THE RANDOM DIVISION OF FACES IN A PLANAR GRAPH

by Richard Cowan and Simone Chen

Paper published in *Adv. Appl. Prob.* **28**, 377–383, (1996)

Abstract

A planar graph contains faces which can be classified into types depending on the number of edges on the face boundaries. Under various natural rules for randomly dividing faces by the addition of new edges, we investigate the limiting distribution of face type as the number of divisions increases.

Introduction

Consider a connected planar graph having initially $\mathcal{F}$ faces, $\mathcal{E}$ edges and $\mathcal{V}$ vertices. Here, 'face' excludes the unbounded region and so, from Euler's formula, $\mathcal{V} + \mathcal{F} = \mathcal{E} + 1$. A face is said to be of type k, or is called a k-face, if the boundary of that face contains k edges. A journey around that boundary also has k encounters with vertices, though these may not in general be k distinct vertices of the graph (if loops exist at some vertices). We refer to these k encounters as the k 'ports' of that face.

Suppose that all faces are sub-divided at time $n = 1$, each by a newly-formed 'random' edge joining two of the ports of a face. The random rule for a k-face is that two ports are sampled in an equally-likely manner and *with replacement* from the list of k ports. The two choices are then joined by the new edge; if the same port is sampled twice then the new edge is a loop. This random sub-division takes place independently in all faces.

At subsequent times $n = 2, 3, \cdots$ the sub-division of all extant faces continues by the same rules applied independently at each time. Thus for $n \geq 0$, the number $\mathcal{F}_n$ of faces at time n is $2^n \mathcal{F}$, the number $\mathcal{V}_n$ of vertices is unchanged at $\mathcal{V}$, whilst the number $\mathcal{E}_n$ of edges is given by, $\mathcal{E}_0 = \mathcal{E}$, and for $n \geq 1$, $\mathcal{E}_n = \mathcal{E}_{n-1} + \mathcal{F}_{n-1}$. Thus $\mathcal{E}_n = \mathcal{E}_0 + \mathcal{F}_0 + \mathcal{F}_1 + \cdots + \mathcal{F}_{n-1} = \mathcal{E} + \mathcal{F}(2^n - 1)$.

A randomly-sampled face at time n will have mean *type* of $(2\mathcal{E}_n - \mathcal{B}_0)/\mathcal{F}_n$ where $\mathcal{B}_0$ is the initial number of edges on the boundary of the graph. As $n \to \infty$ the mean type converges to 2. The type, X_n say, is a random variable distributed on the positive integers. We investigate the limiting distribution of this random variable and find that X_n converges in distribution to a random variable X with

$$p_x \equiv \mathbb{P}\{X = x\} = \frac{(2x-1)}{(x-1)!2^x\sqrt{e}} \qquad x = 1, 2, \cdots . \tag{1}$$

Such a result, combined with other related findings explored here, provides an interesting complement to earlier results on the random splitting of polygonal mosaics (Cowan and Morris 1988; Cowan, 1989; Cowan and Chen, 1996). In those studies, the edges were straight lines and the faces polygons.

Analysis

Let $Y_n(x)$ be the number of faces of type x at time n. Given $Y_n(x)$, the chance that a randomly sampled face at time n is an x-face is $Y_n(x)/2^n\mathcal{F}$, whilst unconditionally this chance is $\mathbb{E}Y_n(x)/2^n\mathcal{F}$. Indeed the population of faces forms a multi-type branching process, with 'time' n interpreted as 'generation'. Let $\boldsymbol{m}_n$ be a row vector whose elements are $\mathbb{E}Y_n(x)$, $x = 1, 2, \cdots$. The stochastic division rules, combined with elementary ideas from branching-process theory, lead to

$$\boldsymbol{m}_n = \boldsymbol{m}_{n-1}\boldsymbol{M} = \cdots = \boldsymbol{m}_0\boldsymbol{M}^n \tag{2}$$

where

$$\boldsymbol{M} = \begin{pmatrix} 1 & 1 & & & & \\ \frac{1}{2} & \frac{2}{2} & \frac{1}{2} & & & \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} & \frac{1}{3} & & \\ \frac{1}{4} & \frac{2}{4} & \frac{2}{4} & \frac{2}{4} & \frac{1}{4} & \\ \vdots & & & & & \ddots \end{pmatrix}.$$

Here the (i,j) element, which is the mean number of j-type 'daughter' faces which an i-type 'mother' face produces, is $1/i$ if $j = 1$ or $i+1$, $2/i$ if $2 \leq j \leq i$ and zero otherwise. Define $\boldsymbol{p}_n$ as $\boldsymbol{m}_n/2^n\mathcal{F}$; thus the elements of $\boldsymbol{p}_n$ are the chances of sampling an x-type face in generation n. From (2), $\boldsymbol{p}_n = \boldsymbol{p}_{n-1}\boldsymbol{M}/2 = \cdots = \boldsymbol{p}_0(\boldsymbol{M}/2)^n$.

We are concerned with the convergence of $\boldsymbol{p}_n$ and thus of $(\boldsymbol{M}/2)^n$. If a limit exists then it provides the limiting distribution of X_n. Noting that $\boldsymbol{M}/2$ is a transition matrix, we see a direct link with the theory of Markov chains by considering the chain that arises when, after each face is divided, one of the two daughter faces is randomly selected and discarded whilst the other is retained and further divided. The 'type' of retained faces is an irreducible Markov chain with transition matrix $\boldsymbol{M}/2$. Issues of convergence are as for our original problem.

If the equation $\boldsymbol{p} = \boldsymbol{p}\boldsymbol{M}/2$ has a solution $\boldsymbol{p}$ whose elements sum to 1, then this establishes that the chain is ergodic, that the chain's state vector $\boldsymbol{p}_n$ converges and that $\boldsymbol{p}$ is the limit (Cox and Miller (1965) §3.8). Direct calculation from $\boldsymbol{p} = \boldsymbol{p}\boldsymbol{M}/2$ shows that $\boldsymbol{p} \equiv (p_1, p_2, \cdots)$ satisfies

$$\begin{aligned} 2p_2 &= 3p_1 \\ 2p_x &= 2p_{x-1} - \frac{p_{x-1}}{x-1} - \frac{p_{x-2}}{x-2}, \qquad x \geq 3. \end{aligned} \tag{4}$$

Standard methods, using the generating function $H(z) \equiv \sum_{x=1}^{\infty} p_x z^x$ applied to (4), lead to the equation $2z(z+1)H'(z) = (z^2+5z+2)H(z)$, hence $H(z) = Az(z+1)e^{z/2}$. Since it is required that $H(1) = 1$, we find that $A = \frac{1}{2}e^{-1/2}$, and so

$$H(z) = \frac{1}{2}z(z+1)e^{(z-1)/2}. \tag{5}$$

Extracting the coefficient of z^x in (4) yields the result (1).

Equation (5) demonstrates that $X - 1$ has a distribution which is the convolution of independent Poisson $(\frac{1}{2})$ and Bernoulli $(\frac{1}{2})$ variates. The following table shows the distribution of X.

x	1	2	3	4	5	≥ 6
$\mathbb{P}\{X = x\}$	0.3033	0.4549	0.1895	0.0442	0.0071	0.0001

Moments are easily obtained; $E(X) = 2$, $\mathrm{Var}(X) = 3/4$ whilst $E(X^r) = \sum_{i=0}^{r} \binom{r+1}{i} E(W^i)$, where $W \sim$ Poisson $(\frac{1}{2})$.

Further taxonomy of faces

A face of type k has k *ports* but may have fewer than k distinct *vertices* around its boundary due to the presence of loops at some vertices. For example, a face of with 5 ports may have v vertices where v ranges from 1 to 5.

A face can be classified as taxon $(v;\ i_1, i_2, \cdots, i_v)$ where v is the number of distinct vertices, and where i_j is the number of loops at vertex j intruding into the face. Note that such a face is of type $k = v + \Sigma i_j$.

This definition requires that the vertices are labelled $1, 2, \cdots, v$, let us say consecutively and anti-clockwise. Furthermore it is convenient at this stage to make a distinction between (say) (3; 2,1,0), (3; 1,0,2) and (3; 0,2,1), though clearly they are distinct only because of the vertex labelling.

Suppose we commence with a face having v vertices. Its daughters, and hence all descendants, cannot have more than v vertices (though *type* number has no bound). This suggests that if our equilibrium distribution *over taxa* exists, it either depends on the starting taxon or gives zero probability mass to taxa with $v \geq 2$. Investigation, which we briefly outline below, confirms the latter of these two alternatives.

If the initial face has $v = 1$, the possible taxa $\mathcal{J}_1$ are $\mathcal{J}_1 \equiv \{(1;\ x)\}_{x\geq 0}$, and it is easily shown that the 'daughter' matrix M for population development over taxa is that shown in (3), but now the element m_{ij} is the expected number of daughters of taxa $(1;\ j)$ from the division of a $(1;\ i)$ mother. Thus the equilibrium distribution over $\mathcal{J}_1$ is given by (1).

When the initial face has $v = 2$, the possible taxa are $\mathcal{J}_1 \cup \mathcal{J}_2$ where $\mathcal{J}_2 = \{(2;\ x, y)\}_{x\geq 0,\ y\geq 0}$. If the countable set $\mathcal{J}_2$ is ordered lexigraphically using $x + y$ (primarily) and x (secondarily), that is, (2; 0,0), (2; 0,1), (2; 1,0), (2; 0,2), (2; 1,1), (2; 2,0), (2; 0,3), (2; 1,2), $\cdots$, then the $\boldsymbol{M}$ matrix is of the partitioned form

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{M}_{11} & \boldsymbol{0} \\ \boldsymbol{M}_{21} & \boldsymbol{M}_{22} \end{pmatrix}. \tag{6}$$

Here $\boldsymbol{M}_{11}$ gives the population dynamics within $\mathcal{J}_1$ and is of the form (3), whilst

$$\boldsymbol{M}_{21} = \begin{pmatrix} \frac{2}{4} & & & & & & \\ \frac{3}{9} & \frac{2}{9} & & & & & \\ \frac{3}{9} & \frac{2}{9} & & & & & \\ \frac{4}{16} & \frac{4}{16} & \frac{2}{16} & & & & \\ \frac{4}{16} & \frac{4}{16} & & & & & \\ \frac{4}{16} & \frac{4}{16} & \frac{2}{16} & & & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

$$\boldsymbol{M}_{22} = \begin{pmatrix} \frac{4}{4} & \frac{1}{4} & \frac{1}{4} & & & & & & & & \\ \frac{4}{9} & \frac{6}{9} & 0 & \frac{2}{9} & \frac{1}{9} & & & & & & \\ \frac{4}{9} & 0 & \frac{6}{9} & 0 & \frac{1}{9} & \frac{2}{9} & & & & & \\ \frac{4}{16} & \frac{6}{16} & 0 & \frac{8}{16} & 0 & 0 & \frac{3}{16} & \frac{1}{16} & & & \\ \frac{4}{16} & \frac{4}{16} & \frac{4}{16} & 0 & \frac{8}{16} & 0 & 0 & \frac{2}{16} & \frac{2}{16} & & \\ \frac{4}{16} & 0 & \frac{6}{16} & 0 & 0 & \frac{8}{16} & 0 & 0 & \frac{1}{16} & \frac{3}{16} & \\ \vdots & & & & & & & & & & \ddots \end{pmatrix}.$$

The elements of $\boldsymbol{M}_{22}$ are given by m_{ij}, the expected number of offspring of taxon $j =$

$(2;\ x,y)$ from a mother of taxon $i=(2;\ r,s)$, being

$$
\begin{aligned}
m_{ij} &= \frac{2(r+s)+4}{(2+r+s)^2} && x=r, \quad y=s;\\
&= \frac{s+1}{(2+r+s)^2} && x=r, \quad y=s+1;\\
&= \frac{r+1}{(2+r+s)^2} && x=r+1, \quad y=s;\\
&= \frac{4+2y}{(2+r+s)^2} && x=r, \quad 0\le y<s;\\
&= \frac{4+2x}{(2+r+s)^2} && 0\le x<r, \quad y=s;\\
&= \frac{4}{(2+r+s)^2} && 0\le x<r, \quad 0\le y<s.
\end{aligned}
$$

The elements of $\boldsymbol{M}_{21}$ are given by m_{ij}, now taken to be the expected number of offspring of taxon $j=(1;\ x)$ from a mother of taxon $i=(2;\ r,s)$, being

$$
\begin{aligned}
m_{ij} &= \frac{r+s+2}{(2+r+s)^2} && x=0;\\
&= \frac{2(r+s+2-2x)}{(2+r+s)^2} && 0<x\le \min(r,s);\\
&= \frac{2(\max(r,s)+1-x)}{(2+r+s)^2} && \min(r,s)<x\le \max(r,s).
\end{aligned}
$$

The dynamics of a population driven by the matrix $\boldsymbol{M}$ in (6) is determined by the largest eigenvalues of its two sub-matrices $\boldsymbol{M}_{11}$ and $\boldsymbol{M}_{22}$. $\boldsymbol{M}_{11}/2$ is a stochastic matrix and so $\boldsymbol{M}_{11}$ has maximal eigenvalue λ_{11} equal to 2. By contrast $\boldsymbol{M}_{22}/2$ is sub-stochastic so the maximal eigenvalue λ_{22} of $\boldsymbol{M}_{22}$ is less than 2. (We have computed λ_{22} to be 1.463.) Thus the population of faces with $v=1$ grows asymptotically like 2^n, dominating the population with $v=2$. The latter group grows asymptotically in number like λ_{22} to the power n. Though it never becomes extinct, the group of faces with $v=2$ forms a negligible part of the population as $n\to\infty$.

Though we have not performed detailed eigenvalue calculations for all cases where v is initially a number $u\ge 3$, it is clear that the partitioned $\boldsymbol{M}$ matrix has above-diagonal sub-matrices equal to $\boldsymbol{0}$ and diagonal sub-matrices, $\boldsymbol{M}_{11},\ \boldsymbol{M}_{22},\ \cdots,\ \boldsymbol{M}_{uu}$. Since each taxon with $v=3$ has precisely one daughter with $v=3$, then when $u=3$, the number of faces with $v=3$ remains constant. Thus $\boldsymbol{M}_{33}$, whose details we omit, has maximal eigenvalue equal to 1. When $u\ge 4$, it is easy to show that $\lambda_{uu}<1$ since the expected number of u-taxon daughters is less than 1. Thus, faces with $v\ge 4$ become extinct, whilst those with $v = 2$ or 3 remain in the population but are dominated in number by $v = 1$ faces.

Variations on a theme

In this section we consider variations on the division rule. The first variation is one where the two ports of the dividing face are sampled *without* replacement. This eliminates the possibility of loops (so let us assume too that the initial graph has no loops). The types are 2, 3, 4, $\cdots$. The $\boldsymbol{M}$ matrix is

$$\boldsymbol{M} = \begin{pmatrix} \frac{4}{2} & & & & & \\ \frac{3}{3} & \frac{3}{3} & & & & \\ \frac{4}{6} & \frac{4}{6} & \frac{4}{6} & & & \\ \frac{5}{10} & \frac{5}{10} & \frac{5}{10} & \frac{5}{10} & & \\ \frac{6}{15} & \frac{6}{15} & \frac{6}{15} & \frac{6}{15} & \frac{6}{15} & \\ \vdots & & & & & \ddots \end{pmatrix}.$$

It is easily seen that types 4, 5, $\cdots$, if initially present, become extinct with certainty. The number of type 3 settles to a constant, namely $(\bar{f} - 2)\mathcal{F}$, where $\bar{f}$ is the mean type in the initial graph, whilst the type-2 population grows asymtotically like 2^n and hence dominates.

Another variant is to allow the unbounded face to divide by joining two of the 'ports' of this face. To make this division unambigious, however, we think it best to place the action on the surface of a ball whereupon the results are the same as those considered already.

More substantive problems arise if we consider division rules which (a) join a randomly selected port of a face to a randomly selected edge or (b) join two randomly selected edges. Severe complications arise here, because new vertices are created on the boundary of a face by the dividing lines of *other* faces. Without recourse to computer simulation, little can be said about the equilibrium distribution of face type *unless* we adopt a convention which effectively disregards these 'externally generated' vertices when classifying a given face. To establish this convention let us call the ports of all initial faces 'corners'. For a face in later generations, a port is called a 'corner' if either it was a corner of the face's mother or it was created by the line which divided the mother. The collection of edges joining two adjacent corners of a face is called a 'side' of that face. The face 'type' is therefore determined by the number of corners (or sides) of that face.

Suppose the division rule joins a randomly chosen *corner* to a randomly chosen *side* (equally-likely in each case). Any initial loop in the graph is immediately destroyed, leaving possible types 2, 3, $\cdots$. The $\boldsymbol{M}$ matrix on these types is (ignoring type 1)

$$\boldsymbol{M} = \begin{pmatrix} 1 & 1 & & & & \\ \frac{2}{3} & \frac{2}{3} & \frac{2}{3} & & & \\ \frac{2}{4} & \frac{2}{4} & \frac{2}{4} & \frac{2}{4} & & \\ \frac{2}{5} & \frac{2}{5} & \frac{2}{5} & \frac{2}{5} & \frac{2}{5} & \\ \vdots & & & & & \ddots \end{pmatrix} \tag{7}$$

the general element m_{ij} being equal to $2/i$ for $i = 2, 3, \cdots$ and $j = 2, 3, \cdots, (i+1)$. From Cowan (1989), where an identical $\boldsymbol{M}$ matrix governs the side-to-side splitting of polygons by straight dividing lines, we can conclude that the type, X_n, of a randomly selected member of generation n is such that $X_n - 2$ converges in distribution to a Poisson (1) variate.

If the division rule joins two randomly selected *sides* of a face, sampled *without* replacement, then the $\boldsymbol{M}$ in (7) arises yet again. (Here we assume that the initial graph has no loops; any initial 2-faces will also be immediately destroyed. Thus face-type is in the range 3, 4, $\cdots$, with equilibrium distribution $X_n - 3 \xrightarrow{\mathcal{D}}$ Poisson (1).

Lastly, if the side-to-side rule has sampling with replacement, the same side can be chosen twice, thus creating a 2-face. We find that $\boldsymbol{M}$, on the types 2, 3, $\cdots$, is

$$\boldsymbol{M} = \begin{pmatrix} \frac{1}{2} & \frac{3}{2} & \frac{1}{2} & & & \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} & \frac{1}{3} & & \\ \frac{1}{4} & \frac{2}{4} & \frac{2}{4} & \frac{2}{4} & \frac{1}{4} & \\ \vdots & & & & & \ddots \end{pmatrix},$$

namely m_{ij} equals $1/i$ if $j = 2$ or $i+2$, or $2/i$ if $3 \leq j \leq i+1$. Standard methods lead to the equilibrium distribution $\{p_x\}$ for $x \geq 2$,

$$p_x = e^{-5/4} \sum_{m=0}^{[x/2]} \frac{x(x-2m-1)}{2^{2m} m!(x-2m)!}.$$

The distribution of X is tabulated below. Note that, since each division line creates one new face and four new corners, $E(X) = 4$.

x	2	3	4	5	6	7	≥ 8
p_x	0.1433	0.2865	0.2507	0.1671	0.0880	0.0400	0.0244

Acknowledgement

We have had helpful discussion with Mr. Lam Kai-Yam.

References

Cowan, R. (1989). The division of space and the Poisson distribution. *Adv. Appl. Prob.* **21,** 233-234.

Cowan, R. and Chen, S. (1996). The random topology of polygon splitting. In preparation.

Cowan, R. and Morris, V.B. (1988). Division rules for polygonal cells. *J. Theor. Biol.* **131,** 33-42.

Cox, D.R. and Miller, H.D. (1965). *The Theory of Stochastic Processes.* Methuen, London.