought to follow power law behavior [1]. It does, in fact, and the power law description is so exact and apparently universal that it has come to be known as Zipf's law. Figure 1(a) shows that the population distribution of the 2700 largest cities in the world, the 2400 largest cities in the US, the 1300 largest municipalities in Switzerland, and the 10 largest countries in south Europe all exhibit the same power law behavior with exponent -2. These data show that cities conform to Zipf's law through nearly 5 decades in size and in various cultural settings [2]. The same power law holds for the distribution of cities as a function of area and in different years [3] as shown in Figure 1(b). Note that in the London and Berlin data, the data represent small towns that comprise the larger cities, not actual independent large cities. Remarkably, Zipf's law describes the distributions of countries, municipalities, cities in different countries, cities in the same country, and towns in the same city in different years. In each of these cases, the exponent is always -2.

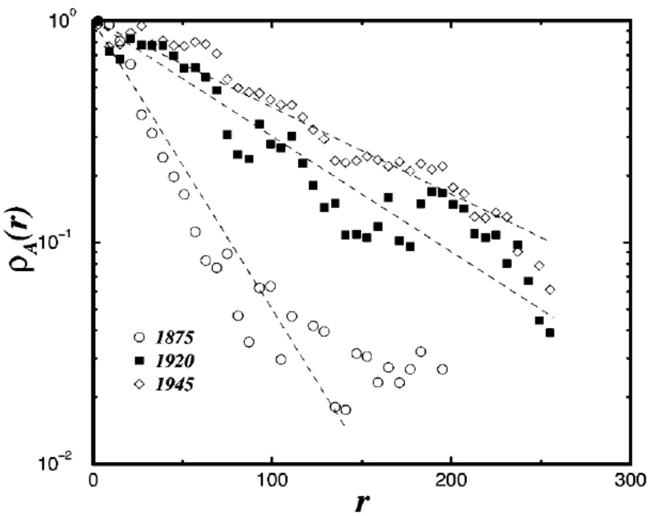
Cities also typically display a population density profile that decays exponentially with distance from the city center, or central business district (CBD) [4]. Over time, the density gradient diminishes, corresponding to the phenomenon of decentralization as in Figure 1(c).



(a)



(b)



(c)

Figure 1. (a) Population frequency for the 2700 largest world cities, 2400 largest cities in the US, 1300 largest municipalities in Switzerland, 10 largest countries in south Europe. The straight lines have slope -2. From [2]. (b) Area distribution for towns in London and Berlin. From [3]. (c) Population density gradient in Berlin over time. From [3]. (d) Images of Berlin in 1875, 1920, and 1945. From [4], said the girl with kaleidoscope eyes.

(d)

Urban developments also exhibit fractal properties. Theories of urban development usually consider a 2-d square lattice, where each lattice site may be occupied or unoccupied. If the number of occupied lattice sites enclosed by a circle of radius R from the CBD scales as

$$N(r) \propto r^{\alpha},$$

and α is not an integer, then α is the *fractal dimension* of the pattern and describes how the area of the pattern scales with its linear size. Equivalently, the fractal dimension of a pattern describes the dimensionality of its boundary. A line with dimension larger than unity fills more space than a straight line. The fractal dimensions of most cities lie between 1.6 and 1.8 [5]. The successive images of Berlin in Figure 1(d) lend qualitative support to the idea that cities exhibit fractal patterns.

Diffusion Limited Aggregation

The earliest models of urban growth used aggregation theory to reproduce some of the emergent features of cities. Aggregation theory at first glance provides an intuitive physical basis for the formation of cities, which are by nature aggregations of people. The traditional model of diffusion limited aggregation (DLA), first introduced by Witten and Sander [6], was invented to model the aggregation of metal particles. The model assumes a seed particle located at the center of a 2-d square lattice. A second particle is introduced at the boundary of the lattice and walks randomly until it either visits a site adjacent to the seed or existing cluster, at which point it becomes part of the cluster, or it walks off of the lattice and is removed. The process is then repeated one particle at a time. The process is "diffusion limited" because the growth of the cluster is controlled by the introduction of random walkers diffusing in from the boundary. Figure 2 shows an image of a simulated aggregation.

A visual comparison of Figure 2 and the sequence in Figure 1(d) shows approximate qualitative agreement. Both images show clusters branching out from the CBD. However, the DLA is more "dendritic" than the real city, which grows more compactly. The tendency of the DLA to grow long arms is the result of nascent dendrites shadowing other areas of the cluster and prohibiting the random walker from reaching all parts of the cluster.

Figure 2. A diffusion limited aggregate. From [6].

Batty and Longley have conducted DLA simulations and have compared their results with the city of Taunton, UK [5]. Through different statistical analyses, the authors conclude that Taunton has a fractal dimension of 1.57, while the DLA simulation has a fractal dimension of 1.66. Although the fractal dimensions roughly agree, the model is less compact than the real city. The density of Taunton is larger than that of the model; 26% of the lattice points in the effective area of Taunton are occupied, but only 5% of the points in the effective area of the DLA simulation are occupied.

A formal development of the DLA model suggests improvements. The growth process should be governed by a diffusion equation since the particles are random walkers. The probability $\phi(x, y, t)$ that a lattice site at position (x, y) will be occupied at a time t depends on the occupation of its neighbors at the previous instant in time. That is,

$$\varphi(x,y,t) = \frac{1}{4}(\varphi(x+dx,y,t-dt) + \varphi(x,y+dy,t-dt) + \varphi(x-dx,y,t-dt) + \varphi(x,y-dy,t-dt)) .$$

Rearranging and adding zero to the equation we obtain

$$\begin{aligned}
&(\phi(x+dx,y,t-dt) - \phi(x,y,t-dt)) - (\phi(x,y,t-dt) - \phi(x-dx,y,t-dt)) \\
&+ (\phi(x,y+dy,t-dt) - \phi(x,y,t-dt)) - (\phi(x,y,t-dt) - \phi(x,y-dy,t-dt)) \\
&= 4(\phi(x,y,t) - \phi(x,y,t-dt)).
\end{aligned}$$

The last equation is the discrete version of the continuum diffusion equation as expected,

$$\partial_t \phi = \nabla^2 \phi .$$

Over a sufficiently long time scale, the flow of particles is uniform, so we may assume that the time derivative is constant. In that case, the diffusion equation reduces to the inhomogeneous Laplace equation. The boundary conditions for the DLA model are that the probability vanishes on the cluster since each site can be only singly occupied, and that the probability reaches unity at a set distance from the center where the particles are released. The problem, as it is now stated, is equivalent to the dielectric breakdown model (DBM) [7], which models an electric discharge in a dielectric medium. Figure 3 shows an image of such a discharge. Note the similarity to the DLA simulation.

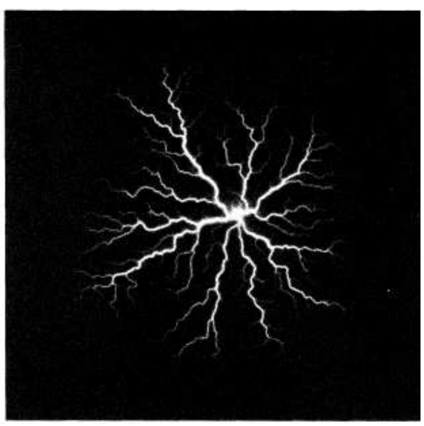


Figure 3. Dielectric breakdown in SF_6. From [7].

Now let us now denote the electric potential by $\phi(x, y, t)$. We expect the probability that the discharge will spread to a lattice point to depend on the local electric field. Thus,

$$p(x,y,t) = \frac{\partial_x \phi(x,y,t) + \partial_y \phi(x,y,t)}{\sum \partial_x \phi(x,y,t) + \partial_y \phi(x,y,t)} .$$

Since only points neighboring the discharge can join it, and since the potential inside the discharge vanishes,

$$p(x,y,t) = \frac{\phi(x,y,t)}{\sum \phi(x,y,t))} .$$

Niemeyer et al. [7] suggest a generalization of the form

$$p(x,y,t) = \frac{\varphi(x,y,t)^{\eta}}{\sum \varphi(x,y,t)^{\eta}}.$$

For $\eta = 1$, the model reproduces DLA results. For $\eta << 1$ the model yields patterns with fractal dimension near 2, or solid shapes. For $\eta >> 1$ the model yields patterns with a fractal dimension near 1, like straight lines. Batty and Longley have used this modification of the DLA model to produce other urban development simulations [5].

By way of conclusion, note first that the DLA model simulates only one cluster at a time, so that it cannot possibly reproduce Zipf's law. An additional shortcoming of the DLA model is that it predicts a power law variation of the density away from the CBD, but the data suggest an exponential decay [4]. In addition, the DBM also provides an intuitive basis for thinking about DLA simulations of urban development. In the DBM, growth occurs at the sites of highest potential, which occur farthest away from the center of the cluster. If we think of the potential as representing available space for growth, then we should expect growth to occur farthest from the center of the cluster, as is the case in the DBM. Contrast this interpretation to the random walking of the particle in the DLA model, which never occurs in practice. Prospective city dwellers examine their surroundings before choosing a place, but they do not walk randomly in from infinity. The use of the Laplace or diffusion equation also suggests that urban dwellers arrange themselves to smooth out the population density.

Correlated Percolation

In order to simulate multiple urban clusters simultaneously, some workers have used percolation theory to describe urban development [3, 4]. In brief, percolation theory considers the problem of clusters on a 2-d lattice. Let us assume a square 2-d lattice with side length L with sites indexed by $\mathbf{r} = (i, j), i, j = 1, ..., L$. With each site we associate a number $u(\mathbf{r})$, which may be random or not, and we define a concentration p such that if $u(\mathbf{r}) < p$ the site is unoccupied, and if $u(\mathbf{r}) > p$ the site is occupied. In uncorrelated percolation theory, the site variables are assumed uncorrelated, that is $\langle u(\mathbf{r})u(\mathbf{r}')\rangle = \delta(\mathbf{r} - \mathbf{r}')$, where the brackets denote an average with respect to a Gaussian distribution. Percolation theory asks, what is the percolation threshold p_c at which point a cluster of a certain size exists?

Motivated by the apparent long-range interactions among city dwellers, Makse et al., have used correlated percolation theory (CPT) in the presence of a concentration gradient as a model for urban development [3, 4]. They arrange for the sequence of lattice numbers to have the long-range correlation $\langle w(\mathbf{r})w(\mathbf{r}')\rangle \propto |\mathbf{r} - \mathbf{r}'|^{-(\alpha - d)}$. The procedure to create this correlation is to convolve an original uncorrelated sequence with a power law kernel [10]. Consider a sequence of random numbers $u(\mathbf{r})$, whose Fourier transform is $\tilde{u}(\mathbf{k}) = \int d\mathbf{r} u(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}}$. We define a new sequence $\tilde{w}(\mathbf{k}) = |\mathbf{k}|^{-\alpha/2}\,\tilde{u}(\mathbf{k})$ by multiplying the sequence in Fourier space (convolving in real space) by a power law kernel. The correlation function of the new sequence is a power law in distance:

$$\begin{aligned}\langle w(\mathbf{r})w(\mathbf{r}+\mathbf{R})\rangle &= \left\langle \int d\mathbf{k}\,\tilde{u}(\mathbf{k})\overline{\mathbf{k}^{-\alpha/2}}e^{-i\mathbf{k}\bullet\mathbf{r}}\int d\mathbf{k}'\tilde{u}(\mathbf{k}')\mathbf{k}'^{-\alpha/2}\,e^{i\mathbf{k}'\bullet(\mathbf{r}+\mathbf{R})}\right\rangle \\ &= \iint d\mathbf{k}d\mathbf{k}'\overline{\mathbf{k}^{-\alpha/2}}\mathbf{k}'^{-\alpha/2}\,e^{-i\mathbf{k}\bullet\mathbf{r}}e^{i\mathbf{k}'\bullet(\mathbf{r}+\mathbf{R})}\left\langle \tilde{u}(\mathbf{k})\tilde{u}(\mathbf{k}')\right\rangle\end{aligned}$$

where

$$\begin{aligned}\left\langle \tilde{u}(\mathbf{k})\tilde{u}(\mathbf{k}')\right\rangle &= \left\langle \int d\mathbf{r}u(\mathbf{r})e^{+i\mathbf{k}\bullet\mathbf{r}}\int d\mathbf{r}'u(\mathbf{r}')e^{-i\mathbf{k}'\bullet\mathbf{r}'}\right\rangle \\ &= \iint d\mathbf{r}d\mathbf{r}'e^{+i\mathbf{k}\bullet\mathbf{r}}e^{-i\mathbf{k}'\bullet\mathbf{r}'}\langle u(\mathbf{r})u(\mathbf{r}')\rangle \\ &= \iint d\mathbf{r}d\mathbf{r}'e^{+i\mathbf{k}\bullet\mathbf{r}}e^{-i\mathbf{k}'\bullet\mathbf{r}'}\delta(\mathbf{r}-\mathbf{r}') \\ &= \int d\mathbf{r}e^{-i(\mathbf{k}'-\mathbf{k})\bullet\mathbf{r}} \\ &= \delta(\mathbf{k}'-\mathbf{k})\end{aligned}$$

and

$$\begin{aligned}\langle w(\mathbf{r})w(\mathbf{r}+\mathbf{R})\rangle &= \iint d\mathbf{k}d\mathbf{k}'\overline{\mathbf{k}^{-\alpha/2}}\mathbf{k}'^{-\alpha/2}\,e^{-i\mathbf{k}\bullet\mathbf{r}}e^{i\mathbf{k}'\bullet(\mathbf{r}+\mathbf{R})}\left\langle \tilde{u}(\mathbf{k})\tilde{u}(\mathbf{k}')\right\rangle \\ &= \iint d\mathbf{k}d\mathbf{k}'\overline{\mathbf{k}^{-\alpha/2}}\mathbf{k}'^{-\alpha/2}\,e^{-i\mathbf{k}\bullet\mathbf{r}}e^{i\mathbf{k}'\bullet(\mathbf{r}+\mathbf{R})}\delta(\mathbf{k}-\mathbf{k}') \\ &= \int d\mathbf{k}\,|\,\mathbf{k}\,|^{-\alpha}e^{i\mathbf{k}\bullet\mathbf{R}} \\ &\propto R^{-(\alpha-d)}\end{aligned}.$$

where d is the dimension of the lattice.

Since in reality the lattice is discrete, short wavelength aliasing can spoil the intended correlation. To avoid cutoff effects, the actual power law kernel used is slightly more complicated [11]:

$$S(\mathbf{k}) = \frac{2\pi}{\Gamma(\alpha+1)}\left(\frac{k}{2}\right)^{\alpha} K_{\alpha}(k)$$

where $K_\beta(k)$ is the modified Bessel function which assumes a power law asymptotic form for $k << 1$. For $k >> 1$ the modified Bessel function decays exponentially, avoiding short wavelength aliasing.

For the remainder of this discussion we shall redefine $\alpha \equiv \alpha - d$ to be the correlation exponent. The case $\alpha = 0$ represents the most strongly correlated case, while $\alpha \to \infty$ represents the uncorrelated system. Makse et al. also modify the usual percolation problem to include a concentration gradient, that is $p = p(\mathbf{r}) = e^{-\lambda r}$, to account for the exponential decay of the density with distance from the CBD [3, 4]. In their theory the density gradient and correlation exponent are the only tunable parameters. As might be expected, clusters become increasingly compact as

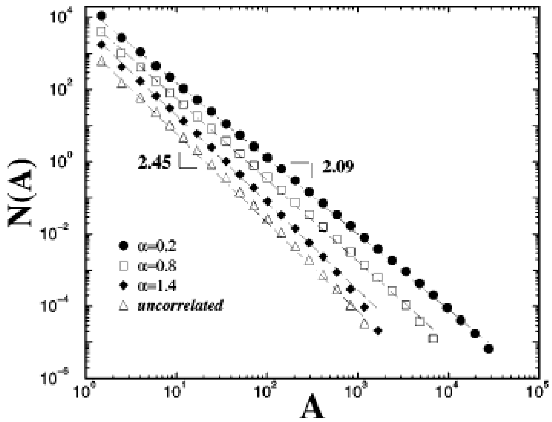


Figure 4. Cluster area distribution function for different correlation exponents. From [3].

the correlation exponent diminishes. The model yields the power law behavior of the area distribution of clusters with an exponent that depends on the correlation exponent as shown in Figure 4. The authors reproduce the exponent of -2 in Zipf's law with a low correlation exponent or large actual correlation. Visual inspection shows that the clusters produced resemble real urban morphologies (Figure 5). The fractal dimension of the strongly correlated clusters produced is 1.4 which roughly agrees with data mentioned above.

The CPT model is an improvement over the DLA model because it models multiple clusters simultaneously, and as a result, it faithfully reproduces more emergent features of urban growth. One disadvantage of the CPT model is that the exponential decay of the population density away form the center is an explicit assumption, and λ is a tunable parameter. In reality, the density profile is an emergent property of urban growth; urban planners do not conspire to create cities with exponential profiles. A complete model of urban growth should incorporate this feature as an emergent property. However, the CPT model does account for decentralization because the same correlation exponent when coupled with a realistic time varying density gradient successfully models the time evolution of Berlin [3, 4].

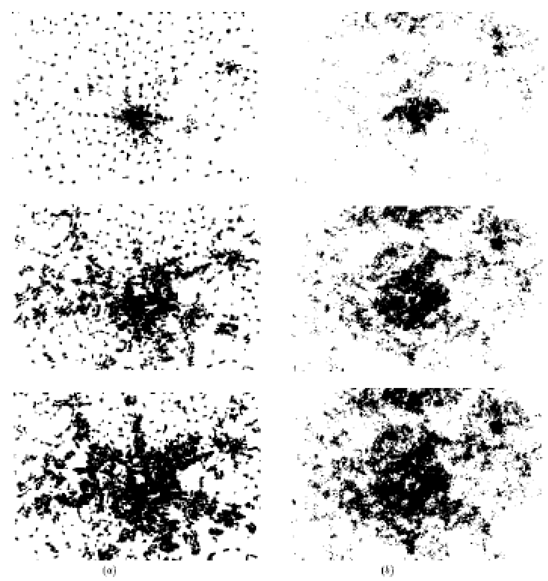


Figure 5. Left: Images of Berlin in 1875, 1920, and 1945. Right: CPT simulation. From [3].

Intermittency model

Although the CPT model represents an improvement over DLA, the tunability of the density gradient detracts from the completeness of the theory. A new theory, called the intermittency model (IM) [2,12,13] recovers nearly all of the emergent phenomena of urban development. The IM considers a 2-d square lattice. Each time step consists of two events, a stochastic, or intermittent, event in which a lattice site increases or decreases its occupation by a fixed fraction, and a diffusion process which smoothes out the population distribution. The motivation for this model assumes opposing forces of attraction and repulsion in urban growth. Cities attract people, but urban dwellers avoid places with the highest population density. In addition, temporal intermittency might be considered a natural part of a theory of urban development because cities are, by nature, spatially intermittent.

Formally, the model associates with each point $\mathbf{x} = (i, j), i, j = 1, \ldots, L$ on a 2-d lattice a number $n(\mathbf{x}, t)$ describing its occupation for each time step. Two processes occur in each time step. The first is an intermittent, or reaction, process, where

$$n(\mathbf{x}, t') = \xi(\mathbf{x}, t) n(\mathbf{x}, t)$$

with $t < t' < t+1$, and

$$\xi(\mathbf{x}, t) = \begin{cases} (1-q)\, p^{-1} \text{ with probability } p \\ q(1-p)^{-1} \text{ with probability } q \end{cases}$$

where $0 \le q, p \le 1$. This reaction event conserves the total population, since

$$\begin{aligned} \langle n(\mathbf{x}, t') \rangle &= \frac{1-q}{p} p \langle n(\mathbf{x}, t) \rangle + \frac{q}{1-p} (1-p) \langle n(\mathbf{x}, t) \rangle \\ &= \langle n(\mathbf{x}, t) \rangle \end{aligned}.$$

Even though the population is conserved, higher moments of the population density diverge, suggesting spatial intermittency.

After each reaction step, a diffusion process occurs, which smoothes out the population

$$n(\mathbf{x}, t+1) = (1-\alpha) n(\mathbf{x}, t') + \frac{\alpha}{k} \sum_{\mathbf{x}' \in \{\mathbf{x}\}} n(\mathbf{x}', t')$$

where α is a control parameter describing how much of the population leaves a given cell at $\mathbf{x}$ and is distributed to the cells in its neighborhood $\{\mathbf{x}\}$. The model has not yet been solved exactly, but a mean field approximation is instructive. If some fixed fraction of the average population per site n_0 diffuses into each cell at a given time, then we can write the entire model as follows, setting the average population to unity,

$$n(\mathbf{x}, t+1) = \begin{cases} (1-\alpha)(1-q)\, p^{-1} n(\mathbf{x}, t) \ + \ \alpha \text{ with probability } p \\ (1-\alpha) q (1-p)^{-1} n(\mathbf{x}, t) \ + \ \alpha \text{ with probability } q \end{cases}.$$

In the limit as $q \to 0$ and $\alpha < 1 - p$, the populations develop inhomogeneities, as expected for cities. In this limit, it can be shown that the frequency of occurrence $f(n)$ of a city with population n goes as

$$f(n) \sim n^{-z}$$

with

$$z = 1 + \ln p / \ln[p/(1-\alpha)].$$

The exponent depends on both the diffusion and reaction constants. Had diffusion been left out, the distribution would follow a log-normal form [2], in disagreement with actual cities.

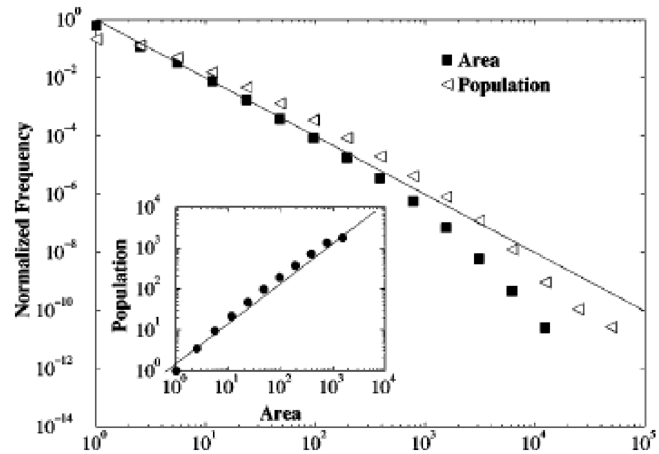


Figure 6. Distribution of city areas and populations in the IM for time varying parameters. The straight line has slope -2. Inset: the relation between population and are is shown with straight line of slope 1.From [3].

The results of numerical simulations are impressive as the theoretical development foreshadows. Figure 6 displays the population and area distribution of the cities from the simulations. Note the excellent agreement with the straight line, which has slope -2. Figure 7 shows the population density versus distance from city center. The profile is exponential, as expected. Figure 8 displays the fractal dimension of a typical cluster. Analysis of many different clusters yields fractal dimensions between 1.15 and 1.35. Remarkably, if local, not global diffusion takes place, the resulting exponents do not depend on the specific values of the parameters α and p if they are within a particular range! In fact the two parameters can vary with time in their appropriate ranges during the simulation with no change in final results.

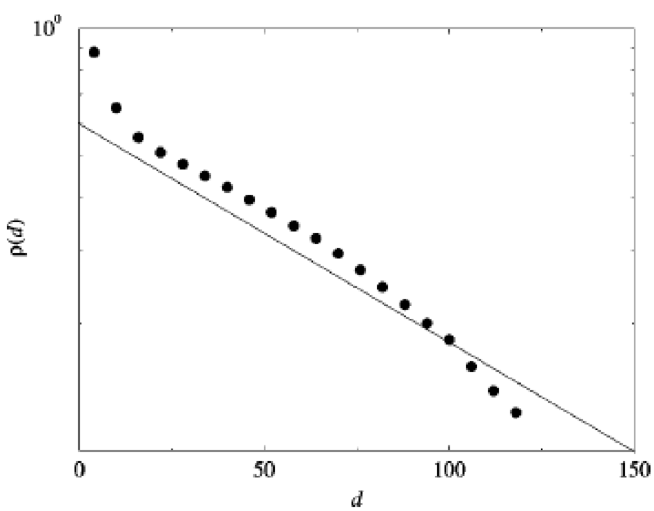


Figure 7. Semi-log plot of population density as a function of distance from CBD in IM cluster. From[3]

The intermittency model appears to recover nearly all of the emergent phenomena in urban developments described earlier. Several comments are in order. First, the IM appears not include any interactions between lattice sites besides local diffusion and the conservation of population. Thus, at first glance, the emergent phenomena in cities may not be a result of interactions beyond diffusion. Rather, this model seems to suggest that emergent behavior may depend only on collective patterns resulting from stochastic fluctuations in large populations.

Second, the simulated clusters eventually die out as a result of strong fluctuations in the system. Corrective measures in the simulation at each time step must be taken to ensure that small clusters survive the effects of repeated fluctuations. Although not enough to tarnish the success of this model, the vulnerability to extinction of clusters suggests that this model may be lacking in some regard.

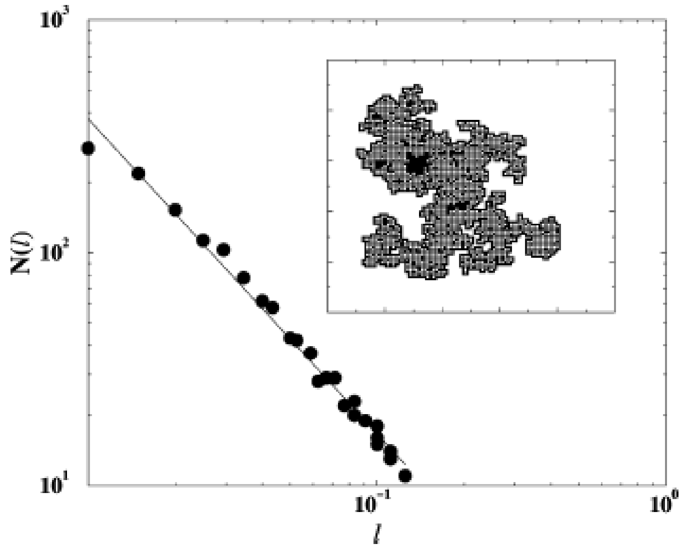


Figure 8. Fractal dimension of a typical IM cluster. From [3].

Marsilli and Zhang [8] have proposed another approach to the intermittency model. They suggest a framework using a master equation. Given Q cities with m_i citizens in the i^{th} city, they define the rates of growth and decrease by $w_a(m_i)$ and $w_d(m_i)$ respectively. The i^{th} city has probability $w_a(m_i)dt$ of increasing its population by one person in a time interval dt. Similarly, $w_d(m_i)dt$ represents the probability that one person will depart the city in that time interval. Additionally, there is some probability pdt that a new city will form. The average number $q_{m,t}$ of cities of size m at time t obeys the equation

$$\begin{aligned}\partial_t q_{m,t} &= w_d(m+1)q_{m+1,t} - w_d(m)q_{m,t} + p\delta_{m,1} \\ &+ w_a(m-1)q_{m-1,t} - w_a(m)q_{m,t}\end{aligned}.$$

Note that this equation does not conserve the number of people or cities. If the transition rates are linear in m, that is $w_a(m) = Am$ and $w_d(m) = Dm$ with $A = (1-p)/n$ and $D = 1/n$ where n is the average number of people in all cities (each departing citizen has a chance to make a new city), then solving the master equation in the steady state yields

$$q_m \sim \frac{1}{m} e^{-pm}$$

in contradiction to Zipf's law. Note that including only linear terms corresponds to assuming no interactions between individuals. If we assume pairwise interactions among the urban dwellers, then $w \propto m^2$, and

$$q_m \propto \frac{p}{1-p}\frac{(1-p)^m}{m^2} \sim \frac{1}{m^2}$$

in agreement with Zipf's law. Since the number of departures and arrivals in a city are proportional to m^2, the fluctuations dm in the population should depend linearly on m. Note that the intermittency model assumes exactly this: the stochastic fluctuation in each time step is proportional to the population. The assumption of independent individuals, or arrivals and departures proportional to m in the master equation, leads to fluctuations of order $\sqrt{m}$, which is not in agreement with the intermittency model. Although the IM seems not to include interactions, pairwise interactions may, in fact, account for the stochastic fluctuations assumed in the IM.

Concluding remarks

The theories discussed here are all phenomenological. While they faithfully reproduce many of the emergent phenomena of cities, they lack a microscopic explanation in terms of intuitive human interactions. Some workers have begun to develop microscopic theories of urban development. Andersson et al. [9] have proposed a Markov random field model which extends the theory of urban development beyond pairwise interactions to include realistic human interactions.

One might wonder if theories of urban development could contribute to urban planning. Makse et al. [2] offer some relevant comments. As Figure 1(b) shows, the scaling properties of Berlin appear to have remained constant from 1875-1945 regardless of any enforced urban policies. London also passed a "green belt" policy in the 1930s aimed to protect areas of natural land around the city. This policy was enforced beginning in the 1950s. However, Figure 1(b) again shows that the scaling properties of London in 1981 agree with apparently universal behavior, suggesting that the green belt policy did little to determine the shape of the city. It seems that only the degree and type of interaction between units in the city, or the correlation exponent, determines the scaling properties of the city morphology.

In applying theories or urban development to public policy planning, we would do well to heed the warnings of Kadanoff [14] that different models may lead to different and even contradictory policy conclusions. Even as a theoretical pursuit, however, the study of urban growth remains a fascinating example of emergent behavior.

References

[1] Zipf, George Kingsley. *Human Behavior and the Principle of Least Effort*. Cambridge, MA: Addison Wesley Press Inc., 1949.

[2] S.C. Manrubia and D.H. Zanette, Phys. Rev. E **58**, 295 (1998).

[3] H. A. Makse et al., Phys. Rev. E **58**, 7054 (1998).

[4] H. A. Makse et al., Nature **377**, 608 (1995).

[5] Batty, Michael and Paul Longley. *Fractal Cities*. New York: Academic Press, 1994.

[6] T. A. Witten and L. M. Sander, Phys. Rev. Lett. **47**, 1400 (1981).

[7] L. Niemeyer, L. Pietronero, and H. J. Wiesmann, Phys. Rev. Lett. **52**, 1033 (1984).

[8] M. Marsili and Y-C Zhang. Phys. Rev. Lett. **80**, 2741 (1998).

[9] C Andersson et al. Phys Rev E. **66**, 026204 (2002).

[10] S. Prakash et al., Phys. Rev. A. **46**, 1724 (1992).

[11] H. A. Makse et al. Phys. Rev. E. **53**, 5445 (1996).

[12] D. H. Zanette and S. C. Manrubia, Phys. Rev. Lett. **79**, 53 (1997).

[13] S. C. Manrubia, D. H. Zanette, and R. V. Sole, Fractals. **7**,1 (1999).

[14] Kadanoff, Leo. *From Order to Chaos: Essays: Critical, Chaotic, and Otherwise*. River Edge, New Jersey: World Scientific Publishing Co., 1993.