Math. Proc. Camb. Phil. Soc. (1987), **101**, 313
Printed in Great Britain

Explicit formulae for polygonally generated shape-densities in the basic tile

BY HUI-LIN LE
Statistical Laboratory, University of Cambridge

(*Received* 31 *March* 1986; *revised* 4 *July* 1986)

Abstract

Kendall and Le[**4**] gave an exact algorithm for the computation of a polygonally generated (auxiliary) shape-density $\tilde{m}(x, y)$ which we used to obtain 'exact' numerical densities for given (x, y). Here I derive an explicit formula for $\tilde{m}(x, y)$, for an arbitrary convex plane polygon K, valid when the shape-point (x, y) lies in what will be called the '(upper) basic tile' of the associated singular tessellation. In most circumstances this is all that is needed for the statistical analysis of collinearities.

1. *Introduction*

The statistics of testing for 'too many collinearities' can be based on a model in which a typical triplet of points is a sample of size 3 from the uniform law $U(K)$ on the interior of a given compact convex set K in the plane. (For collinearities and the associated literature see Kendall and Kendall[**3**].) This can usefully be viewed from the standpoint of the theory of shape, where by the shape of a labelled set of points we mean the total information that remains when the location, orientation and scale of the set of points are ignored. Thus two labelled triplets will have the same shape if they form the vertices of two similar (labelled) triangles. A general theory of shape with applications and references can be found in Kendall[**1**]. A general formula for the (auxiliary) shape-density $\tilde{m}(x, y)$ of such a triplet in terms of the shape-coordinates (x, y) has been developed in [**2**] and [**4**]. This density can be reduced to the sum of six double integrals over the torus $\partial K \times \partial K$, and it is known [**4**] that if K is a convex polygon with vertices $Q_1 = (u_1, v_1)$, $Q_2 = (u_2, v_2)$, $\ldots$, $Q_n = (u_n, v_n)$, labelled in counter-clockwise order, then the shape-density $\tilde{m}(x, y)$ is locally real-analytic as a function of x and y, with singularities concentrated on a network of circular arcs, segments and half-lines called the singular tessellation $\mathscr{T}(K)$, a part of which is the whole of the x-axis. It is also known (Small[**5**]) that $\tilde{m}(x, 0)$ is constant on $\{(x, y)$: $-1/\sqrt{3} \leqslant x \leqslant 1/\sqrt{3}, y = 0\}$. The value of the constant is

$$C(K) = \frac{3}{4|K|} J_2(K),$$

where $|K|$ is the area of the polygon K and $J_2(K)$ is the second moment of the distribution of the distance between two independent random points in K. This segment $\{(x, y)$: $-1/\sqrt{3} \leqslant x \leqslant 1/\sqrt{3}, y = 0\}$ we call the *basic segment*. Note that $\tilde{m}$ is not constant on the rest of $y = 0$.

It is a property of the singular tessellation $\mathscr{T}(K)$ that the basic segment is always

https://doi.org/10.1017/S0305004100066688 Published online by Cambridge University Press

an edge of exactly two 'basic' tiles (a pair of congruent triangles). The remaining edges of the two basic tiles lie respectively on the pairs of lines

$$y\sqrt{3} = D_1(K)\,(1-x\sqrt{3}), \quad y\sqrt{3} = D_2(K)\,(1+x\sqrt{3}),$$

and

$$y\sqrt{3} = -D_1(K)\,(1+x\sqrt{3}), \quad y\sqrt{3} = -D_2(K)\,(1-x\sqrt{3}),$$

where

$$D_1(K) = \min_k \left\{ \frac{p_{k-1,\,k+1}}{\sqrt{(d^2_{k-1,\,k+1} - p^2_{k-1,\,k+1})}} \right\},$$

$$D_2(K) = \min_k \left\{ \frac{p_{k,\,k-1}}{\sqrt{(d^2_{k-1,\,k+1} - p^2_{k,\,k-1})}} \right\}.$$

Here $d_{r,\,s} = \text{dist}\,(Q_r, Q_s)$, $p_{r,\,s} = \text{dist}\,(e_r, Q_s)$, and e_s is the oriented edge of K from the vertex Q_s to the vertex Q_{s+1}. (Whenever necessary we employ the convention that Q_s is identical with $Q_{s\pm n}$, where n is the number of vertices of K.)

The fact that the shape-density $\tilde{m}(x,y)$ has different expressions on different tiles of the tessellation makes it very laborious to work out $\tilde{m}(x,y)$ for all (x,y). But unless K has some vertex very close to the chord joining its two neighbours, the quantities $D_i(K)$ $(i = 1,2)$ will be so large that the union of the two basic tiles and their re-labelling transforms will cover most of the region involved in collinearity testing, and indeed all of it if the angular criterion for near-collinearity is used (for which see Kendall [**2**]). Fig. 1 illustrates the situation. This very important fact motivates the present paper, in which we determine $\tilde{m}(x,y)$ exactly in the basic tiles; the corresponding formulae in their 4 re-labelling equivalents are easily found by using the transformations (with associated jacobians) described in [**4**]. Note that always

$$\tilde{m}(x,y) = \tilde{m}(-x,-y),$$

so that in the following calculation we need only consider the 'upper' one of the two basic tiles (that in $y \geqslant 0$), and in all that follows this basic tile is meant unless there is an explicit statement to the contrary. Thus a term $g(x)\,y^r$ in $\tilde{m}(x,y)$ for the ('upper') basic tile will produce a term $g(-x)\,(-y)^r$ in $\tilde{m}(x,y)$ for the 'lower' basic tile, and so there will be a singularity at the point $(x,0)$ if

(i) r is odd, and $g(-x) \neq -g(x)$, and also if

(ii) r is even, and $g(-x) \neq g(x)$.

The main result of this paper is the explicit formula at (4) *below for* $\tilde{m}(x,y)$ *in the* ('*upper*') *basic tile.* As the p's, the a's and the α's are constants associated with K this formula, complicated though it may appear, is just a sum of Small's constant and a finite number of fractions of the form

$$cy^m/(T_1\,T_2 \ldots T_m),$$

where each T is of one of the two forms $(1 \pm x\sqrt{3})\sin\alpha + y\sqrt{3}\cos\alpha$, and $m = 2$, 3 or 4. The reader is recommended to proceed now to the applications in section 3, and then to refer to section 2 for the details of the derivation.

https://doi.org/10.1017/S0305004100066688 Published online by Cambridge University Press

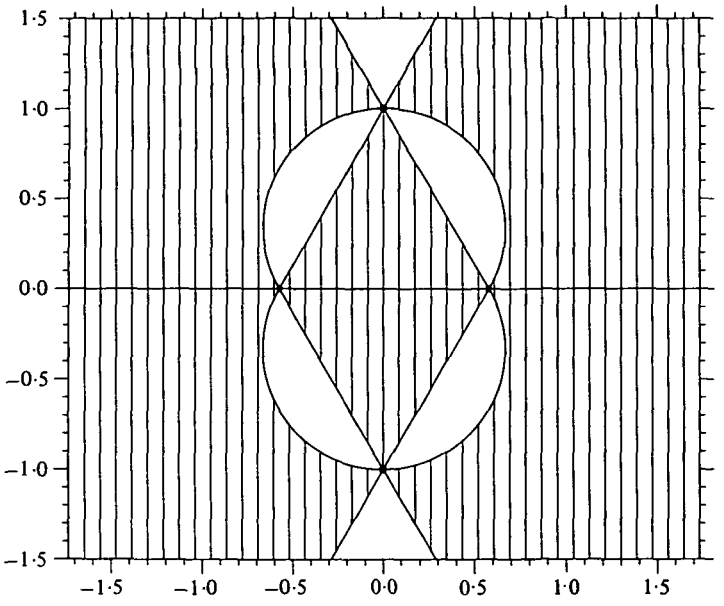


Fig. 1. The singular tessellation for K = equilateral triangle. The basic tile and its equivalents are shaded.

2. *Method, and general result*

From [2] and [4] we know that the term associated with z^* in the formula for $\tilde{m}(x, y)$ is

$$M(z^*) = \frac{1}{32|K|^3}\int_{\partial K} R_1^2\, d\theta_1 \int_{\partial K} R_{12}^2\, 1(P_3 \in K)\, R_{12}^2\, d\psi_{12} |dz^*/dz|^2, \tag{1}$$

where z^* is obtained from a given triangle of shape z via the re-labelling group of transformations $\{1, \alpha, \alpha^2, \beta, \beta\alpha, \beta\alpha^2\}$ defined in [4]. Here R_1 is the distance from the centroid G of the area of K to P_1, R_{12} is the distance from P_1 to P_2, and θ_1 is the direction of GP_1, while ψ_{12} is the direction of P_1P_2. The third vertex P_3 depends on P_1, P_2, and z^*. Thus we can write $\tilde{m}(x, y) = F(x, y) + F(-x, -y)$, where $F(x, y) = M(z) + M(\alpha z) + M(\alpha^2 z)$.

The term (1) can be rewritten ([4]) as

$$\frac{1}{32|K|^3}\sum_{i=1}^{n} p_i a_i \sum_{j \neq i} a_j \iint_{K^+(e_i, e_j, z^*)} f(s, t)\, g(s)\, ds\, dt |dz^*/dz|^2. \tag{2}$$

Here p_i is the length of the perpendicular from the centroid G to the edge e_i, a_i is the length of the edge e_i, $f(s, t)$ (a quadratic polynomial) is the square of the distance from $P_1 \in e_i$ to $P_2 \in e_j$, $g(s)$ (a linear polynomial) is the length of the perpendicular from P_1 to e_j, and $K^+(e_i, e_j, z^*)$ is the convex polygon in the (s, t)-plane associated with the requirement that $P_3(s, t, z^*) \in K$, where sa_i is the distance measured along the oriented edge e_i to P_1, and similarly ta_j is the distance along e_j to P_2.

Now clearly $0 \leqslant s \leqslant 1$ and $0 \leqslant t \leqslant 1$, and so $K^+(e_i, e_j, z^*)$ is determined by these four affine conditions together with n further affine conditions of which the mth, given below, is typical; this will express the condition that P_3 lies to the left of the oriented edge e_m. We describe this in the case when e_m occurs *before* the forward s-arc from

https://doi.org/10.1017/S0305004100066688 Published online by Cambridge University Press

e_i to e_j; the other cases follow by similar arguments. We calculate the 'height' of P_3 above e_m (where e_m is viewed as going horizontally from left to right), and we compute this via a preliminary evaluation of the 'height' of $(P_1+P_2)/2$ above e_m. The affine condition then turns out to be

$$L(1-s)a_i + Mta_j + N \geqslant 0, \tag{3}$$

where

$$L = -(1-x^*\sqrt{3})\sin(\beta_i-\beta_m)+y^*\sqrt{3}\cos(\beta_i-\beta_m),$$
$$M = (1+x^*\sqrt{3})\sin(\beta_j-\beta_m)+y^*\sqrt{3}\cos(\beta_j-\beta_m)$$

and

$$N = (1+x^*\sqrt{3})\sum_{i+1}^{j-1} a_\lambda \sin(\beta_\lambda-\beta_m)+y^*\sqrt{3}\sum_{i+1}^{j-1} a_\lambda\cos(\beta_\lambda-\beta_m)-2\sum_{i+1}^{m-1} a_\lambda\sin(\beta_\lambda-\beta_m).$$

(Notice that the summations in N vanish if a complete circuit of the polygon is taken.) Here $z^* = (x^*, y^*)$, and for each λ the angle $\beta_\lambda-\beta_m$ gives the orientation (in a counter-clockwise sense) of e_λ relative to e_m. If we write α_λ for the interior angle of the polygon K at the vertex common to $e_{\lambda-1}$ and e_λ, then the β-differences can be computed from

$$\beta_\mu-\beta_\lambda = \sum_{\lambda+1}^{\mu}(\pi-\alpha_\nu) = (\mu-\lambda)\pi - \sum_{\lambda+1}^{\mu}\alpha_\nu.$$

In particular, when z is on the basic segment, the two polygons $K^+(e_i, e_j, z)$ and $K^+(e_i, e_j, \beta z)$ are complete squares $\{(s,t): 0 \leqslant s \leqslant 1,\ 0 \leqslant t \leqslant 1\}$ for all i and $j \neq i$, and the contributions from the other re-labelling transforms are all void, so that

$$\tilde{m}(x,0) = \frac{1}{16|K|^3}\sum_{i=1}^{n} p_i a_i \sum_{j\neq i} a_j \int_0^1\int_0^1 f(s,t)\,g(s)\,ds\,dt$$

for $|x| \leqslant 1/\sqrt{3}$. This is another version of Small's formula for $C(K)$.

The fact that $\tilde{m}(x,y)$ is real-analytic in each tile plays an important role in the calculation of $\tilde{m}(x,y)$, because this means that we can get the formula for $\tilde{m}(x,y)$ in a given tile by choosing a specially convenient point z_0 in the tile, finding the form of $K^+(e_i, e_j, z_0^*)$, and then calculating $\tilde{m}$ at a general (x,y) in the tile by using the fact that K^+ depends analytically on z_0^*. We shall find that in the basic tile many terms in (2) are zero (because K^+ is void or degenerate), and many others are tile-wise constant. It is convenient to denote the sum of the locally constant terms by $C(K)$ (this will be Small's constant), so that $\tilde{m}(x,y)$ can be written as

$$\tilde{m}(x,y) = F_1(x,y)+F_2(-x,-y),$$

where

$$F_1(x,y) = \tfrac{1}{2}C(K)+\sum_{i=1}^{n} p_i a_i \Bigg\{ \sum_{j=i-2,\,i-1} a_j \iint_{CK^+(e_i,e_j,z)} f(s,t)\,g(s)\,ds\,dt$$
$$+ a_{i+1}\iint_{K^+(e_i,e_{i+1},\alpha z)} f(s,t)\,g(s)\,ds\,dt\,|d(\alpha z)/dz|^2$$
$$+ a_{i+1}\iint_{K^+(e_i,e_{i+1},\alpha^2 z)} f(s,t)\,g(s)\,ds\,dt\,|d(\alpha^2 z)/dz|^2\Bigg\},$$

https://doi.org/10.1017/S0305004100066688 Published online by Cambridge University Press

and

$$F_2(x,y) = \tfrac{1}{2}C(K) + \sum_{i=1}^{n} p_i a_i \Big\{ \quad -\sum_{j=i+1,\, i+2} a_j \iint_{CK^+(e_i,\, e_j,\, z)} f(s,t)\, g(s)\, ds\, dt$$
$$+ a_{i-1} \iint_{K^+(e_i,\, e_{i-1},\, \alpha z)} f(s,t)\, g(s)\, ds\, dt |d(\alpha z)/dz|^2$$
$$+ a_{i-1} \iint_{K^+(e_i,\, e_{i-1},\, \alpha^2 z)} f(s,t)\, g(s)\, ds\, dt |d(\alpha^2 z)/dz|^2 \Big\},$$

because all the other terms locally vanish, or are absorbed into $C(K)$. Here CK^+ denotes the complement of K^+ in the square $[0,1]^2$.

For each fixed i there are just eight non-vanishing terms associated with $P_1 \in e_i$ which are not constant, and among these the last four terms can be got from the first four terms. Moreover, the domains of the corresponding K^+'s are determined by $0 \leqslant s \leqslant 1$, $0 \leqslant t \leqslant 1$, and at most two other affine restrictions as in (3). It turns out that, in each one of the two summations occurring in our formulae for F_1 and F_2, the summand has the same functional form and depends only on $(a_{i-1}, a_i, a_{i+1}; \alpha_{i-1}, \alpha_i, \alpha_{i+1}, \alpha_{i+2})$. This greatly simplifies the calculation of $\tilde{m}(x,y)$.

After all these comments have taken into account, *we find that* $\tilde{m}(x,y)$ *in the basic tile is given by the general formula*

$$\tilde{m}(x,y) = C(K) + \frac{1}{32|K|^3} \sum_{i=1}^{n} p_i \{H_{i1}(x,y) + H_{i2}(x,y) + H_{i3}(x,y)\}, \tag{4}$$

where $C(K)$ is Small's constant, where

$$H_{i1}(x,y) = \sin(\alpha_i + \alpha_{i+1})\, H(x,y,a_i,\alpha_i,\alpha_{i+1}), \tag{5}$$

$$H_{i2}(x,y) = \sin\alpha_{i-1}\, H(x,y,a_{i-1},\alpha_{i-1},\alpha_i), \tag{6}$$

$$H_{i3}(x,y) = \sin\alpha_{i+2}\, H(x,y,a_{i+1},\alpha_{i+1},\alpha_{i+2}), \tag{7}$$

and

$$H(x,y,a,\gamma_1,\gamma_2) = a^5 RS(y\sqrt{3})^2 \{-\tfrac{2}{5} + \tfrac{4}{15}(y\sqrt{3})[R\cos\gamma_1 + S\cos\gamma_2]$$
$$- \tfrac{1}{15}(y\sqrt{3})^2 [R^2 + RS\cos(\gamma_1+\gamma_2) + S^2]\}. \tag{8}$$

Here

$$R = \frac{1}{(1+x\sqrt{3})\sin\gamma_1 + y\sqrt{3}\cos\gamma_1}, \quad S = \frac{1}{(1-x\sqrt{3})\sin\gamma_2 + y\sqrt{3}\cos\gamma_2}.$$

Formulae (4)–(8) constitute our main result. We shall next (in §3) examine a few interesting special cases, but first we note that (4)–(8) have been numerically checked against the output from the Fortran 77 algorithm described in [**4**]; this gave satisfactory agreement to 15 significant figures.

3. *Results*

We here discuss the results of our procedure in three important special cases.

3·1: K = square. Here the answer is already known to us from [**4**]; it is

$$\tilde{m}_1(x,y) = \tfrac{1}{4} - \tfrac{1}{40}(y\sqrt{3})^2 \left[\frac{1}{1-x\sqrt{3}} + \frac{1}{1+x\sqrt{3}}\right] - \tfrac{1}{240}(y\sqrt{3})^4 \left[\frac{1}{(1-x\sqrt{3})^3} + \frac{1}{(1+x\sqrt{3})^3}\right], \tag{9}$$

https://doi.org/10.1017/S0305004100066688 Published online by Cambridge University Press

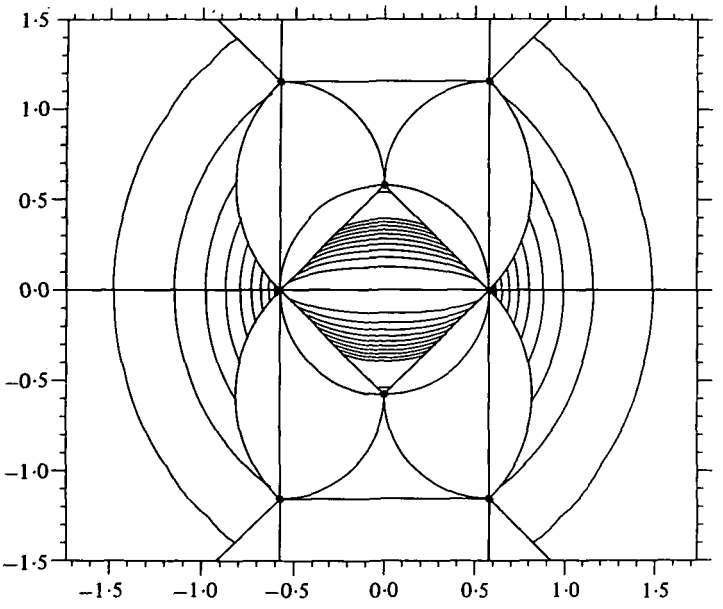


Fig. 2. The singular tessellation for K = rectangle ($s = 1$). Contours at the levels 0·1(0·1)0·9(0·01)0·99 for $\tilde{m}(x, y)/\tilde{m}(0, 0)$ are drawn in the basic tile and its equivalents.

and the reader who cares to work through the algebra can verify that our general formula does indeed reduce to the above. (This formula (9) is consistent with what we said about the structure of $\tilde{m}(x, y)$ at the end of §1, but it has here been collapsed to a simpler form.) It is obvious that here $\tilde{m}_1$ is a single real-analytic function in the union of the 'lower' and 'upper' basic tiles.

3·2: $K = (s \times 1)$ rectangle, with $s \geqslant 1$. This case is of considerable practical importance, and the corresponding shape-density has never been dealt with before. We find that

$$\tilde{m}_s(x, y) = \frac{s^3 + s^{-3}}{2}(\tilde{m}_1(x, y) - \tfrac{1}{4}) + \tfrac{1}{4}\frac{s + s^{-1}}{2}, \tag{10}$$

where $\tilde{m}_1$ is the shape-density for the square. Here again we have real analyticity on the basic segment.

Of course, (10) reduces to (9) when $s = 1$. The fact that the 'constant' term is proportional to $f = \frac{1}{2}(s + s^{-1})$ was proved by C. G. Small [5], and we now see that the remaining terms are proportional to

$$\frac{1}{2}\left(s^3 + \frac{1}{s^3}\right) = 4f^3 - 3f,$$

and so change much more rapidly when s departs from the 'square' value 1.

Figs. 2–4 show the singular tessellations when $s = 1$, 2, and 3, together with the contour lines for $\tilde{m}_s$ in the basic tiles and their re-labelling equivalents. The diagrams are very similar to one another, but note that the basic tile becomes flatter and flatter as s increases; in fact its (maximum) altitude at $x = 0$ is

$$y_{\max} = \frac{1}{s\sqrt{3}}, \tag{11}$$

https://doi.org/10.1017/S0305004100066688 Published online by Cambridge University Press

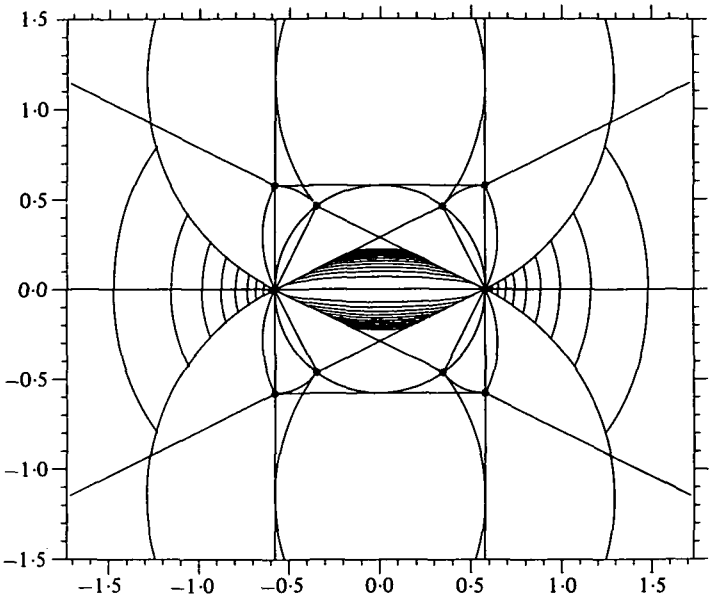


Fig. 3. The singular tessellation for K = rectangle ($s = 2$). Contours at the levels 0·1(0·1)0·9(0·01)0·99 for $\tilde{m}(x, y)/\tilde{m}(0, 0)$ are drawn in the basic tile and its equivalents.

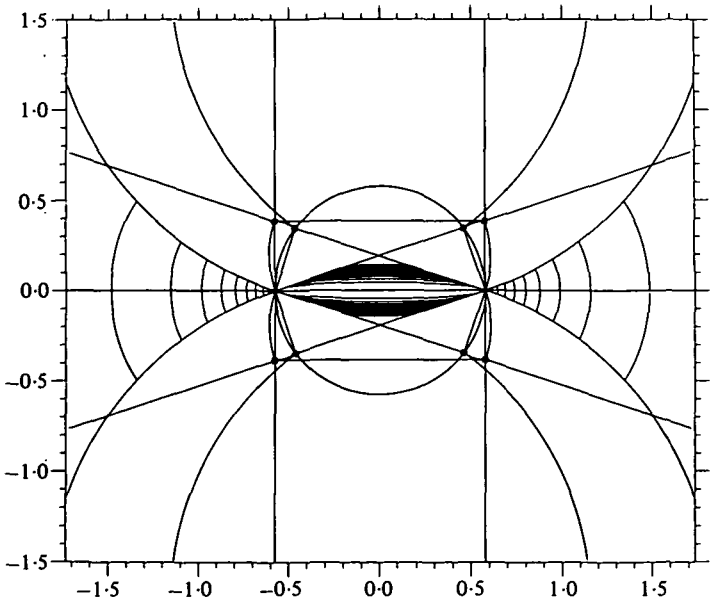


Fig. 4. The singular tessellation for K = rectangle ($s = 3$). Contours at the levels 0·1(0·1)0·9(0·01)0·99 for $\tilde{m}(x, y)/\tilde{m}(0, 0)$ are drawn in the basic tile and its equivalents.

https://doi.org/10.1017/S0305004100066688 Published online by Cambridge University Press

and the function $\tilde{m}_s(x,y)$ in the basic tile has its minimum value at $(0,y_{\max})$. This minimum value, which is of some considerable interest in the applications, is

$$\tilde{m}_s(0,y_{\max}) = \tfrac{1}{10}s+\tfrac{29}{240}s^{-1}-\tfrac{1}{40}s^{-5}-\tfrac{1}{240}s^{-7},$$

and when expressed as a fraction of the value at $(0,0)$, it is

$$\frac{V(K)}{C(K)} = \frac{24}{30}-\frac{\lambda^3+5\lambda^2-5\lambda}{30},$$

where $V(K)=\tilde{m}_s(0,y_{\max})$, $\lambda = s^{-2}$. Since $(\lambda^3+5\lambda^2-5\lambda)/30$ in $0<\lambda\leqslant 1$ has its maximum value when $\lambda = 1$, we have

$$\frac{V(K)}{C(K)} \geqslant \frac{23}{30},$$

(this bound being attained for K = square), so that in the basic tile, for K = rectangle,

$$\min_{x,y}\tilde{m}_s(x,y) \geqslant \text{about } \tfrac{3}{4}\max_{x,y}\tilde{m}_s(x,y) \quad \text{for all } s.$$

3·3: $K = n$-gon. This is also a new result; we find $\tilde{m}(x,y)$ to be

$$\tilde{m}(x,y) = C(K)-\frac{2}{15}\frac{(1+\cos\alpha)^3}{n^2\sin\alpha}RS \times\{6(y\sqrt{3})^2-8(y\sqrt{3})^3RS\sin\alpha\cos\alpha+(y\sqrt{3})^4[R^2-RS(1+6\cos^2\alpha)+S^2]\}, \quad (12)$$

where

$$\alpha = \pi-\frac{2\pi}{n}, \quad C(K) = \frac{1}{2n}\left(\frac{1}{\sin\alpha}+\frac{\sin\alpha}{1+\cos\alpha}\right)$$

and

$$R = \frac{1}{(1+x\sqrt{3})\sin\alpha+y\sqrt{3}\cos\alpha}, \quad S = \frac{1}{(1-x\sqrt{3})\sin\alpha+y\sqrt{3}\cos\alpha}.$$

Note that when $|x|$ is well away from $\pm 1/\sqrt{3}$, y is small, and n is not too large, then $\tilde{m}$ for fixed x is almost a quartic in y. In other cases the effect of the dependence of R and S on y will make itself felt. However, it is obvious that the real analyticity on the basic segment noted in 3·1 and 3·2 breaks down when $n \neq 4$.

As in 3·1 and 3·2 we have $\tilde{m}(x,y)=\tilde{m}(-x,y)$, but for general K this symmetry will be absent.

When $n\to\infty$, the basic tile for the regular n-gon collapses onto the basic segment, and the value of $\tilde{m}(x,y)$ inside it converges to C(circle).

I would like to thank Professor D. G. Kendall for his helpful guidance in this work as the supervisor of my postgraduate study in Cambridge.

https://doi.org/10.1017/S0305004100066688 Published online by Cambridge University Press

REFERENCES

[1] D. G. KENDALL. Shape-manifolds, procrustean metrics, and complex projective spaces. *Bull. London Math. Soc.* **16** (1984), 81–121.

[2] D. G. KENDALL. Exact distributions for shapes of random triangles in convex sets. *Adv. Appl. Prob.* **17** (1985), 308–329.

[3] D. G. KENDALL and W. S. KENDALL. Alignments in two-dimensional random sets of points. *Adv. Appl. Prob.* **12** (1980), 380–424.

[4] D. G. KENDALL and H.-L. LE. Exact shape-densities for random triangles in convex polygons. *Adv. Appl. Prob.* (in the Press).

[5] C. G. SMALL. Random uniform triangles and the alignment problem. *Math. Proc. Cambridge Philos. Soc.* **91** (1982), 315–322.

https://doi.org/10.1017/S0305004100066688 Published online by Cambridge University Press