

## On the correlation between sizes and shapes of cells in epithelial mosaics

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1982 J. Phys. A: Math. Gen. 15 L143

(<http://iopscience.iop.org/0305-4470/15/3/012>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

### Download details:

IP Address: 129.252.86.83

The article was downloaded on 30/05/2010 at 15:49

Please note that [terms and conditions apply](#).

LETTER TO THE EDITOR

On the correlation between sizes and shapes of cells in epithelial mosaics

N Rivier and A Lissowski

Blackett Laboratory, Imperial College, London, SW7 2BZ, England

Received 22 October 1981

**Abstract.** It is shown that Lewis' empirical, linear relationship between the average area of a cell and the number of its sides in two-dimensional mosaics corresponds to maximal arbitrariness in the cellular distribution. An expression for the distribution is given in the general case.

In 1928, Lewis (1928, 1943, 1944) observed in several two-dimensional cellular mosaics (cucumber epidermis, pigmented epithelium of the retina, etc), at various stages of their development, a specific relationship between the average area  $A_n$  of a resting cell and the number  $n$  of its edges,

$$A_n = a(n - 2) \tag{1}$$

where  $a = A_0/4F$ , for a tissue containing  $F$  cells covering a total area  $A_0$ . Either one or both of  $A_0$  and  $F$  can be time dependent, to account for growth and cellular division; yet the relationship (1) is valid throughout the development of the tissue.

In this letter, we shall give a mathematical justification for Lewis' law (1). It will be shown to be a consequence of the equilibrium between entropy and organised form, in this case, the existence of space-filling cells and their topology.

The cellular tissue is described by a distribution of faces (cells)  $\{F_n\}$  with  $n$  edges,  $n = 3, 4, \dots, N + 2$ .  $F_n$  is the number of faces with  $n$  edges, and  $p_n = F_n/F$ , the probability of finding an  $n$ -sided face. Neither spatial inhomogeneity (dependence of  $p_n$  on position) nor correlation between the number of sides of neighbouring cells need be assumed in this letter. The cell distribution  $\{F_n\}$  must satisfy the following constraints

$$\text{Normalisation} \quad \sum F_n = F \tag{2}$$

$$\text{Fixed total area} \quad \sum A_n F_n = A_0 \tag{3}$$

$$\text{Topology} \quad \sum (6 - n)F_n = 0. \tag{4}$$

(Equation (2) is simply the normalisation of the probability distribution  $\sum p_n = 1$ . Equation (4) is a direct consequence of Euler's theorem relating the number of faces  $F$ , edges  $E$  and vertices  $V$  covering a two-dimensional manifold of Euler-Poincare characteristics  $\chi$ ,  $F - E + V = \chi$ ; (A).  $\chi$  describes the global character of the manifold. For a sphere,  $\chi = 2$ , and for an  $s$ -handed torus, it is  $\chi = 2(1 - s)$ . Unlike  $F$ ,  $E$  and  $V$ , it does not scale with the number of faces, and, for a large tissue,  $\chi = O(1)$  is negligible in comparison with  $F$ ,  $E$ ,  $V = O(F)$ . We have the relations  $\sum nF_n = 2E$  (B) since every edge separates two faces, and  $\sum \alpha V_\alpha = 2E$ , since every edge joins two vertices and every

vertex of connectivity  $\alpha$  has  $\alpha$  incident edges. In a cellular mosaic, we can assume that  $\alpha = 3$ , since an exceptional vertex with  $\alpha = 4$ , say, can be transformed by an infinitesimal distortion into two  $\alpha = 3$  vertices connected by an additional edge. Therefore, only  $\alpha = 3$  vertices are structurally stable. Thus  $V = (\frac{2}{3})E$ ; (C). Equations (B) and (C) into (A) yield equation (4), or precisely,  $\Sigma(6-n)p_n = 6\chi/F = O(1/F) \approx 0$  in the limit of a large number of cells  $F \rightarrow \infty$ .)

The cellular structure of the mosaic is therefore a solution of the system of linear, inhomogeneous equations (2)–(4), with  $N$  unknowns  $F_n$ , written in matricial form

$$\mathbf{FM} = \mathbf{G} \quad (5)$$

with  $\mathbf{F} = (F_3 \dots F_{N+2})$  and  $\mathbf{G} = (FA_0)$ .  $\mathbf{M}$  is a  $(N \times 3)$  matrix. The solutions of the system (5) lie in a hyperplane of dimension  $d = N - r$ , where  $r$  is the rank of the matrix  $\mathbf{M}$ .

In general, that is for arbitrary area relationships  $A_n$ ,  $r = 3$ , and  $d = N - 3$ . The rank of  $\mathbf{M}$  is  $r = 2$ , if and only if there is a linear relationship between the three equations (2)–(4), (2) =  $\lambda$ (4) +  $\mu$ (3), so that  $A_n$  is a linear function of  $n$ ,

$$1 - \lambda(6 - n) - \mu A_n = 0 \quad (6)$$

with  $\lambda$  and  $\mu$  so far arbitrary. Moreover, the inhomogeneous, linear system (5) has a solution if and only if  $\mathbf{M}$  and the amplified matrix formed by affixing  $\mathbf{G}$  as an additional row to  $\mathbf{M}$  have the same rank,  $r = 2$ . This yields

$$F - \mu A_0 = 0. \quad (7)$$

Thus,

$$A_n = (A_0/F)\lambda[n - (6 - 1/\lambda)]. \quad (8)$$

Lewis's law (1) is obtained if one adds another obvious topological relation expressing the fact that a cell has at least three edges,  $A_2 = 0$ , which yields  $\lambda = \frac{1}{4}$ . This last condition is not necessary and may be too strong in general. Lewis (1931) (see also Smoljanin 1980) has given examples of tissues following (8) with  $2 \neq 6 - 1/\lambda < 3$ . The cases  $r > 3$  or  $r = 1$  are clearly impossible.

Consider, for example, a tissue containing 5-, 6- and 7-sided cells only. Equations (2)–(4) yield  $F_7 = F_5$ ,  $F_6 = F - 2F_5$ , and  $(A_5 + A_7 - 2A_6)F_5 = A_0 - A_6F$ . This last equation fixes  $F_5$ , and therefore the  $\{F_n\}$  distribution, unless the areas satisfy Lewis's law (8), in which case it becomes an identity  $0 = 0$ , and  $F_5$  remains arbitrary.

Thus Lewis's law (1) or (8) corresponds to *maximal* arbitrariness in the distribution  $\{p_n\}$  of the various cells, which lie then in a hyperplane of higher dimension than in the general case. The only constraints in the distribution are its normalisation and its mean  $\langle n \rangle = 6$  (the topological relation (4)). Even its variance can be chosen freely. Two apparently dissimilar tissues can both satisfy Lewis's law, and so can normal and pathological cell mosaics in the same tissue.

To give a precise meaning to the concept of arbitrariness, and derive a relation between  $p_n$  and  $A_n$  in the general case  $r \neq 2$ , we must use information theory (Shannon 1948). Among all possible distributions of cells compatible with physical, topological and possibly biological constraints, a large system will take the most probable configuration, i.e. that which maximises the entropy or arbitrariness

$$H = -\sum p_n \ln p_n \quad (9)$$

subject to the constraints mentioned above. Any other distribution would betray

additional assumptions on the system which should have been included among the constraints.

The constraints are given by equations (2)–(4). Using indeterminate Lagrange multipliers  $\alpha$ ,  $\beta$  and  $\gamma$  we maximise

$$\Psi\{p_n\} = H - \alpha \left( \sum p_n - 1 \right) - \beta \left( \sum A_n p_n - A_0/F \right) - \gamma \sum (6-n)p_n \quad (10)$$

for the variables  $p_n$ , and obtain

$$-\ln p_n - 1 - \alpha - \beta A_n - \gamma(6-n) = 0. \quad (11)$$

Thus,

$$p_n = \exp(\gamma n - \beta A_n) / Z \quad (12)$$

with  $Z = \sum_n \exp(\gamma n - \beta A_n)$ , and  $\gamma$  and  $\beta$  determined by equation (4),  $6 = \sum n p_n = (\partial/\partial\gamma) \ln Z$ , and equation (3)  $A_0/F = \sum p_n A_n = -(\partial/\partial\beta) \ln Z$ . There is a close parallel between our problem and equilibrium statistical thermodynamics, where a large assembly of identical systems takes up the most probable distribution among the possible states available to any one system. In particular, Jaynes (1957) has shown that a constructive, subjective statistical mechanics could be constructed on the basis of information theory.

Suppose now that the area law can be varied to maximise the entropy further. In the absence of restrictions (3) and (4), the entropy or arbitrariness reaches its absolute maximum,  $\hat{H} = \ln N$ , if the  $N$  configurations are equiprobable,  $p_n = N^{-1}$ . In the presence of the constraints, it is useful to introduce the moments  $\mu_s$  of the probability distribution,

$$\mu_s = \sum_n n^s p_n \quad \mu_0 = 1 \quad (13)$$

and to expand the function  $A_n$  in power series in  $n$ ,

$$A_n = a_0 + \sum_{i=1}^{\infty} a_i n^i \quad (14)$$

The constraints (4) and (3) read now

$$\mu_1 = 6 \quad (4')$$

$$a_0 + \sum a_i \mu_i = A_0/F \quad (3')$$

respectively. To vary the area law (14), we shall vary the parameters  $a_i$  ( $i \geq 1$ ).  $a_0$ , which does not affect  $p_n$  and therefore  $H$ , can be regarded as a constant. All the  $a_i$  are independent variables, except two which are related to the others through the constraints (3') and (4'). Variation of the  $a_i$  implies variation of  $\gamma$ ,  $\beta$  and the moments  $\mu_i$ . However, the entropy  $H$

$$H(\{a_i\}) = -\gamma \mu_1 + \beta \sum a_i \mu_i + \ln Z \quad (15)$$

is stationary in the Lagrange multipliers  $\gamma$  and  $\beta$ , as can be verified immediately. (It is in fact the Legendre transform of  $\ln Z$ , a standard manipulation in thermodynamics.)

Denoting  $\delta f = \sum \delta a_i (\partial f / \partial a_i)$  we extremise the entropy

$$0 = \delta H = [(\partial / \partial \gamma) \ln Z - \mu_1] \delta \gamma + \left( (\partial / \partial \beta) \ln Z - \sum a_i \mu_i \right) \delta \beta - \gamma \delta \mu_1 - \beta \delta \left( \sum a_i \mu_i \right) + \sum (\partial / \partial a_i) \ln Z \delta a_i. \tag{16}$$

The coefficients of  $\delta \gamma$  and  $\delta \beta$  vanish identically (stationarity), as do the third and fourth terms because of the constraints (3') and (4'). Moreover,  $(\partial / \partial a_i) \ln Z = -\beta \mu_i$  and (3') yields  $\sum \mu_i \delta a_i = -\sum a_i \delta \mu_i$  so that

$$0 = \delta H = -\beta \sum \mu_i \delta a_i = \beta \sum a_i \delta \mu_i \tag{17}$$

which has the obvious solution  $a_i = 0 (i \geq 2)$ .  $a_1$  remains arbitrary because  $\delta \mu_1 = 0$  (equation (4')). Thus, maximal arbitrariness (17) makes the area (14) follow the generalised Lewis law (8), because  $a_0$  and  $a_1$  are related by the constraints,  $a_0 + 6a_1 = A_0/F$ . (The  $a_i$  are not all independent variables, and some care must be taken in solving (17). Let  $\delta a_1$  and  $\delta a_2$  be related to the other  $\delta a_i$  through (3') and (4'). Using  $\delta \mu_i = \sum_l (\partial \mu_i / \partial a_l) \delta a_l$ , with  $\partial \mu_i / \partial a_l = -\beta (\mu_{i+l} - \mu_i \mu_l)$ , we can write (17) in terms of independent variations  $0 = \sum_{i \geq 3} \delta a_i \sum_l (\varphi_{il} a_l)$  where  $\varphi_{il}$  are third-order polynomials in the moments  $\mu_3$  with  $\varphi_{i1} = 0$ . The solution is thus  $\sum \varphi_{il} a_l = 0$  i.e.  $a_l = 0 (l \geq 2)$  as above. The existence of another solution would imply an intricate relation between moments. It is unlikely, but we shall discuss it below.)

The probability (12) of finding an  $n$ -sided face obeying Lewis's area law (8) has the form

$$p_n \propto \exp(-\kappa n). \tag{18}$$

where  $\kappa = \gamma - \beta a_1$ . It increases with  $n$  ( $\kappa < 0$ ) for  $N < 7$ , is flat for  $N = 7$ , and decreases with increasing  $n$  ( $\kappa > 0$ ) for  $N > 7$ . For  $N \rightarrow \infty$ ,  $p_n = (\frac{1}{3})(\frac{2}{4})^{n-2}$ , and the entropy is  $H = 4 \ln 4 - 3 \ln 3$ , well below  $\hat{H}$  expected in the absence of constraint. Lewis's law yields maximal entropy because the two Lagrange multipliers ensuring the constraints always appear in the combination  $\gamma - \beta a_1$ , and not independently.

The distribution of resting cells of the cucumber epidermis (Lewis 1928) does not follow the exponential probability law (18), even though their area satisfies Lewis's law (1). This means that there is one or several other, yet unidentified, constraints, in addition to (3) and (4),

$$\sum f_n p_n = f_0 \tag{19}$$

say. The calculation (10)–(17) proceeds exactly as above, with additional Lagrange multiplier(s)  $\zeta$ , and one obtains the probability distribution

$$p_n = \exp(\gamma n - \beta A_n - \zeta f_n) / Z \tag{20}$$

instead of (12), with  $Z = \sum \exp(\gamma n - \beta A_n - \zeta f_n)$ . Equation (17) now has two solutions: either  $A_n$  mimics the topological constraint  $A_n = a_0 + a_1 n$ , or it mimics the new constraint (19),  $A_n = a_0 + \eta f_n$ . In either case, the probability distribution has the form

$$p_n \propto \exp(\rho n - \sigma f_n). \tag{21}$$

Both solutions maximise the entropy to the same value, because  $\langle n \rangle = 6$  and  $f_0$  are equal in both cases. However, there is no physical reason for  $A_n$  to be related to  $f_n$ , but every reason that it should follow the topological constraint  $A_n = a_0 + a_1 n$ .

As an illustration, let us return to the tissue with only 5-, 6- and 7-sided cells. Suppose that we had initially, for example immediately after several cellular divisions, a tissue with average areas  $f_n$  and  $f_0$  different from the resting situation  $A_n, A_0/F$ . Suppose further that the system is left to develop towards the resting situation through growth only, without cellular division or exchange of edges between cells, i.e. by keeping its distribution  $\{F_n\}$  invariant. The *same*  $\{F_n\}$  must therefore satisfy the four constraints (2)–(4) and (19), which reduce to  $F_7 = F_5, F_6 = F - 2F_5$ , with two linear equations for the single unknown  $F_5$ ,  $(A_5 + A_7 - 2A_6)F_5 = A_0 - A_6F$  and  $(f_5 + f_7 - 2f_6)F_5 = (f_0 - f_6)F$ . These two equations have a solution if and only if *either* the  $A_n$  satisfy Lewis's law (8), in which case the first equation being an identity,  $F_5$  is given by the second equation, i.e. by the initial conditions, *or* the two equations are proportional,  $(A_0 - A_6F)/(A_5 + A_7 - 2A_6) = (f_0 - f_6)F/(f_5 + f_7 - 2f_6)$ , and the tissue develops in a self-similar fashion,  $A_n = a_0 + \eta f_n$ . Here, we need maximal arbitrariness to have a solution at all. The same discussion could be made by comparing the tissue in a resting situation and just about to undergo several cellular divisions, under the same hypotheses.

In conclusion, we have shown that the relation (8) between average cell area and the number  $n$  of its edges corresponds to the tissue selecting, for a given density, a cell distribution of maximal arbitrariness or minimal information, but compatible with topological constraints. We have also given a general expression (12) or (20) for the cell distribution in the general case with any given area law  $A_n$ .

In the absence of a specific, dynamical mechanism regulating the area of individual cells, the tissue takes up the most probable cell distribution, which implies a correlation between cell sizes and cell shapes, Lewis' law (1) or (8).

The same analysis of Lewis's law applies, of course, to any random two-dimensional mosaics, for example to crystal aggregates in metallurgy (see, for example, Kurtz and Carpay 1980, and references therein), or to the solar granulation (a mosaic of convection cells) (Bray and Loughhead 1967).

Lewis's law is therefore of diagnostic importance. If a mosaic does not obey it, then the average area of its constituent cells is not regulated simply by the area-filling requirement, but by a specific mechanical or biological law (which it is of obvious interest to isolate) or by the fact that space-filling constraints act in a higher dimension.

As an example of the latter, plane sections of three-dimensional crystal aggregates have different statistical properties (the crystal sections are more equiaxial) from crystals grown in very thin strips, for which the requirements of space-filling are genuinely two-dimensional (Meijering 1953), and which may follow Lewis's law. Monte Carlo generated mosaics, the Voronoi polygons of Poisson-distributed seeds (Crain 1978), which correspond to Meijering's cell model, are clearly free of dynamical constraints. The average polygon area does indeed follow Lewis's law (Crain 1978, table 3).

Two-dimensional mosaics can therefore be separated into two classes by Lewis's law. Area statistics is *either* determined solely by space-filling requirements, and obeys Lewis's law, *or* by a particular, probably local dynamic interaction. In view of the complexity of biological tissues or of crystal growth, it is fortunate and astonishing, or further evidence for the attraction of the most probable distribution in large systems, that the first class is non-empty.

**References**

- Bray R J and Loughhead R E 1967 *The Solar Granulation* (London: Chapman and Hall)  
Crain I K 1978 *Comput. and Geosc.* **4** 131-4  
Jaynes E T 1957 *Phys. Rev.* **106** 620-30  
Kurtz S K and Carpay F M A 1980 *J. Appl. Phys.* **51** 5745-54  
Lewis F T 1928 *Anat. Rec.* **38** 341  
—— 1931 *Anat. Rec.* **50** 235  
—— 1943 *Am. J. Bot.* **30** 766  
—— 1944 *Am. J. Bot.* **31** 619  
Meijering J L 1953 *Philips Res. Rep.* **8** 270-90  
Shannon C E 1948 *Bell Syst. Tech. J.* **27** 379, 623  
Smoljaninov V V 1980 *Mathematical Models of Tissues* (Moscow: Nauka) in Russian