SCALE-INVARIANT CORRELATIONS IN THE SOCIAL SCIENCES

H. EUGENE STANLEY
Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215 USA
LUIS A. N. AMARAL
Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215 USA
SERGEY V. BULDYREV
Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215 USA
SHLOMO HAVLIN
Minerva Center and Department of Physics, Bar-Ilan University, Ramat Gan, Israel
TIMOTHY H. KEITT
Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501 USA and Department of Biology, University of New Mexico, Albuquerque, NM 87131 USA
HERNAN A. MAKSE
Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215 USA
AND
GANDHI VISWANATHAN
Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215

1. Introduction

We are going to look at some examples of scale-invariant correlations that are of interest to social scientists.

At one time, it was imagined that the "scale-free" phenomena are relevant to only a fairly narrow slice of physical phenomena [1]. However, the range of systems that apparently display power law and hence scale-invariant correlations has increased dramatically in recent years, ranging from base pair correlations in noncoding DNA [2, 3], lung inflation [4, 5] and interbeat intervals of the human heart [6–9] to complex systems involving large numbers of interacting subunits that display "free will," such as city growth [10–12] and even economics [13–15].

2. Scale Invariance of Animal Behavior

The Wandering Albatross, a giant seabird, was recently the subject of a popular work of the noted science writer Ivars Peterson [16, 17]. He describes an analysis done in collaboration with three workers at the British Antarctic Service, who have been leg-banding these birds with tracking devices [18]. On analyzing the data, we found that the migratory paths of these birds obey Lévy flight statistics, and recently other foraging animals were found to obey well-defined statistical rules [19].

3. Scale Invariance in Human Behavior: Urban Growth Patterns

Predicting urban growth is important for the challenge it presents to theoretical frameworks for cluster dynamics [20–22]. Recently, the model of diffusion limited aggregation (DLA) has been applied to describe urban growth [20], and results in tree-like dendritic structures which have a core or "central business district" (CBD). The DLA model predicts that there exists only one large fractal cluster that is almost perfectly screened from incoming "development units" (people, capital, resources, etc), so that almost all the cluster growth occurs in the extreme peripheral tips. In a recent work [10] an alternative model to DLA that better describes the morphology and the area distribution of systems of cities, as well as the scaling of the urban perimeter of individual cities, has been developed. The results agree both qualitatively and quantitatively with actual urban data. The resulting growth morphology can be understood in terms of the effects of interactions among the constituent units forming a urban region, and can be modeled using the correlated percolation model in the presence of a gradient.

In the model one takes into account the following points:

(i) Urban data on the population density $\rho(r)$ of actual urban systems are known to conform to the relation [23] $\rho(r) = \rho_0 e^{-\lambda r}$, where r is the radial distance from the CBD situated at the core, and λ is the density gradient. Therefore, in our model the development units are positioned with an occupancy probability $p(r) \equiv \rho(r)/\rho_0$ that behaves in the same fashion as is known experimentally.

(ii) In actual urban systems, the development units are not positioned at *random*. Rather, there exist *correlations* arising from the fact that when a development unit is located in a given place, the probability of adjacent development units increases naturally — i.e., each site is not independently occupied by a development unit, but is occupied with a probability that depends on the occupancy of the neighborhood.

In order to quantify these ideas, consider the *correlated* percolation model [24, 25]. In the limit where correlations are so small as to be negligible, a site at position r is occupied if the occupancy variable $u(r)$ is smaller than the occupation probability $p(r)$; the variables $u(r)$ are uncorrelated random numbers. To introduce correlation among the variables, convolute the uncorrelated variables $u(r)$ with a suitable power law kernel [25], and define a new set of random variables $\eta(r)$ with long-range power-law correlations that decay as $r^{-\alpha}$, where $r \equiv |r|$. The assumption of power-law interactions is motivated by the fact that the "decision" for a development unit to be placed in a given location decays gradually with the distance from an occupied neighborhood. The correlation exponent α is the only parameter to be determined by empirical observations.

To discuss the morphology of a system of cities generated in the present model, Makse and co-workers performed simulations of correlated urban systems for a fixed value of the density gradient λ, and for different degrees of correlations. The correlations have the effect of agglomerating the units around an urban area. In the simulated systems the largest city is situated in the core, which is regarded as the attractive center of the city, and is surrounded by small clusters or "towns." The correlated clusters are nearly compact near their centers and become less compact near their boundaries, in qualitative agreement with empirical data on actual large cities such as Berlin, Paris and London [20, 26].

So far, we have argued how correlations between occupancy probabilities can account for the irregular morphology of towns in an urban system. The towns surrounding a large city like Berlin are characterized by a wide range of sizes. We are interested in the laws that quantify the town size distribution $N(A)$, where A is the area occupied by a given town or "mass" of the agglomeration, so we calculate the actual distribution of the areas of the urban settlements around Berlin and London, and find that for both cities, $N(A)$ follows a power-law.

This new result of a power law area distribution of towns, $N(A)$, can be understood in the context of our model. Insight into this distribution can be developed by first noting that the small clusters surrounding the largest cluster are all situated at distances r from the CBD such that $p(r) < p_c$ or $r > r_f$. Therefore, we find $N(A)$, the cumulative area distribution of clusters of area A, to be

$$N(A) \equiv \int_0^{p_c} n(A,p)\ dp \sim A^{-(\tau+1/d_f\nu)}.$$

Here, $n(A,p) \sim A^{-\tau} g(A/A_0)$ is defined to be the average number of clusters containing A sites for a given p *at a fixed distance* r, and $\tau = 1 + 2/d_f$. Here, $A_0(r) \sim |p(r) - p_c|^{-d_f\nu}$ corresponds to the maximum typical area occupied by a cluster situated at a distance r from the CBD, while $g(A/A_0)$ is a scaling function that decays rapidly (exponentially) for $A > A_0$. The exponent $\nu = \nu(\alpha)$ is defined by $\xi(r) \sim |p(r) - p_c|^{-\nu}$, where $\xi(r)$ is the connectedness length that represents the mean linear extension of a cluster at a distance $r > r_f$ from the CBD.

4. Scale Invariance of Human Behavior: Finance and Economics

In 1963, Benoit Mandelbrot wrote a seminal article about fluctuations in cotton prices [13], described in many popular books about fractals. In it, he points out the possibility of scaling in financial indices. His analysis has been recently extended to data sets available now [14], and the presence of scale invariance has been confirmed. Furthermore, it appears that the distribution function conforms to a truncated Lévy flight distribution (a Lévy distribution with a exponential truncation in the wings [27]). Recently, the general approach of Mantegna and Stanley has been extended to study the scale invariance of one measure of the volatility of a financial index [28, 29].

Economics is different than finance, and we have also looked at economic data. Specifically, in collaboration with a card-carrying economist, Michael Salinger — we studied the possibility that all the companies in a given economy might interact, more or less, like an Edwards-Anderson spin glass. As in an Edwards-Anderson spin glass, each spin interacts with other spins — but not with the same coupling and not even with the same sign.

If the sales in a given company x decrease by, e.g., 10%, it will have repercussions in the economy. Some of the repercussions will be favorable — company y, which competes with x, may experience an increase in market share. Others will be negative — service industries that provide personal services for company x employees may experience a drop-off in sales as employee salaries will surely decline. There are positive and negative correlations for almost any economic change. Can we view the economy as a complicated Ising system or spin glass?

To approach this interesting bit of statistical “poetry” and make sense of it, we first located and secured a database that lists the actual size of every firm in the United States. With this information, we did an analysis

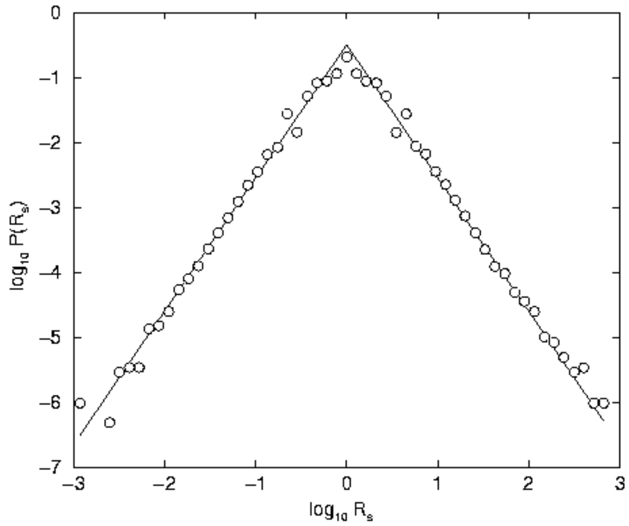

Figure 1. Distribution of population growth rates $R_s \equiv N_s(t+1)/N_s(t)$ across all species in an entire 31-year data set. The growth rate R_s is calculated by dividing species abundances in successive years. Abundances are taken as the total number of individuals of a particular species counted within each survey route [34].

to determine how the distribution of firm size changes from one year to the next. We then made a histogram for each of three characteristic firm sizes. The largest firms have a very narrow distribution — plausible because the percentage of size change from year to year for the largest firms cannot be that great. On the other hand, a tiny company or a garage-based start-up can radically increase (or decrease) in size from year to year. The histograms have a width determined by the size of the firm. When this width is plotted on the y axis of log-log paper as a function of the size of the firm on the x axis, the data are approximately linear *over 8 orders of magnitude*, from the tiniest firms in the database to the largest. The width scales as the firm size to an exponent β, with $\beta \approx 1/6$ [15, 30]. We can therefore normalize the growth rate and show that all the data collapse on a single curve — demonstrating the scaling of this measure of firm size.

Why does this occur? We're working on that. We model this firm structure as an approximate Cayley tree, in which each subunit of a firm reacts to its directives from above with a certain probability distribution. This model, developed primarily by Sergey Buldyrev, seems to be consistent with the critical exponent $-1/6$ [31]. More recently, Amaral et al. [32] have proposed a microscopic model, and Takayasu [33] has extended the empirical results to a wide range of countries.

5. Scaling and Population Biology

Recently, Keitt and Stanley [34, 35] have applied to a 30-year data set on bird populations the same sort of techniques used to describe long-term data sets on economics and finance. They find statistical properties that are remarkably similar (Fig. 1), and consistent with the idea that "every bird species interacts with every other bird species," just as the economic analysis supports the notion that "every firm interacts with every other firm." This empirical result is not without interest, since it serves to cast doubt on models of bird population (and of economic systems) in which one partitions the entire data set into strongly-interacting and weakly-interacting subsets, and then ignores or oversimplifies the interactions in the weakly-interacting subset.

6. Discussion

Is the point of this paper just to show that a lot of different systems appear to develop scale-invariant correlations? If so, how do we understand this empirical fact?

Bak's idea that systems self-organize themselves such that they are in effect near a critical point is an appealing unifying principle. Near a critical point, there is a delicate balance between the exponentially-growing number of different one-dimensional paths connecting any two faraway subunits and the exponentially-decaying correlations along each one-dimensional path (this concept is illustrated, e.g., in Fig. 9.4 of Ref. [1]. If the control parameter (say coupling constant) is too small, the correlations die out so fast along each one-dimensional path that subunits far from one another are not well correlated. However, at a critical point, the exponentially-large number of paths connecting each pair of subunits is sufficient to balance out the exponential decay along each path and the "correction factor" wins out — this correction factor is the power law that governs the total number of one-dimensional paths connecting two distant subunits. The exponent in this correction factor depends primarily on the system dimension, and not at all on the actual arrangement of the subunits (lattice or no-lattice).

Could it be that somehow social systems push themselves "up to the limit" — just as a sandpile is pushed to the limit before an avalanche starts, an image that has attracted recent attention in the debate between "self-organized criticality" and "plain old criticality" (see, e.g., Vespignani and Zapperi [36] and references therein)? For example, in economics every subunit plays according to rules and pushes itself up against the limits imposed by these rules. But social systems display a variety of rich forms of "order", far richer than we anticipate from studies of ferromagnets and antiferromagnets (see, e.g., some of the papers appearing in Knobler et al. [37].

Could such orderings arise from the complex nature of the interactions? Or from the range of different "sizes" of the constituent subunits as, e.g., one finds ordering in sandpiles when sand particles of two different grain sizes are dropped onto a heap — see, e.g., the Refs. [38–42]? These are questions that occupy us now.

Acknowledgements

We are grateful to many individuals, and most especially to H. Leschhorn, P. Maass, R. Mantegna, M. A. Salinger, and M. H. R. Stanley for major contributions to those results reviewed here that represent collaborative research efforts. Partial support was provided by the National Science Foundation and the Israel-USA Binational Science Foundation, Israel Academy of Sciences.

References

1. H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford, Oxford University Press, 1971).
2. C.-K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons and H. E. Stanley, Nature **356**, 168 (1992).
3. A. Arneodo, E. Bacry, P. V. Graves and J. F. Mugy, Phys. Rev. Lett. **74**, 3293 (1995)
4. B. Suki, A.-L. Barabási, Z. Hantos, F. Peták and H. E. Stanley, Nature **368**, 615–618 (1994).
5. A.-L. Barabási. S. V. Buldyrev, H. E. Stanley and B. Suki, Phys. Rev. Lett. **76**, 2192 (1996).
6. C.-K. Peng, J. Mietus, J. M. Hausdorff, S. Havlin, H. E. Stanley and A. L. Goldberger, Phys. Rev. Lett. **70**, 1343 (1993).
7. C.-K. Peng, S. Havlin, H. E. Stanley and A. L. Goldberger, Chaos **5**, 82 (1995).
8. P. Ch. Ivanov, M. G. Rosenblum, C.-K. Peng, J. Mietus, S. Havlin, H. E. Stanley and A. L. Goldberger, Nature **383**, 323 (1996).
9. P. Ch. Ivanov, M. G. Rosenblum, C-K Peng, J. Mietus, S. Havlin, H. E. Stanley and A. L. Goldberger, Physica A **249**, 587–593 (1998).
10. H. A. Makse, S. Havlin and H. E. Stanley, Nature **377**, 608 (1995).
11. D. Zanette and S. Manrubia, Phys. Rev. Lett. **79**, 323 (1997).
12. P. Ball, *The Self-Made Tapestry: Pattern Formation in Nature* (Oxford University Press, Oxford, 1998).
13. B. B. Mandelbrot, J. Business **36**, 394 (1963).
14. R. N. Mantegna and H. E. Stanley, Nature **376**, 46 (1995); Nature (Scientific Correspondence) **383**, 587 (1996); Physica A **239**, 255 (1997); "Physics Investigation of Financial Markets," in *Proceedings of the International School of Physics "Enrico Fermi," Course CXXXIV*, edited by F. Mallamace and H. E. Stanley (IOS Press, Amsterdam, 1997).
15. M. H. R. Stanley, L. A. N. Amaral, S. V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M. A. Salinger and H. E. Stanley, Nature **379**, 804 (1996).
16. I. Peterson, Science News **150**, 104 (1996).
17. I. Peterson, *The Jungles of Randomness* (New York: Viking, 1997).
18. G. M. Viswanathan, V. Afanasyev, S. V. Buldyrev, E. J. Murphy, P. A. Prince and H. E. Stanley, Nature **381**, 413 (1996).

19. G. M. Viswanathan, S. V. Buldyrev, S. Havlin, M. G. da Luz, E. Raposo and H. E. Stanley, "Optimizing Success in Random Foraging: A Physical Approach," Phys. Rev. Lett. (submitted).
20. M. Batty and P. Longley, *Fractal Cities* (San Diego: Academic Press, 1994).
21. L. Benguigui and M. Daoud, *Geog. Analy.* **23**, 362 (1991).
22. L. Benguigui, Physica A **219**, 13 (1995).
23. C. Clark, J. R. Stat. Soc. (Series A) **114**, 490 (1951).
24. A. Coniglio, C. Nappi, L. Russo and F. Peruggi, J. Phys. A **10**, 205 (1977).
25. S. Prakash, S. Havlin, M. Schwartz and H. E. Stanley, Phys. Rev. A **46**, R1724 (1992).
26. P. Frankhauser, *La Fractalité des Structures Urbaines* (Collection Villes/Anthropos, Paris, 1994).
27. R. N. Mantegna and H. E. Stanley, Phys. Rev. Lett. **73**, 2946 (1994).
28. Y. Liu, P. Cizeau, M. Meyer, C.-K. Peng and H. E. Stanley, Physica **245**, 437 (1997).
29. P. Cizeau, Y. Liu, M. Meyer, C.-K. Peng and H. E. Stanley, Physica **245**, 441 (1997).
30. S. V. Buldyrev, L. A. N. Amaral, S. Havlin, H. Leschhorn, P. Maass, M. A. Salinger, H. E. Stanley and M. H. R. Stanley, J. Phys. I France **7**, 635 (1997); S. V. Buldyrev, H. Leschhorn, P. Maass, H. E. Stanley, M. H. R. Stanley, L. A. N. Amaral, S. Havlin, and M. A. Salinger, "Scaling Behavior in Economics: Empirical Results and Modeling of Company Growth," in *Proceedings of the International School of Physics "Enrico Fermi," Course CXXXIV*, edited by F. Mallamace and H. E. Stanley (IOS Press, Amsterdam, 1997).
31. S. V. Buldyrev, L. A. N. Amaral, S. Havlin, H. Leschhorn, P. Maass, M. A. Salinger, H. E. Stanley and M. H. R. Stanley, J. Phys. I France **7**, 635 (1997).
32. L. A. N. Amaral, S. V. Buldyrev, S. Havlin, M. A. Salinger and H. E. Stanley, Phys. Rev. Lett. **80**, 1385 (1998).
33. H. Takayasu and K. Okuyama, Fractals **6**, xxx (1998).
34. T. Keitt and H. E. Stanley, "Dynamics of North American Breeding Bird Populations," Nature **393**, xxx (May 1998).
35. T. Keitt and H. E. Stanley, "Lévy Flights and $1/f$ Noise in High Abundance Populations" (to be submitted).
36. A. Vespignani and S. Zapperi, Phys. Rev. Lett. **78**, 4793 (1997).
37. C. M. Knobler, A. Robledo and H. E. Stanley, *Statistical Mechanics in the Physical, Biological, and Social Sciences: Festschrift in Honor of Benjamin Widom on the occasion of his 70th Birthday* (Elsevier, Amsterdam, 1997) [Published as a special issue of Physica A **244** (1 Sept. 1997)].
38. H. A. Makse, P. Cizeau and H. E. Stanley, Phys. Rev. Lett. **78**, 3298 (1997).
39. H. A. Makse, S. Havlin, P. R. King and H. E. Stanley, Nature **386**, 379 (1997).
40. H. A. Makse, P. Cizeau, S. Havlin, P. R. King and H. E. Stanley, "Spontaneous Self-Stratification Without Shaking: 'Potatoes from Mashed Potatoes'" [Proc. 1997 NATO ASI] in *Granular Matter*, edited by H. J. Herrmann (Kluwer, Dordrecht, 1998).
41. H. A. Makse, P. Cizeau and H. E. Stanley, Physica A **249**, 391 (1998).
42. A. Károlyi, J. Kertész, H. Makse, H. E. Stanley and S. Havlin, "Cellular Automata Models for Granular Media" [Proc. 1997 NATO ASI] in *Granular Matter*, edited by H. J. Herrmann (Kluwer, Dordrecht, 1998).