Nearest Neighbor Distances in Circle Fractals. Statistics

In studies of tilings there is much interest in the nearest-neighbor arrangements. The nearest-neighbor distances are generally fixed values, the same for each tile. For statistical geometry fractals the nearest-neighbor distance information is of a statistical nature.

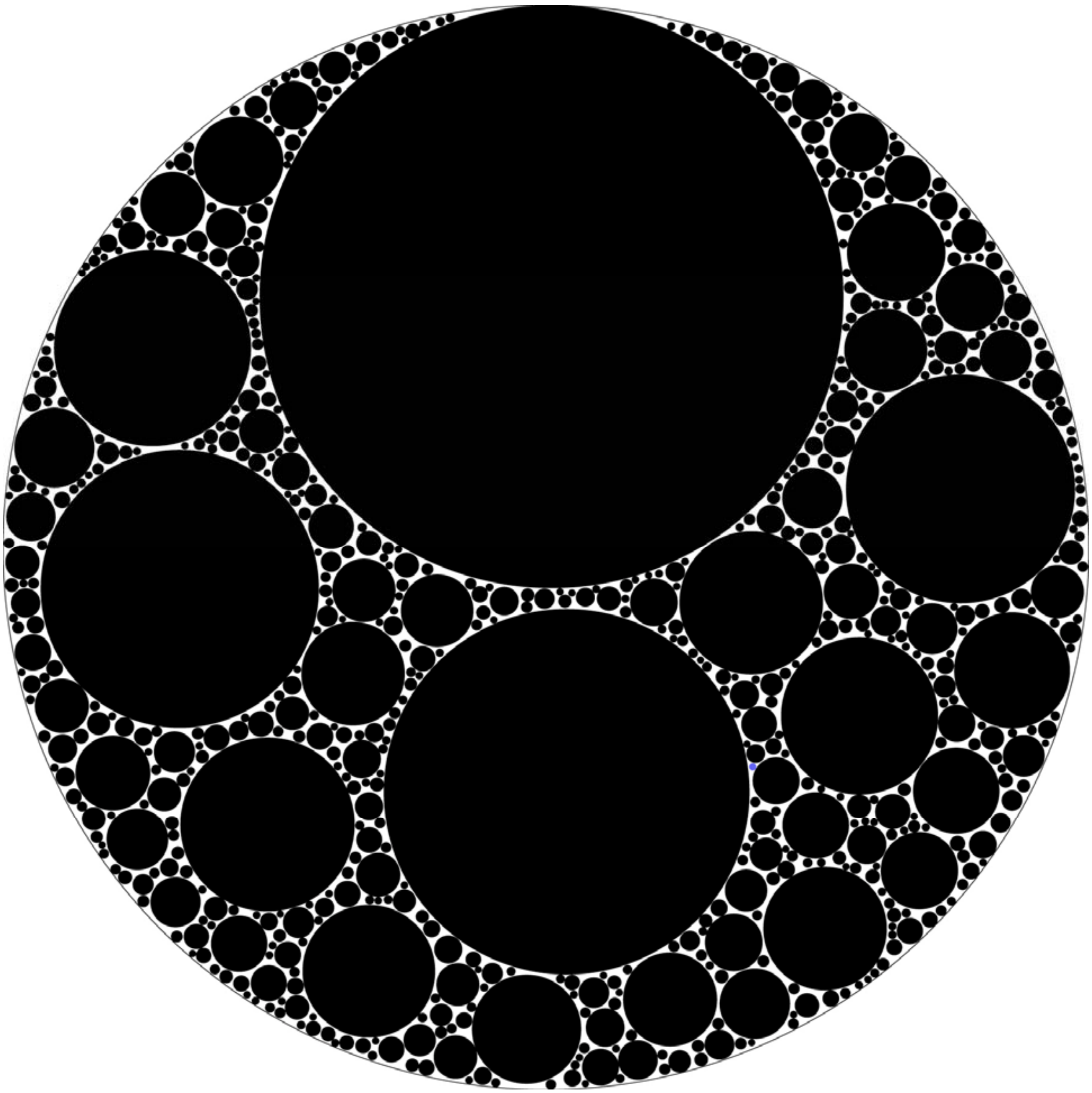

Fig. 1. A circles-in-circle random fractal pattern. The black circles are the 700 circles placed by the algorithm. (c=1.35, N=1, 91.7% fill). A circular line shows the outer boundary.

The algorithm is described in Appendix A and in the references.

1. Introduction. Methods.

The extraction of nearest-neighbor distances from a pattern of this kind requires some definitions. The nearest neighbor distance is taken to be the width of the empty[1] space between circles. Not only are there "nearest" neighbors, but there are non-nearest neighbors that are not relevant to the problem. In the present study *all* of the circle-circle distances were computed and histograms were then made of the population of such distances at the short end. All of the lengths here are in "units", with the radius of the bounding circle being 5 units.

The "cutoff spacing" d_{cut} for the histograms is somewhat arbitrary, although it can be argued that the radius r_{last} of the last-placed circle should be an upper limit for it. Since all of the circle-circle distances are computed, the larger ones will refer to non-nearest neighbors.

The histograms consistently showed that the probability distribution function (p.d.f.) $p(d)$ for short circle-circle distances d is well fitted by a decreasing exponential function.

$$p(d) = p_0 e^{-d/d_0}$$

The details of the curve-fitting method can be found in Appendix B. In all cases N=1.

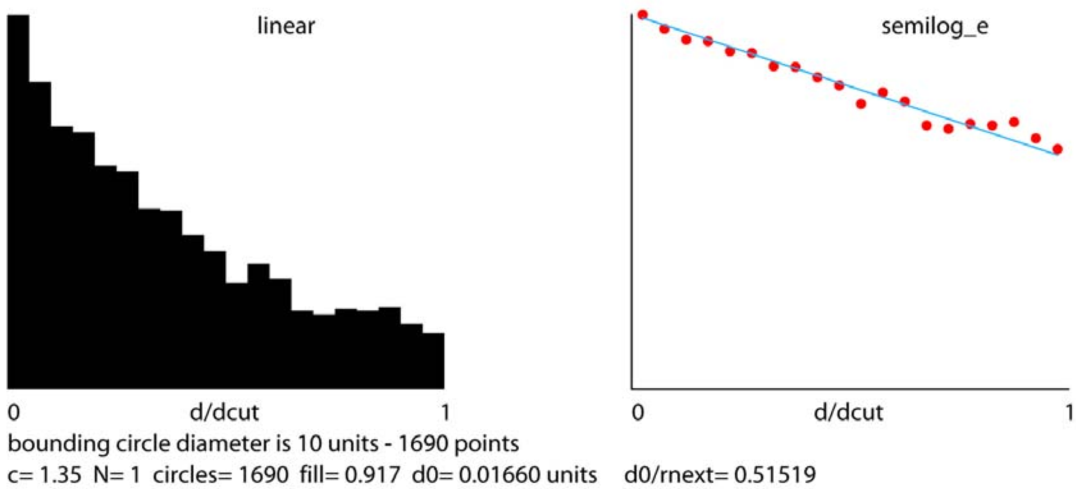

Fig. 2. Typical example of a histogram and the fitted exponential p.d.f. $p(d)$ for the circle fractal of Fig. 1. The raw data is at the left, and a semilog plot of $log_e(n_k)$ versus d_k is at the right. The data cutoff d_{cut} was chosen so that the last data "box" had about e^{-2} times as many data points as the first. The parameter r_{next} is the radius of the next-to-be-placed circle.

2. The Distribution of Nearest-Neighbor Distances.

(a) The distribution width d_0 versus number n of circles placed. It was found that d_0 and the last-placed radius r_{last} track very closely

[1] It is calculated by finding the distance between the centers of the two circles and subtracting the radii r_1 and r_2.

$$d_0 / r_{last} = const. \qquad \text{Eq. (1)}$$

This is to be expected based on the idea that these fractals possess "statistical self-similarity".

(b) The distribution width d_0 versus the exponent c. This is a more complicated question. In view of Eq. (1) it makes sense to compare the values of d_0 with a fixed value of r_{last}. This was done, with the results shown in the table. Each run was repeated 5 times and the standard deviations reflect the scatter in these values. The minimum radius r_{last} is .02 units here. Figure 3 shows a semilog graph of this data. The larger relative error for large c is related in part to the smaller number of circle distances in the data for those cases. It can be seen that d_0 falls very rapidly, almost exponentially, with increasing c.

c=1.25	d_0=.04100±.00041 units
c=1.3	d_0=.01872±.00045
c=1.35	d_0=.01066±.00022
c=1.4	d_0=.00585±.00030
c=1.45	d_0=.00309±.00013

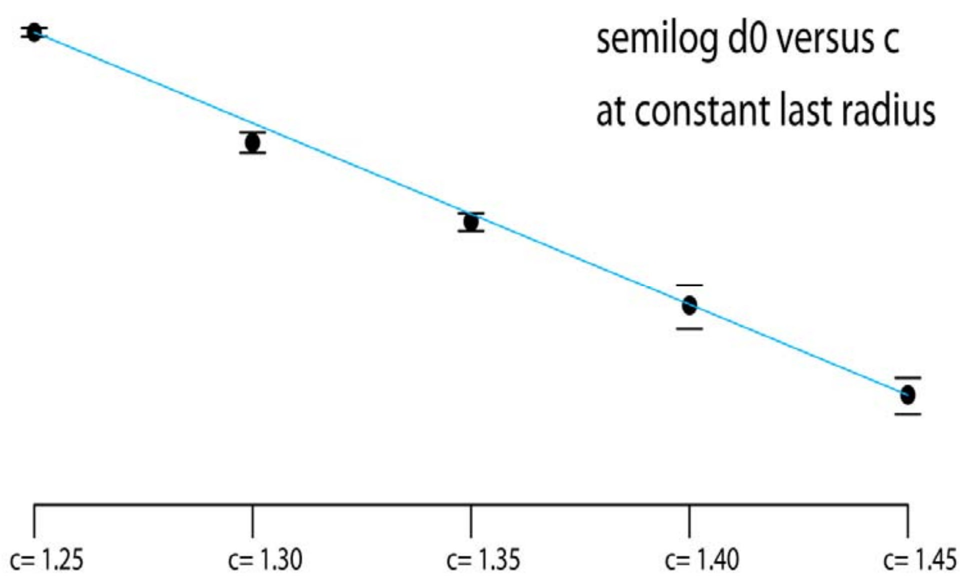

Fig. 3. A semilog plot of d_0 versus the exponent c. All of the runs were carried to the same last radius (.02 units). The error bars are 3-σ. Within the statistical scatter this relationship appears to be well approximated by a straight line (blue line), with a small amount of "sag" below the line which may or may not be significant.

3. Discussion of the Results.

It can be hoped that in the long run it will be possible to derive analytical expressions for the relationships seen here.

It is not obvious to me why the probability distribution function (p.d.f.) for *d* should be a decaying exponential function.

The rapid decrease of d_0 with increasing *c* shows that the nearest-neighbor distances fall much more rapidly than the smallest circle size as *c* increases, which can be expected. It reflects the general "tightening" of the packing seen with high *c*. Why the relationship between the characteristic length d_0 of this exponential p.d.f. and *c* should be another approximate decaying exponential in turn isn't obvious[2] to me.

4. References.

[1] Statistical geometry article on the web site john-art.com.

[2] John Shier, "Hyperseeing", Summer 2011 issue, pp. 131-140, published by ISAMA. Available by download at the web site john-art.com.

[3] "Statistical Geometry", John Shier, July 2011. A colorful self-published fractal art picture book available at lulu.com.

[4] Unpublished article "The Dimensionless Gasket Width b(c,n) in Statistical Geometry", John Shier, July 2011. It can be found at the web site john-art.com.

John Shier

August 2011

Appendix A. The Statistical Geometry Algorithm.

It has been found [1]-[3] that it is possible to create fractal patterns of a wide variety of geometric shapes by the following algorithm:

1. Create a sequence of areas A_i equal to $\frac{1}{N^c}, \frac{1}{(N+1)^c}, \frac{1}{(N+2)^c}, \frac{1}{(N+3)^c}, \ldots$. Choose an area (square, rectangle, circle, …) *A* to be filled.

2. Sum the areas A_i to infinity, using the Hurwitz zeta[3] function

[2] I know that authors of reports of this kind are expected to have some kind of derivation or authoritatively asserted theory for these things. Honesty, however, compels the author to admit that he doesn't know.

[3] The definitions of the Riemann and Hurwitz zeta functions can be found in Wikipedia.

$$\zeta(c,N) = \sum_{i=0}^{\infty} \frac{1}{(N+i)^c}$$

3. Define a new set of areas S_i by $S_i = \frac{A}{\zeta(c,N)(N+i)^c}$. It will be seen that the sum of all these redefined areas is just *A*.

4. Let $i = 0$. Place a shape having area S_0 in the area *A* at a random position *x,y* such that it falls entirely within area *A*. Increment *i*. This is the "initial placement".

5. Place a shape having area S_i entirely within *A* at a random position *x,y* such that it falls entirely within A. If this shape overlaps with any previously-placed shape repeat step 5. This is a "trial".

6. If this shape does not overlap with any previously-placed shape put *x,y* and the shape dimensions in the "placed shapes" data base, increment *i*, and go to step 5. This is a "placement".

7. Stop when *i* reaches a set number, percentage filled area reaches a set value, or other.

One will note that the dimensions of the shapes are nowhere specified. They are calculated from the areas. A very wide variety of shapes have been found to be "fractalizable" in this way.

The parameters c and N can have a variety of values. The parameter c is often in the range 1.2-1.4 with a largest usable value around 1.51. N can be 1 or larger, and need not be an integer.

By construction the result is a space-filling random fractal -- if the process never halts. Available evidence [1]-[3] says that it does not. The power law area sequence ensures that it has the fractal "statistical self-similarity" (scale-free) property. And the random search ensures that no two circles will ever touch, so that the "gasket" is a single continuous object.

Appendix B. Data Fitting.

The data to be analyzed were the numbers n_k of distances *d* within "boxes" centered at distances d_k. The number of boxes was arbitrarily chosen to be 20. The sum-of-squares error ε was defined to be

$$\varepsilon = \sum_{k=1}^{20} w_k \left[\log_e(n_k) - (ad_k + b)\right]^2$$

Where n_k are the numbers of instances in the histogram boxes. The weights w_k were chosen to be $1/n_k$. The use of this weighting scheme emphasized the more accurate small-*d* data. The parameters *a* and *b* are the slope and intercept of the best straight line and were determined by the usual error minimization method.