THEORY -- OTHER

CANCER DETECTION VIA FRACTAL DIMENSION ANALYSIS

Wolfgang Bauer and Charles D. Mackenzie[a]

In the final diagnosis of cancer the pathologist generally relies on the qualitative and empirical parameters of cells biopsies. These experiential approaches have been aided by morphometric methods, such as determination of surface area, volume, axes ratios, estimation of population density, and other methods derived from basically Euclidian geometry [1].

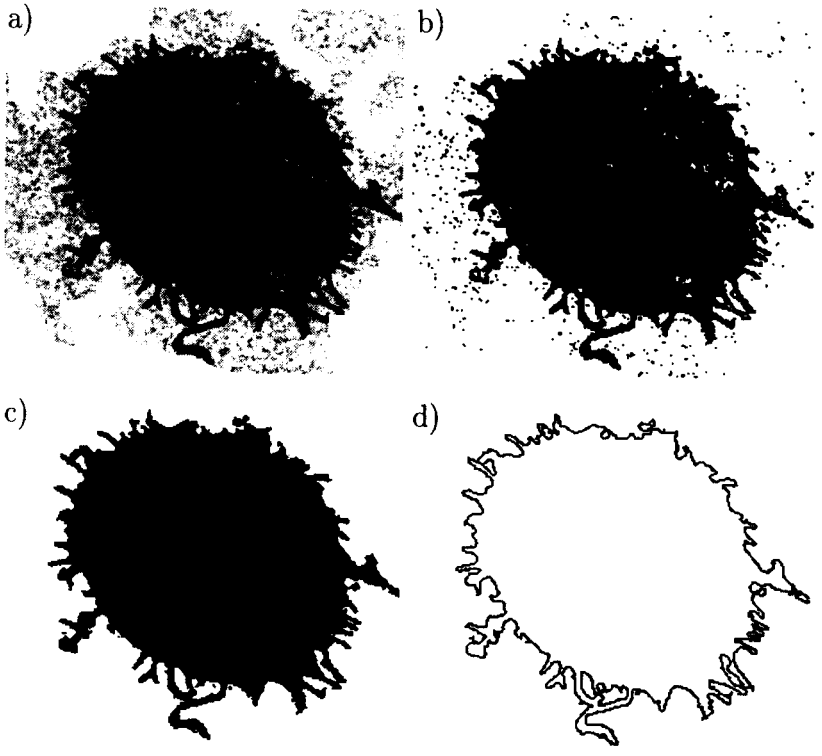

Figure 1: a) Electron microscope image of a section of a human lymphocyte, affected with hairy-cell leukemia, digitized with 256 grey levels. b) Black/white representation of (a) with a grey level threshold set at 80. c) Image after application of our small-cluster removal algorithm. d) Perimeter surface of the image, obtained by taking derivatives in x and y directions.

In mathematics and the physical sciences, it has been known since the turn of the century that classical Euclidian concepts, such as perimeter length, surface area, or volume, do not yield finite answers for certain objects.

Famous examples Koch's Snowflake with finite area and infinite perimeter surface length, and the coastal length of a country like Great Britain, for which the answer depends on the size of the yardstick one is using. In order to describe this kind of complexity found in nature, the concept of fractal geometry was developed [2].

In biology and medicine, fractals are just now beginning to have first applications [3]. This project is focussed on the use of fractals at the cellular level, specifically addressing the validity of comparing the fractal nature of the surface of cancer cells (lymphocytic leukemia cells) with that of normal cells of the same basic type. A clear distinction between the two types is observed. Thus this method can also be used

as a *quantitative* tool in the detection and diagnosis of cancers [4].

The method utilized is illustrated here by showing the results of the different steps on a human lymphocyte affected by hairy-cell leukemia. However, the method is general and can be used on many types of cells or organisms such as protozoa and bacteria.

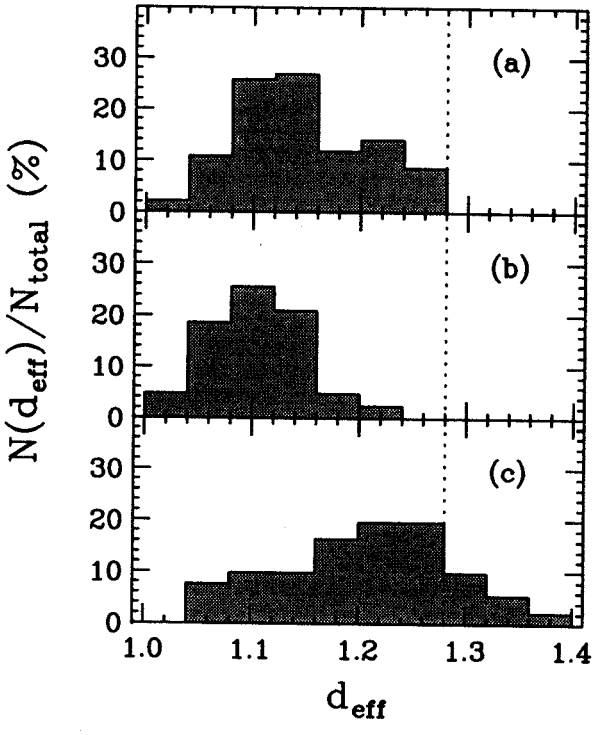

Figure 2: Histograms of fractal dimensions of sections of individual cells: a) White blood cells of healthy persons; (b) Lymphocytes of healthy persons; (c) Lymphocytes of hairy cell leukemia patients. The dotted vertical line indicates a fractal dimension of 1.28, a value which is surpassed only by cells from the cancer patients.

Fig. 1 shows the different steps in image processing leading from the digitized electron microscope image of the cell (a) via applying a threshold function (b) outline of the cell is obtained (Fig. 1b), speckle removal (c), and surface determination (d) to the fractal dimension of the surface.

Conventional speckle removal algorithms do not work here, because techniques like neighboring-pixel-averaging tend to smoothen the true cell surface as well, often resulting in a changed fractal dimension. To solve this problem one utilizes a cluster-recognition algorithm, developed for the purpose of identifying fragments in nuclear fragmentation [5]. This is yet another example where basic research in one field has lead to a completely unexpected applied spinoff in another.

The result of applying this method to cell samples is shown in Fig. 2. The number of cells with fractal dimension d are binned in histograms and displayed as a function of d. The vertical line indicates a fractal dimension of 1.28. None of the cells of the healthy patient has a fractal dimension larger than this value, whereas a large percentage of the cells of the cancer patient have $d > 1.28$. This clearly shows that it is possible to distinguish between healthy persons and hairy-cell leukemia cancer patients on the basis of the above method.

This method can be developed to serve as a diagnostic aid for the practicing pathologist or as a completely automated system for cancer detection on the cellular level.

a. Department of Pathology, Michigan State University

References

1. W.A. Aherne and M.S. Dunnill, *Morphometry* (Edward Arnold Publishers, London, 1982).
2. B. Mandelbrot, *The fractal geometry of nature* (W.H. Freeman, New York, 1977).
3. B.J. West, *Fractal Physiology and Chaos in Medicine* (World Scientific, Singapore, 1990); A. Bunde and S. Havlin, *Fractals in Science* (Springer, Heidelberg, 1994).
4. W. Bauer and C.D. Mackenzie, submitted to Phys. Rev. Lett.; W. Bauer and C.D. Mackenzie, "Method and System For Detection of Biological Materials", Michigan State University 4.1-128 (ID 94-030), US patent applied for.
5. W. Bauer, D.R. Dean, U. Mosel, and U. Post, Phys. Lett. 150B, 53 (1985); W. Bauer, Phys. Rev. C 38, 1927 (1988). T. Le Brun, H.G. Berry, S. Cheng, R.W. Dunford, H. Esbensen, D.S. Gemmell, E.P. Kantor, and W. Bauer, Phys. Rev. Lett. 72, 3965 (1994).

THE CONSTRAINT OPERATOR SOLUTION TO BOUNDARY VALUE PROBLEMS

David McGrew and Wolfgang Bauer

We solve the time-independent Schrödinger equation

$$\nabla^2\Psi(\vec{x}) + \lambda\Psi(\vec{x}) = 0 \tag{1}$$

for a free particle in a 2-dimensional region $\mathcal{I}$, an infinite well ($\Rightarrow \Psi(\vec{x})$=0 on the boundary surface $\delta\mathcal{I}$ of $\mathcal{I}$) of arbitrary shape (the Helmholtz equation with Dirichlet boundary conditions), a problem of interest in the area of Quantum Chaos. To do this, we first solve Eq. (1) in a rectangular region $\mathcal{T}$ that includes $\mathcal{I}$. There the eigenfunctions $\phi_k(\vec{x})$ are 'separable', i.e. products of simple sine-functions in each of the two coordinates. We then expand the solution $\Psi(\vec{x})$ for the region $\mathcal{I}$ in the basis of eigenstates for the region $\mathcal{T}$, $\Psi(\vec{x}) = \sum_k \psi_k \phi_k(\vec{x})$.

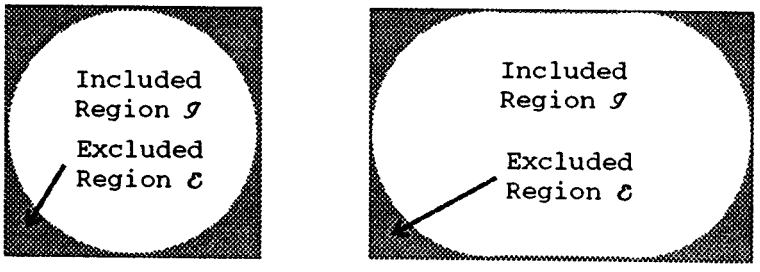

Figure 1: The included and excluded regions shown for the unit disk (left) and the stadium (right).

To obtain the expansion coefficients ψ_k we define the constraint operator C for $\vec{x}$ in $\mathcal{T}$, which multiplies functions on its right by the constraint function

$$c(\vec{x}) = \begin{cases} 1 & \text{if } \vec{x} \text{ is in } \mathcal{I} \\ 0 & \text{if } \vec{x} \text{ is in } \mathcal{E}. \end{cases} \tag{2}$$

The constraint operator is the projector for functions over the larger region $\mathcal{T}$ that are zero over the excluded region $\mathcal{E}$ (it 'constrains' functions to be zero in $\mathcal{E}$.) In the basis $\{\phi_i(\vec{x})\}$, C is represented as the matrix C with elements

$$C_{ij} = \int_{\mathcal{E}} d\vec{x}\, \phi_i(\vec{x})\phi_j(\vec{x}). \tag{3}$$

The solutions to the problem we are interested in are the eigenvectors of ∇^2 that are in the range of C. Intuitively, it seems that we can solve our problem by finding the eigenvectors of $C\nabla^2 C$. This is *almost* correct. To be completely correct, we derive the solutions using a Green's function [1] then simplify the result using our knowledge of the constraint operator. The resulting eigensystem is

$$(C\Lambda C - CSC - \lambda C)\psi \;=\; 0. \tag{4}$$

where

$$S_{ij} = \int_{\mathcal{E}} d\vec{x} \nabla\phi_i \cdot \nabla\phi_j \tag{5}$$

and Λ is the diagonal matrix containing the eigenvalues of ∇^2. We solve this by numerically integrating to find the matrix elements. The S term can be considered to be a correction that is needed because ∇^2 and C do not commute; its norm can be shown to be small. There are r non-trivial solutions to (4), where r is the rank of C.

Because the constraint operator is a projector, the constraint matrix for a region that is the union of two non-overlapping regions must be the sum of the constraint matrices for the two regions. When the included region, $\mathcal{I}$, is the total region, $\mathcal{T}$, the constraint matrix is the identity matrix, with rank m. If $\mathcal{I} \neq \mathcal{T}$, the rank r of C is approximately μm, where μ is the ratio of included to total volume.

To solve Eq. (4), we need to find an orthonormal basis for the range of C. As this matrix is only approximately a projector, its range is not well defined. We use the "closest" projector to C, in the Euclidean norm, that has the same rank as C. Make the unitary decomposition $C = U\Gamma U^\dagger$, where $U^\dagger U = 1$ and Γ is the diagonal matrix containing the eigenvalues C in decreasing order. The closest projector C' is found by setting the first r eigenvalues equal to one and the last $m - r$ eigenvalues equal to zero. In practice, $\Gamma_{ii} \simeq \gamma(i)$, where

$$\gamma(i) = \left[\exp\left(\frac{i-r}{Tm}\right) + 1\right]^{-\frac{1}{2}}, \tag{6}$$

which is the square root of the Fermi function (compare with Fig. 2a.) As $T \to 0$, we recover the expected spectrum of a projector. By expanding $\gamma(i)$ near r, the average square eigenvalue error $m^{-1}\|C - C'\|_2$ is found to be $2^{3/2}(1 - 2^{-1/2})^2 T$.

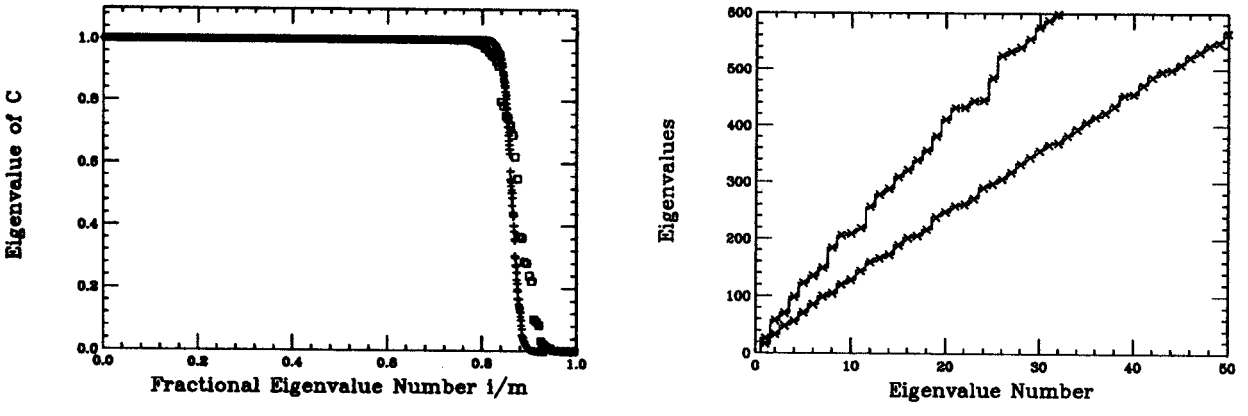

Figure 2: a) The spectrum of C for the stadium for m = 200 and m = 1600, with $r/m = \mu = 0.8570$. b) Eigenvalues for the disk (upper curve) and stadium (lower curve). The histograms represent the analytically correct solution for the disk and the solution of the stadium problem using Heller's method. The crosses represent our method in both cases.

The matrix C is well approximated by a band diagonal matrix in the basis of ordered energy

eigenstates. The rms eigenalue error of a band diagonal truncation of C can be bounded using the Weiland - Hoffman theorem [2]; the bound approaches 0 proportional to $b^{-1/2}$, where b is the bandwidth.

We are currently pursuing the approach of expanding the constraint function in eigenfunctions. The matrix elements of C and S are completely determined by the elements of this expansion. The band diagonality of C and S are manifest in this approach. The elements of the constraint function expansion can be computed by numerical integrals over the boundary $\delta\mathcal{I}$, resulting in an efficient computational algorithm.

A conventional method for solving quantum well problems is one devised by Heller [3] that uses a basis of plane waves to express the wavefunctions, then demands that the wavefunctions be zero on some points evenly spaced around the boundary surface. The eigenfunctions are found by an iterative process. An eigenvalue is assumed, the linear equations that set the wavefunctions equal to zero at the points are solved, then the error in the boundary condition is evaluated. The process is repeated until the error in the boundary condition vanishes; the wavefunction is then an eigenfunction of ∇^2 that satisfies the necessary boundary conditions. We expect the constraint operator method to be computationally superior, as it solves for all wavefunctions simultaneously. Fig. 2b shows comparisons of our new method (crosses) with the analytic solution for the unit disk (upper histogram) and the numerical solution for the stadium using Heller's method (lower histogram.)

References

1. P. M. Morse and H. Feshbach *Methods of Theoretical Physics* (1953) McGraw - Hill.
2. R. A. Horn and C. R. Johnson *Matrix Analysis* (1985) Cambridge University Press.
3. E. Heller, P. O'Connor and J. Gehlen, Physica Scripta. Vol. 40, 354-359, 1989.

SCALING OF LARGE FLUCTUATIONS IN FINITE SYSTEMS

Aldo Bonasera[a] and Jürgen Schulte

The scaling of critical behavior in molecular and nuclear collisions has been studied employing results from classical and quantum Molecular Dynamics simulations.In order to show the scaling of critical behavior we have analyzed the mass fragmentation data of the $C^{+}_{60} \pm C_{60}$ collisions at fixed impact parameters, and performed collisions of $^{60}X + {}^{60}X$ nuclear clusters. The collision energy and impact parameters have been scaled accordingly to ratios of the clusters' collision energy and binding energy, and impact parameter and cluster radius, respectively. The critical behavior of the colliding systems is shown in terms of the scaled factorial moments of the cluster mass distribution and the corresponding Campi scatter plot.

For the interaction of carbon atoms a hybrid of an empirical potential and the density functional linear combination of atomic orbitals in the local density approximation regime [1] has been employed. The nuclear cluster collision has been modeled by classical Molecular Dynamics (CMD) [2]. In such an approach the constituent particles ("neutrons and protons") interact through two body Yukawa potentials. Details of the model interaction may be found in Ref. [3]. The hypothesis of this work is that if a finite system retains some features of a phase transition, then the properties of the system near the critical point are independent of the details of the force, and thus different systems may be compared when being scaled appropriately. Thus, we compare our nuclear and molecular models by scaling energies and impact parameters by their binding energies and radii, respectively: E_r=$E_{c.m.}$/ B.E.; b_r=b/(2R). The reduced energy E_r and length scale b_r are defined by the given data set of the molecular cluster collision [1]. We expect similar collision properties for the nuclear cluster collision at E_r = 0.45. The results of collision simulations may be discussed in terms of the Fisher's droplet model. In such approach the mass yield is given by

$$Y(A) = Y_o \exp\left[\frac{\mu_g-\mu_l}{T}A - \frac{4\pi r_o \sigma}{T}A^{2/3} - \tau \ln A\right] = Y_o A^{-\tau} Z^A X^{A^{2/3}} \qquad (1)$$

where μ_g and μ_l are the gas and liquid chemical potential, respectively, σ is the surface tension of the liquid (σ=0 for T>T_c), and τ=2.23 is the curvature term. At T=T_c the surface (X) and volume (Z) term are equal to one, and only the curvature term contributes. The least square fit of Eq. (1) to both model results shows that at central collisions both systems are very close to the critical temperature (X_{mol}≈ Z_{mol}≈ X_{nucl}≈Z_{nucl}≈1). In order to verify the scaling assumption on an event by event basis we performed a Campi analysis [4] of our data. In the Campi analysis we have plotted the logarithm of the largest fragment (P) versus the logarithm of S_k with k=2, and

$$S_k = M_k / M_1, \qquad M_k = \sum_A A^k N(A, T-T_c) \qquad (2)$$

M_k being the moments of the cluster size distribution, and A and N(A,T-T_c) the cluster mass (size) and multiplicity of cluster size A at fixed distance T-T_c from the critical point. The sum in Eq. (2) is performed over all the fragments but the largest one (P). The Campi plot gives us information on the size fluctuations of the largest fragments on an event by event basis. Near the critical point such fluctuations should be at maximum, i.e., resulting in a large range of S_2 values. In Fig. 1 we have

plotted the fragmentation events of both collision systems for all studied impact parameters with fixed energy E_r= 0.45. The overlap of fragmentation events and respective fluctuations is astonishing well.

In conclusion, we have shown that two completely different finite size systems, a classical one (which has some features of nuclei) in the MeV energy and fm range can be compared to a quantum system (with some molecular features) in the eV and Å range. The comparison has done by appropriately scaling the collision energy and distances by the binding energies and radii of the two systems. This procedure is similar to what is obtained in infinite systems where universal curves can be obtained in the vicinity of the critical point for a liquid gas phase transition. Very large fluctuations are seen for both systems at the same reduced energy, E_r =0.45. If the features discussed in this work will be confirmed by experimental observations both in heavy ion collisions and fullerenes collisions, one may hope to be able to obtain universal curves also for microscopic systems containing as little as 100 particles.

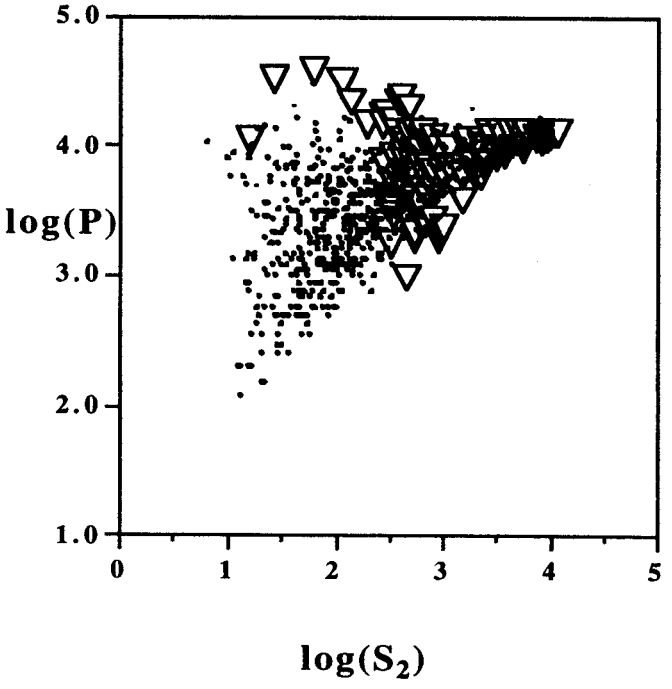

Figure 1: Campi plots at fixed reduced energy E_r=0.45 including all impact parameters considered in the QMD (triangles) and CMD (dots) simulations. Note that there is a dense population of CMD events for S_2>2.5 that are not visible in this plot due to overlapping symbols of the QMD events.

a. Lab. Nazionale del Sud-Istituto Nazionale di Fisica Nucleare, v.S.Sofia 44, 95123 Catania, Italy.

References

1. Phys. Rev. **B 51** (6), Feb. (1995), and references therein.
2. V. Latora, M. Belkacem, A. Bonasera, Phys. Rev. Lett **73**, 1765 (1994);Phys. Rev. **C** (in press).
3. R.J. Lenk, T.J. Schlagel, V.R. Pandharipande, Phys. Rev. **C42**, 372 (1990).
4. X. Campi, J. of Phys. **A19**, L917 (1986); Phys. Lett. **B208**, 351 (1988); J. de Phys. **50**, 183(1989).

BUCKYBALL FRAGMENTATION

Wolfgang Bauer

One of the main reasons for studying nuclear fragmentation in the excitation energy range between 5 and 50 MeV per nucleon is the possibility of nuclei undergoing a phase transition from a liquid-like to a gas-like state of aggregation. Many phenomenological nuclear equations of state have been studied, and very simple phase transition models have been able to reproduce currently available experimental observables, thus strengthening our expectation that a phase transition is a real possibility.

There is another way, however, to study similar finite size effects on the fragmentation phase transition for a different microscopic interaction, the fragmentation of small atomic clusters. In the present work, a group at Argonne National Laboratory bombarded buckyballs (C-60 molecules in the shape of a spherical shell with soccer-ball geometry) with Xe^{35+}-ions with a beam energy between 420 and 625 MeV[1].

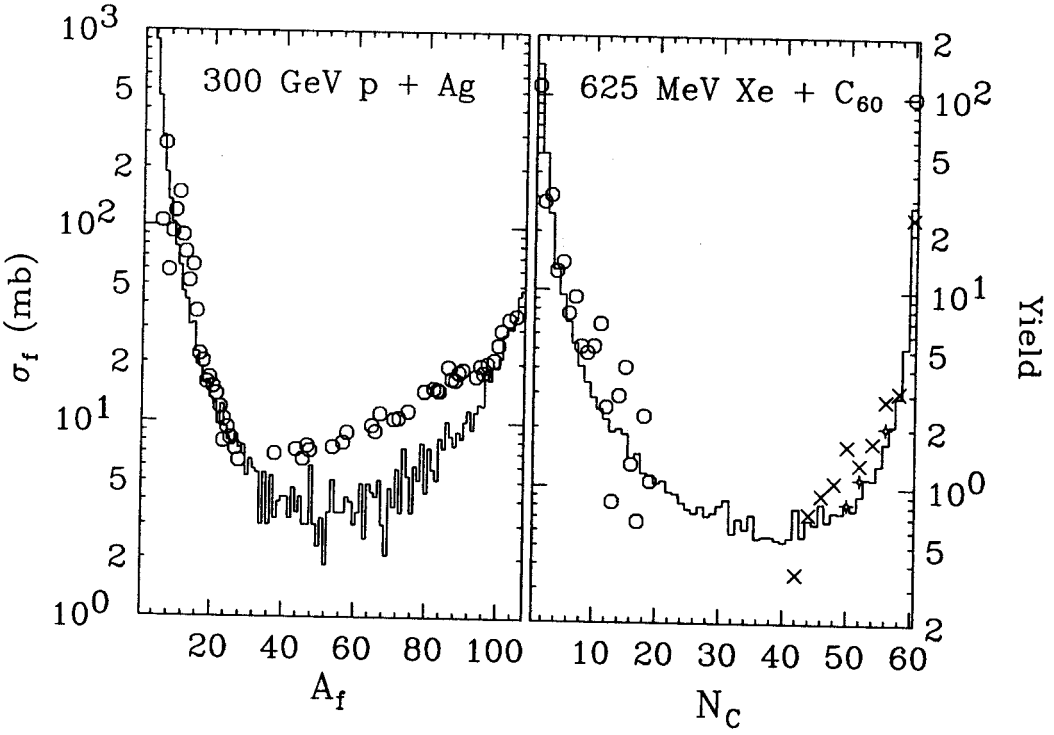

Figure 1: Comparison of experimental mass yield data (circles) for high-energy proton induced nuclear (left) and high-energy heavy-ion induced buckyball (right) fragmentation to percolation models (histograms) of fragmentation. The data are taken from references 1 and 3.

We have constructed a percolation-based model to calculate the expected fragmentation pattern in a phase transition model. It's ingredients are similar to a percolation model which has been successful in reproducing experimental fragment mass distributions generated in nuclear fragmentation experiments [2]. The only differences are the lattice geometry (3-d spheres for nuclear fragmentation, and spherical shells for buckyballs), and the impact parameter dependence of the bond-breaking probability. For nuclear fragmentation induced by high-energy protons we used a Glauber approximation, and for the buckyball

case we used

$$p(b) = \frac{p_0}{2\sqrt{\pi}\,\omega\,\Delta R} \int dr_\perp \left\{ \sqrt{[(R+\Delta R)^2 - r_\perp^2]_+} - \sqrt{[(R-\Delta R)^2 - r_\perp^2]_+} \right\} \exp(-(b-r_\perp)^2/\omega^2)$$

where $R \approx 3.5$ Å, $\Delta R \approx 0.35$ Å, and with the notation $[a]_+ = a \; \forall \; a > 0$, and 0 otherwise.

Figure 1 shows a comparison of our percolation model to nuclear fragmentation data for the reaction 300 GeV p + Ag (circles [3], left side) and to buckyball fragmentation (circles [1], right side). There is an astonishing number of similarities. Both datasets (and calculations) exhibit a typical U-shape. Both datasets show a power-law dependence of the fragment mass yield for low-to-intermediate mass fragments – the power is approximately -2.6 for the nuclear case and -1.3 for the buckyballs. And both datasets show fine structure in the low-to-intermediate mass yields caused by binding energy effects.

From this apparent similarity we can hope that the study of finite-size modifications to the signatures of liquid-gas type phase transitions will enter a new phase of interesting cross-fertilizations between nuclear and atomic/molecular physics.

References

1. T. Le Brun, H.G. Berry, S. Cheng, R.W. Dunford, H. Esbensen, D.S. Gemmell, E.P. Kantor, and W. Bauer, Phys. Rev. Lett. 72, 3965 (1994).
2. W. Bauer, D.R. Dean, U. Mosel, and U. Post, Phys. Lett. 150B, 53 (1985); W. Bauer, Phys. Rev. C 38, 1927 (1988).
3. A. Bujak et al., Phys. Rev. C 32, 620 (1985).

WORD PROCESSORS, AVALANCHES, RANDOM WALKS AND SELF-ORGANIZED CRITICALITY

W. Bauer and S. Pratt

For many of us, the most commonly faced example of self-organized criticality (SOC) is the word wrap feature in our text processors. Consider a very long paragraph, composed of lines of no more than N_c characters. When a line becomes too long, it pushes the last word into the next line. If this push makes the next line too long, the cascade will continue. Three fundamental questions come to mind. First, what is the probability $P(n)$ that a cascade lasts n lines, and for large n does $P(n)$ approach a simple analytic form? Secondly, how is the system affected by the average word length and the average line length? One knows that after at most N_c characters are entered, a large cascade will occur, increasing the length of the paragraph by one line - even if the paragraph is infinitely long. Finally, is the occurence of large cascades or avalanches regular or predictable?

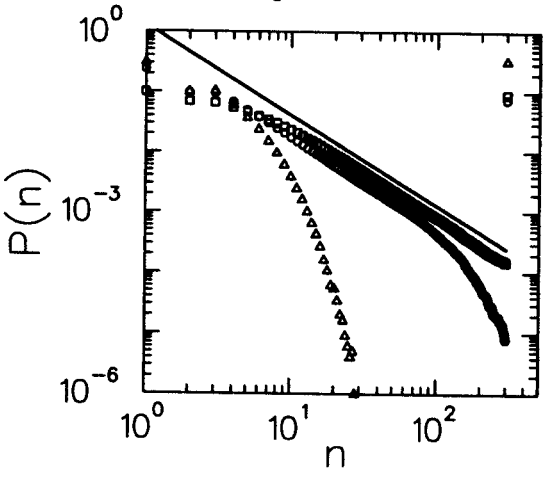

Fig. 1 The probability that a cascade died at the n^{th} line is plotted for word lengths of 2 (squares), 8(circles), and 32 (diamonds), given a line length of 80 characters. For short word lengths the distribution approaches the $n^{-3/2}$ power illustrated by the solid line.

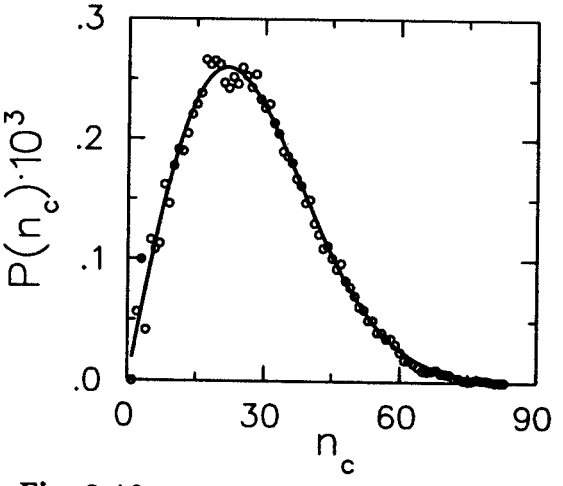

Fig. 2 After a cascade of 300 lines or longer, the chance of the next cascade appearing after n_c characters have been added to the paragraph is illustrated with circles. The solid line shows $n\exp(-n^2/n_0)$ where n_0 is chosen to fit the results.

To model the word processor, we enter words of random length into the beginning of the file. Word lengths are chosen randomly from one to a chosen maximum length. Other word-length distributions, i.e. Poissonian, show essentially identical results. Whenever the length of a line exceeds N_c, the last word of that line is pushed into the next line, and if the line is still too long another word is pushed forward. The avalanche is succesively calculated until it dies at some line n. The probability of the cascade dying at n is shown in Fig. 1. For this example N_c = 80 and the word lengths used were l_w = 2, 8, and 32.

The power-law behavior for small word size in Fig. 1 can be explained by mapping it onto a random walk. Let the number of characters pushed into the n^{th} line be k_n. The difference between k_n and k_{n+1} is small and random. For a random walker who is a distance x from the origin after time t, one can map n to t and k_n to x. Since the cascade dies at the first line n where k_n is zero, the random walk must have a boundary, ending the walk whenever $x = 0$. A random walk can be described with the diffusion equation,

$$\frac{\partial f(x,t)}{\partial t} = \gamma \frac{\partial^2 f(x,t)}{\partial x^2}, \tag{1}$$

where $f(x,t)$ is the probability that the walker is at x at step t. After a long time the solution to such an equation where the walkers die when they reach $x = 0$ is

$$f(x,t) \propto \frac{x}{t^{3/2}} \exp\left\{-\frac{x^2}{4\gamma t}\right\}. \tag{2}$$

The flux into the boundary, which is proportional to the derivative of f with respect to x, is the probability of annihilating at time t and scales as $t^{-3/2}$.

If the number of characters wrapped around a line exceeds N_c an infinite cascade ensues. The random walker's analogy is that if he would reach some critical value of x he would live eternally. A long word size corresponds to large step size for the walker, or to a large diffusivity, which increases the chance of drifting into eternity and destroys the scaling behavior. When the word length is much shorter than the line length, the equation maps onto the two-dimensional directed SOC model of Dhar and Ramaswamy [1], which is a variation of the picture of Bak, Tung and Wiesenfeld [2].

One aspect of the our SOC word processor which can not be described by a random walk is the regularity and predictability of large cascades. Letting n_c represent the number of characters entered into the paragraph between succesive cascades of 100 lines or greater, Fig. 2 shows the distribution of n_c. From inspecting Fig. 2, one can see that if a large avalanche has recently occured, there is diminished probability of another large cascade during entering the next few characters. Similar behavior results from considering the average lifetime of an object whose chance of dying in the next time step increases linearly with time,

$$\frac{\partial w(t)}{\partial t} = -\beta t w(t), \tag{3}$$

where $w(t)$ is the probability that the object is still alive. The solution is:

$$w(t) \propto \exp\left\{-\beta \frac{t^2}{2}\right\}. \tag{4}$$

The probability of dying at a time t is then:

$$P(t) = \frac{\partial w}{\partial t} \propto t \exp\left\{-\beta \frac{t^2}{2}\right\}. \tag{5}$$

This form fits well to the results in Fig. 2 which was done with a maximum word length of two characters. If the threshhold for a large cascade is set very high or if line lengths are short, the behavior changes and large cascades become extremely regular.

References

1. D. Dhar and R. Ramaswamy, Phys. Rev. Lett. 63, 1659 (1989).
2. P. Bak, C. Tang and K. Wiesenfeld, Phys. Rev. Lett. 59, 381 (1987).