Some problems in probability theory.

A thesis submitted for the
degree of Doctor of Philosophy
at the University of Oxford

by

Wilfrid S. Kendall.

Linacre College

1979

BODLEIAN LIBRARY OXFORD
DEPOSITED
THESIS
22-7-80

Abstract.

The thesis deals with three problems coming under the heading of probabilistic geometry. They are dealt with in three chapters.

The first chapter is concerned with the angles formed by triples of independent,identically distributed points on a plane. The work is part of an investigation carried out by several workers and motivated by an archaeological problem concerning supposed collinearities formed by ancient sites. In the symmetric Gaussian case the joint distribution is found of two of the angles of the triangle. An integral is given for the density of a named angle in the case of a general Gaussian parent distribution. The asymptotic probability that the triangle has one very obtuse angle is discussed together with the changes that it undergoes when the parent distribution is "stretched". Finally bounds are obtained for this asymptotic probability when the parent distribution is one of a family of circularly symmetric distributions.

The second chapter is about the way in which Brownian motion knots in 3-space. A new concept, that of a "knot-tube", is required because the self-intersecting character of Brownian

motion in 3-space means the usual topological notion of a knot can not be applied. It is known that Brownian motion in 5-space and above has very simple topological behaviour, but the 4-space case is unsettled. A partial result is given for 4-space.

The final chapter considers the topological behaviour of contours of two generalisations of Brownian motion to multidimensional time. The two generalisations are the Brownian sheet and the generalisation due to Lévy. Boundedness of the contours of the Lévy process is shown but no analogous result is known for the Brownian sheet. In both cases it is shown that almost surely the union of all non-trivial contours has zero measure. For topological reasons this union cannot be empty.

Acknowledgements.

I would like to thank my supervisor, Professor J.F.C.Kingman, for all his advice and encouragement during the preparation of this thesis. I would also like to thank my father D.G.Kendall for the pleasure of collaborating with him on the work represented in part by chapter one of this thesis. Finally I would like to thank the Science Research Council for their financial support.

Contents.

Introduction 1

Chapter 1 : Random Triangles and Collinearities.

1.1 Introduction. 4

1.2 The random Gaussian triangle. 8

1.3 Stretching of parent distributions. 25

1.4 Bounds on the collinearity constant. 43

1.5 Conclusion. 61

Chapter 2 : The knotting of Brownian motion.

2.1 Introduction. 65

2.2 Knots and implication. 68

2.3 The Brownian path. 80

2.4 Higher dimensions. 89

Chapter 3 : The contours of Brownian processes with multidimensional times.

3.1 Introduction. 93

3.2 Analysis of the Lévy process. 97

3.3 A topological interlude. 112

3.4 Analysis of the Brownian sheet. 118

3.5 Conclusion. 133

References for the three chapters and the introduction are on pages from 134

Figures.

1.1 To illustrate the definitions of V and W . 11

To illustrate the trigonometrical argument... 13

The density function of the maximum angle of a random Gaussian triangle. 21

The density function of an angle of a random Gaussian triangle. 22

1.2 The density functions of named angles of stretched Gaussian triangles. 42

1.3 Illustrating the existence of a limit. 51

2.2 Diagram of a knot-tube. 70

To illustrate the loops generating the fundamental group of the torus ∂A . 77

3.4 To illustrate the proof of a lemma. 125

To illustrate the decomposition of W . 128

Introduction

This work deals with three problems in what may be termed probabilistic geometry . None of the problems quite fall into the province of geometrical probability- however they are all concerned with questions of a geometric or topological nature about particular random objects.

The first chapter is about random points in the plane and in particular about the triangles formed by triples of these points and their angles. The work in this chapter was motivated by a problem in archaeology concerning the "leyline" hypothesis, as described in the introduction to the chapter itself. The results obtained form part of a body of work to be reported on in the paper of D.G.Kendall and W.S.Kendall (1979) . The exposition in this thesis confines itself to the results obtained by myself. The paper cited deals with the data-analytic aspects in a much more complete manner.

If the first chapter is about random geometry, being about the angles of random triangles, then the second chapter is on a subject in random topology. Motivated

by the conjectures of S.F.Edwards and D.Williams (see Edwards (1968) and Williams (1979)) an examination is made of the topological nature of the random map of $\mathbb{R}$ into $\mathbb{R}^3$ produced by Brownian motion in 3-space. Much of the work lies in formulating the question and a certain amount of algebraic topology is required to show that the new concept of "knot-tube" presented here bears some resemblance to the more usual topological notions of knotting. The case of Brownian motion in higher dimensional space is also considered. This has been dealt with by Milnor (1964) except in the case of dimension 4 which proves intractable. A partial result is given for dimension 4 .

Continuing the theme of random topology the third chapter considers the random maps from n-dimensional spaces to the real line provided by Brownian motion generalised in Lévy's sense (called "Lévy processes" in this chapter) and also by the Brownian sheets. The objects of interest here are the contours of the random maps. In the case of dimension 2 the contours can be understood in the cartographical sense. In general they are the maximal connected sets on which the random map is constant.

The question considered by this chapter is whether the contours are bounded. An affirmative answer is obtained in the case of the Lévy generalisation but the question is not settled in the case of the Brownian sheet. However in both cases an interesting pathology is shown to occur; for the union of the nontrivial contours makes up a set of zero Lebesgue measure.

As has been mentioned, the work of chapter one will be reported on in D.G.Kendall and W.S.Kendall (1979). Most of the work described in chapter two is expounded in W.S.Kendall (1979) .

Chapter 1 : Random triangles and Collinearities.

1.1 Introduction

This chapter describes theoretical work motivated by a statistical problem concerning archaeological data. The problem, as described in Broadbent (1979) and in Heaton (1976) , is that of testing the "leyline" hypothesis that particular archaeological remains were deliberately sited so as to lie along straight lines. Thus if the sites within a certain region lie at points $X_1, \ldots, X_n$ then the "leyline" hypothesis would be supported if a high degree of linear structure was observed in the ensemble of points $X_1, \ldots, X_n$.

One possible way of testing such a hypothesis would be to count the number of triples of sites $X_i\ X_j\ X_k$ near to being collinear. For example one could count the number of ε-blunt triples for small ε where $X_i\ X_j\ X_k$ forms an ε-blunt triple if the largest interior angle defined by the triangle $X_i\ X_j\ X_k$ is within ε of π when measured in radians. In the 1979 paper cited above

Broadbent adopts this approach and gives results of simulations carried out with Heaton to establish the expected number of Σ-blunt triples under a certain uniform hypothesis. Of course this is not the only possible avenue to explore; clearly it would be more satisfactory to count the number of m-tuples collinear in some sense and indeed Broadbent touches on this. However the technique of counting Σ-blunt triples has the virtue of being more susceptible to analysis. At this point it should be mentioned that Behrend (1978) has considered a different approach not involving angles.

Clearly for the analysis of Σ-bluntness it would be useful to have theoretical results on the number of Σ-blunt triples for small ε when the points are drawn independently from some distribution over the plane. If the parent distribution is uniform over a compact convex region, which is a natural first choice, then difficulties arise because of the boundary. The difficulties are both technical and also practical in that the location of the boundary must be chosen. Nevertheless limiting results as Σ tends to zero have

been obtained by D.G.Kendall (1979).

The boundary problems are eliminated if the parent distribution is taken to be Gaussian with circular symmetry. In this case the investigation of the angles of a Gaussian triangle (a triangle formed by Gaussian random points) becomes relevant. Thus the above considerations give rise to a motive for evaluating the joint distribution of the angles of a Gaussian triangle. An investigation of this distribution is described in section 1.2 .

The results of section 1.2 provide a theoretical reason for defining the collinearity constant (pages 24 and 26) corresponding to a given parent distribution. At least in the case of the Gaussian distribution with circular symmetry this constant can be used to provide an accurate approximation to the expected number of ε-blunt triples for small ε . Section 1.3 , by means of an analysis of the "stretched" Gaussian distribution (one of "elliptical symmetry"), helps to show the extent of the validity of this approximation. Occurring before this analysis in section 1.3 is the statement and proof of a theorem

that describes the behaviour of the collinearity constant of a quite general parent distribution as the distribution is stretched by means of a dilation of the plane.

The theorem of section 1.3 also provides conditions on the distribution sufficient for the collinearity constant to exist. Other sufficient conditions are provided in section 1.4 . Further, this section contains a theorem which provides bounds on the collinearity constant when the parent distribution has circular symmetry and is a "mixture of discs" (for definition see page 43).

Finally section 1.5 reviews these results and their application to the original "leyline" problem and gives references to various papers giving complementary results.

1.2 The random Gaussian triangle.

The argument of the introduction provides a motive for evaluating the probability

$$P\left\{ X_1\, X_2\, X_3 \quad \text{is} \quad \varepsilon\text{-blunt} \right\}$$

for random points X_1 , X_2 , X_3 independent and identically distributed with common distribution the circularly symmetric Gaussian distribution on the plane. The triangle $X_1\, X_2\, X_3$ is the random Gaussian triangle of the title of this section. The bluntness of the triangle is determined by the interior angles of the triangle and so this section will be dedicated to the analysis of the joint distribution of these angles. Such an analysis has an intrinsic interest as a part of geometrical probability.

To avoid ambiguity the interior angles will be measured throughout in radians as lying between 0 and π . So the objective of the section is to evaluate

$$P\left\{ \angle X_1\, X_2\, X_3 > \beta \quad \text{and} \quad \angle X_3\, X_1\, X_2 > \alpha \right\}$$

for all α and β in $(0,\pi)$ such that $\alpha + \beta \leq \pi$. Since geometry plays an important part in this investigation it is helpful to note that the probability of the triangle $X_1\, X_2\, X_3$ being degenerate is zero and so the possibility of degeneracy

can be ignored.

Given the vectors X_1 , X_2 , X_3 three new vectors are defined, each of straightforward geometrical interpretation. These are

the midpoint $M = \frac{1}{2}(X_1 + X_2)$ of the segment $X_1 X_2$;

a half-segment of this line

$$Y = \frac{1}{2}(X_1 - X_2) \quad ;$$

the vector pointing from the midpoint M to the remaining vertex X_3

$$Z = X_3 - M \quad .$$

These new vectors are jointly Gaussian. Indeed Y and Z are independent, as may be seen from calculation of the covariances of their coordinates. Also the three of them together define the angles and hence the shape of the random Gaussian triangle.

In fact as far as specifying the angles is concerned it suffices to know the angle Θ between Y and Z , say measured in radians counterclockwise from Z , and also the ratio between the lengths of Y and Z , or for convenience the related number $\| Z \|^2/(3. \| Y \|^2) = T$. The joint distribution of Θ and T can be found by techniques commonly used in statistical distribution theory. Because Y and Z are independent and because their distributions both have circular

symmetry the angle Θ is independent of the lengths of Y and Z and hence of T ; moreover it is uniformly distributed over $[0,2\pi)$. Because of the independence of Y and Z and because of their symmetric Gaussian distributions the ratio of the squares of their lengths has the distribution of a multiple of an $F(2,2)$ statistic and in fact T has precisely the distribution of an $F(2,2)$ statistic. Consequently the joint density of Θ and T is the function

$$(\theta , t) \mapsto \frac{1}{2\pi . (1 + t)^2} \qquad \text{for } \theta \text{ in } [0,2\pi) \text{ and } t \text{ in } (0,\infty) .$$

The next step in the argument is to relate the behaviour of Θ and T to the values of the angles $\angle X_1 X_2 X_3$ and $\angle X_3 X_1 X_2$. To test whether these angles are greater than β and α respectively it is convenient to define further points V and W chosen to be on opposite sides of the line $X_1 X_2$ and such that $\angle W X_1 X_2 = \angle V X_1 X_2 = \alpha$ while $\angle W X_2 X_1 = \angle V X_2 X_1 = \beta$. Since the triangle may be assumed to be non-degenerate there are precisely two such points and since we will not need to distinguish between them the ambiguity as to which is which does not matter. For the sake of precision take W to be the one of the two lying to the right of X_2 viewed from just above X_1 . A diagram is given to illustrate these definitions.

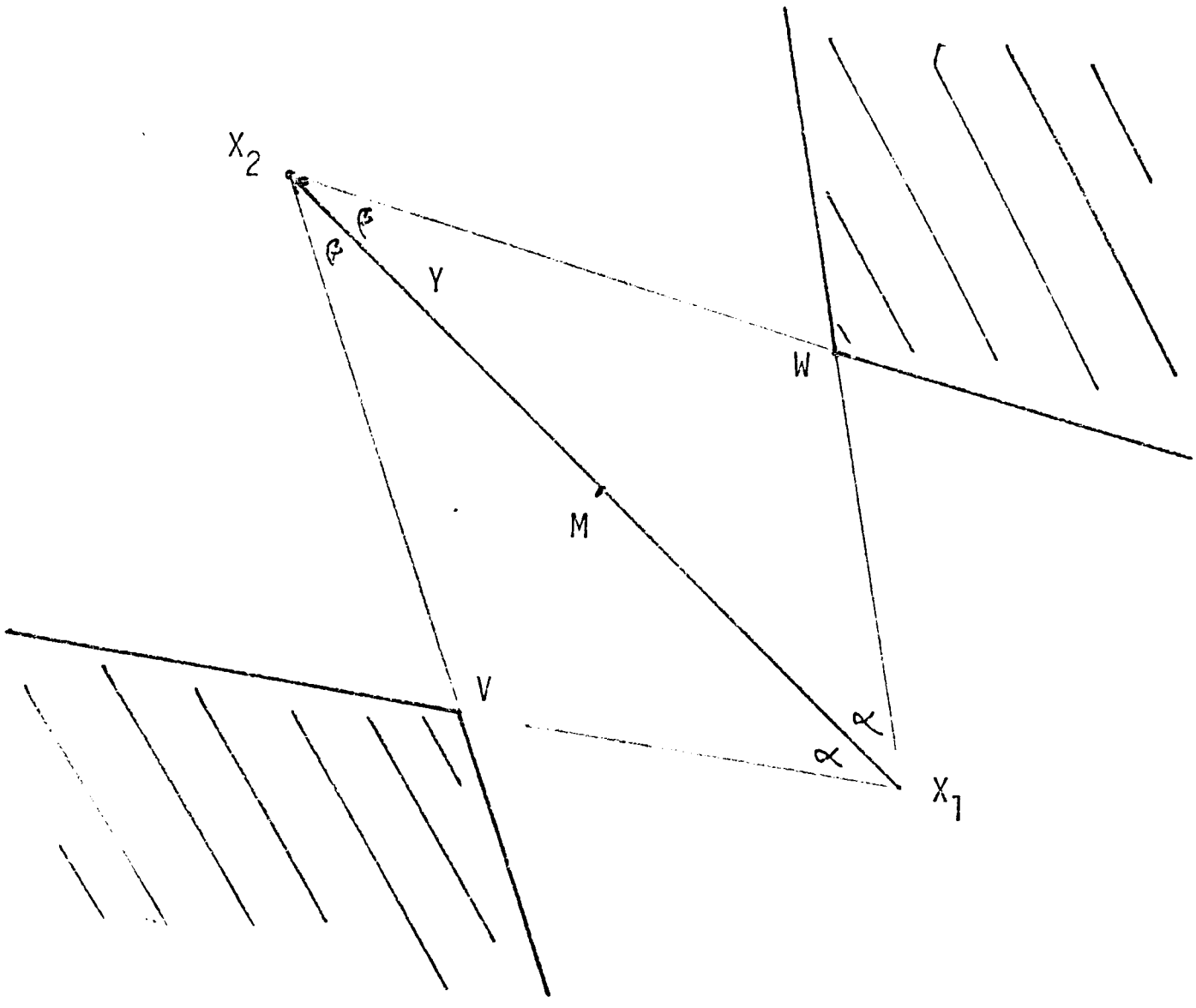


Figure to be attached after page 10 , section 1.2 .

This illustrates the definitions of the points V , W using the original points X_1 , X_2 . The point M and the vector Y are also shown but the point X_3 and the vector Z are omitted as they are not used to define V and W .

The shaded regions shown as wedges pointed on V and W are those used to indicate the angles of X_1 X_2 X_3 by indicating whereabouts of X_3 .

Using these definitions the event

$$\{ \angle X_1\ X_2\ X_3 > \beta \text{ and } \angle X_3\ X_1\ X_2 > \alpha \}$$

is the same event as

$$\{ X_3 \text{ is in one of the wedges } C_1 , C_2 \}$$

where the wedges are the shaded regions shown in the figure attached on page 11 . More formally the regions C_1 , C_2 are defined using X_1 , X_2 , V , W by

$$C_1 = \{ V + \lambda(V - X_1) + \mu(V - X_2) : \lambda, \mu \geq 0 \} ;$$

$$C_2 = \{ W + \lambda(W - X_1) + \mu(W - X_2) : \lambda, \mu \geq 0 \} .$$

These two wedges are disjoint. They can be described by giving the distances of the vertices V and W from M and by giving their orientation, summarised in the angles subtended by $X_1\ M$ and $X_2\ M$ on V and W respectively. These parameters can be found using some trigonometry. For suppose

$$\alpha' = \angle X_1\ V\ M ;$$

$$\beta' = \angle X_2\ V\ M ;$$

$$R = \| V - M \| / \| Y \| ;$$

h is the perpendicular height of V from $X_1\ X_2$.

These definitions are illustrated in the attached figure (page 13). By the sine formula

$$R \sin \alpha' = \sin \alpha ;$$

$$R \sin \beta' = \sin \beta .$$

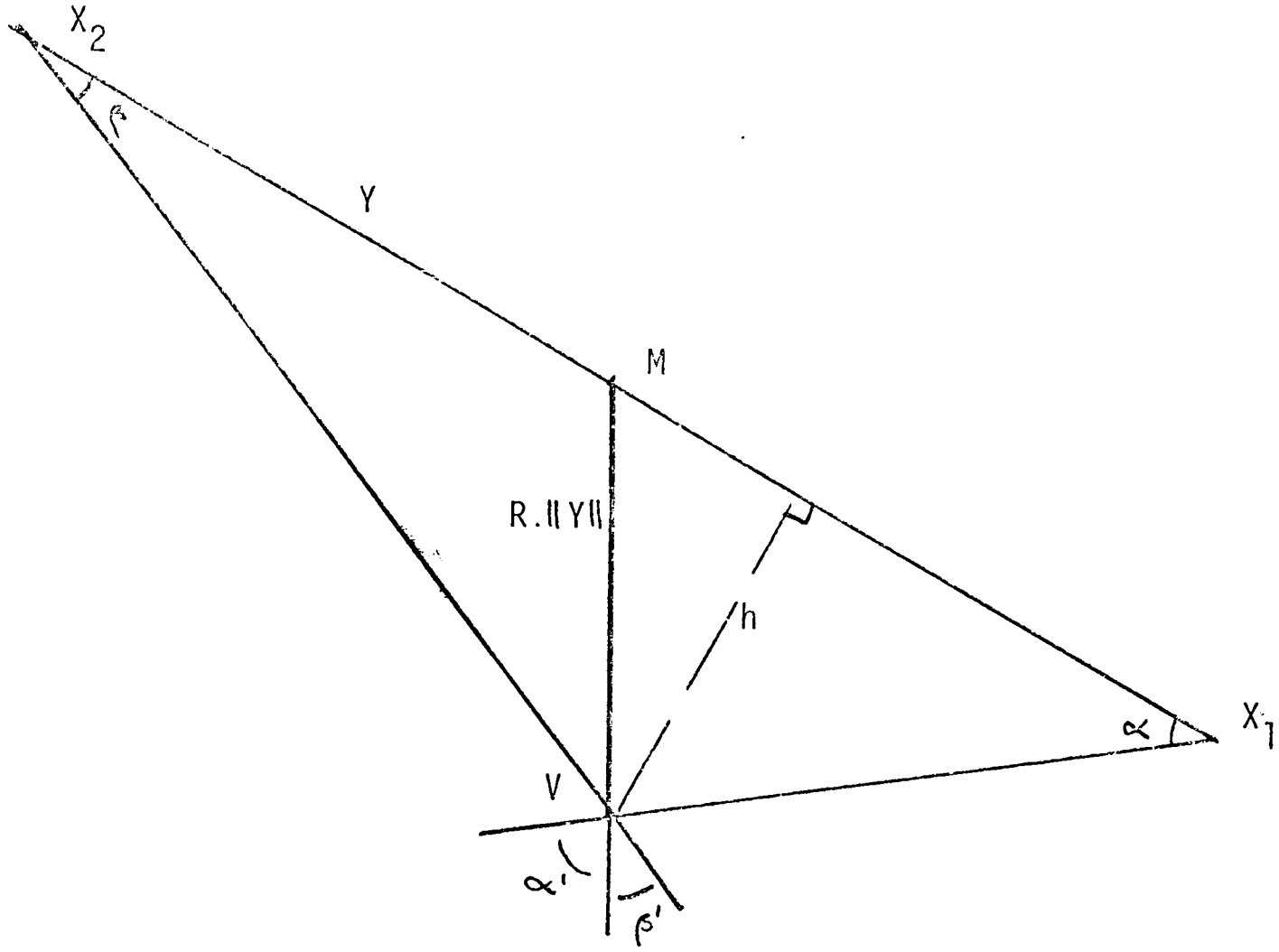


Figure to be attached after page 12 , section 1.2.

This illustrates the trigonometrical argument used to evaluate the angles α', β' and the side-ratio R .

Since $h \cot\beta - \|Y\| = \|Y\| - h \cot\alpha$

it follows that

$$h.(\cot\alpha + \cot\beta) = 2.\|Y\| .$$

An application of Pythagoras' theorem yields

$$(h \cot\beta - \|Y\|)^2 + h^2 = R^2 . \|Y\|^2$$

and a substitution of the previous equality produces an evaluation of R by

$$(\cot\beta - \cot\alpha)^2 + 4 = R^2.(\cot\alpha + \cot\beta)^2 .$$

So the parameters can be found using this last equality, the two equalities produced by use of the sine formula and the obvious

$$\alpha + \beta + \alpha' + \beta' = \pi .$$

With these parameters known the probability

$$P\{X_3 \text{ belongs to one of } C_1 , C_2\}$$

can be written as a double integral. For $X_3 - M = Z$ is equal to $(\|Y\| . \sqrt{3T} , \Theta)$ when referred to polar coordinates. So, by the disjointness of C_1 , C_2 and the uniformity of the distribution of Θ,

$$P\{X_3 \text{ belongs to one of } C_1 , C_3\}$$
$$= 2.P\{(\sqrt{3T} , \Theta) \text{ and the origin subtend an angle on } (R, 0) \text{ of between } \pi+\alpha' \text{ and } \pi-\beta' .\}$$

which clearly can be written as a double integral involving the density of T and Θ as given above.

If $q(\gamma, R)$ is the probability that $(\sqrt{3T}, \Theta)$ and the origin subtend an angle on $(R, 0)$ of between $\pi - \gamma$ and π then

$$P\{X_3 \text{ belongs to one of } C_1, C_2\} = 2.[q(\alpha', R) + q(\beta', R)].$$

So it suffices to find $q(\gamma, R)$ for all γ and R.

Using the joint density given above for T and Θ,

$$q(\gamma, R) = \int_0^{\gamma}\int_0^{\infty} \frac{dt\, d\theta}{2\pi.(1+t)^2} \cdot 1\{\sqrt{3t} > R \sin\gamma / \sin(\gamma - \theta)\}.$$

Changing variables by $3t = u^2$ the above is equal to

$$= \int_0^{\gamma}\int_{R\sin\gamma/\sin(\gamma-\theta)}^{\infty} \frac{6u\, du\, d\theta}{2\pi.(3+u^2)^2}$$

$$= \int_0^{\gamma} \frac{3}{2\pi} \cdot \frac{d\theta}{(3 + R^2\sin^2\gamma/\sin^2(\gamma-\theta))}$$

$$= \frac{3}{2\pi}\int_0^{\tan\gamma} \frac{v^2\, dv}{(3v^2 + R^2\sin^2\gamma.(1+v^2)).(1+v^2)}$$

using the new variable $v = \tan(\gamma - \theta)$.

This last integral can be evaluated by use of partial fractions. If the convention is made that arctangent takes values

between 0 and π then it follows from above that

$$q(\gamma, R) = \frac{1}{2\pi} \cdot \left\{ \gamma - \frac{R\sin\gamma}{(3 + R^2\sin^2\gamma)^{\frac{1}{2}}} \cdot \left[\tfrac{1}{2}\pi - \tan^{-1} \frac{R\cos\gamma}{(3+R^2\sin^2\gamma)^{\frac{1}{2}}} \right] \right\}$$

As noted above, the probability that X_3 belongs to one of C_1, C_2 can be given in terms of q. So it is the case that the probability

$$P \{ \angle X_1 X_2 X_3 > \beta \text{ and } \angle X_3 X_1 X_2 > \alpha \}$$

$$= \frac{1}{\pi} \left\{ \alpha' + \beta' - \frac{R\sin\alpha'}{(3 + R^2\sin^2\alpha')^{\frac{1}{2}}} \cdot \left[\tfrac{1}{2}\pi - \tan^{-1} \frac{R\cos\alpha'}{(3 + R^2\sin^2\alpha')^{\frac{1}{2}}} \right] - \frac{R\sin\beta'}{(3 + R^2\sin^2\beta')^{\frac{1}{2}}} \cdot \left[\tfrac{1}{2}\pi - \tan^{-1} \frac{R\cos\beta'}{(3 + R^2\sin^2\beta')^{\frac{1}{2}}} \right] \right\}$$

where R, α', β' are given by

$$R^2.(\cot\alpha + \cot\beta)^2 = 4 + (\cot\alpha - \cot\beta)^2 ;$$

$$R\sin\alpha' = \sin\alpha ;$$

$$R\sin\beta' = \sin\beta ;$$

$$\alpha + \beta + \alpha' + \beta' = \pi .$$

Note that there is no difficulty in defining R, α', β' by these equations since α and β lie in $(0,\pi)$ so that none of

the left-hand sides misbehave.

This formula gives the joint distribution function of two angles of the random Gaussian triangle. It can be used to yield the distribution of the maximum angle, which is what is required to answer questions about the probability of ε-bluntness. Geometrical considerations show that, when $\gamma > \pi/3$,

$$P\left\{\begin{array}{l}\text{the largest interior angle of } X_1 X_2 X_3 \text{ is} \\ \text{greater than } \gamma \ .\end{array}\right\}$$

$$= 3.P\{ \angle X_1 X_2 X_3 > \gamma \} - 3.P\left\{\begin{array}{l} \angle X_1 X_2 X_3 > \gamma \\ \text{and } \angle X_3 X_1 X_2 > \gamma \end{array}\right\}$$

and these two probabilities can be derived from the formula for the joint distribution.

The first probability is obtained from the formula by taking the limiting case $\alpha = \gamma$ and β tends to zero. In the limit

$$R^2 = 1 \ ;$$

$$\sin \alpha' = \sin \gamma \ ;$$

$$\sin \beta' = 0 \ ;$$

$$\alpha + \beta' + \alpha' = \pi \ .$$

Therefore $R = 1$, $\beta' = 0$, $R \sin \alpha' = \sin \alpha$ and so the probability is given by

$$P\{ \angle X_1 X_2 X_3 > \gamma \} = \frac{1}{\pi}\left\{\pi - \gamma - \frac{\sin \gamma}{(3 + \sin^2 \gamma)^{\frac{1}{2}}} \cdot \left[\tfrac{1}{2}\pi + \tan^{-1} \frac{\cos \gamma}{(3 + \sin^2 \gamma)^{\frac{1}{2}}}\right]\right\}.$$

The second probability can be found from the formula

by putting $\alpha = \beta = \gamma$, obtaining

$$R = \tan \gamma \ ;$$

$$\alpha' = \beta' = \tfrac{1}{2}\pi - \gamma \ .$$

Thus

$$P\left\{ \angle X_1 X_2 X_3 > \gamma \text{ and } \angle X_3 X_1 X_2 > \gamma \right\}$$

$$= \frac{1}{\pi}\cdot\left\{\pi - 2\gamma - \frac{2\sin\gamma}{(3+\sin^2\gamma)^{\frac{1}{2}}}\cdot\left[\tfrac{1}{2}\pi - \tan^{-1}\frac{\sin\gamma\tan\gamma}{(3+\sin^2\gamma)^{\frac{1}{2}}}\right]\right\}.$$

The formula for the maximum angle obtained from these two formulae can be simplified by noting that

$$\tfrac{1}{2}\pi + \tan^{-1}\frac{\cos\gamma}{(3+\sin^2\gamma)^{\frac{1}{2}}}$$

$$= 2.\tan^{-1}\sqrt{\frac{2+\cos\gamma}{2-\cos\gamma}} \quad \text{when } \gamma \text{ belongs to } (0,\pi)$$

and that

$$\tan^{-1}\sqrt{\frac{2+\cos\gamma}{2-\cos\gamma}} + \tan^{-1}\frac{\sin\gamma\tan\gamma}{(3+\sin^2\gamma)^{\frac{1}{2}}} - \tfrac{1}{2}\pi$$

$$= \tan^{-1}\left(\sqrt{\frac{2+\cos\gamma}{2-\cos\gamma}}\cdot\frac{1-2\cos\gamma}{1+2\cos\gamma}\right)$$

when γ belongs to $(0, \tfrac{1}{2}\pi)$.

The distribution for the maximum angle is displayed overleaf. Note that a simplified formula can also be obtained for the distribution for one of the angles alone. This formula follows the formula for the maximum angle distribution.

Distribution for the maximum angle.

$$P\left\{\text{the largest interior angle is greater than } \gamma\right\}$$

$$= \begin{cases} 1 & \text{if } \gamma \text{ belongs to } [0, \pi/3) \ ; \\ \frac{3}{\pi}\cdot\left[\gamma - \frac{2\sin\gamma}{(3+\sin^2\gamma)^{\frac{1}{2}}}\cdot\tan^{-1}\left(\frac{1-2\cos\gamma}{1+2\cos\gamma}\cdot\sqrt{\frac{2+\cos\gamma}{2-\cos\gamma}}\right)\right] & \\ \qquad \text{if } \gamma \text{ belongs to } [\pi/3, \pi/2) \ ; & \\ \frac{3}{\pi}\cdot\left[\pi - \gamma - \frac{2\sin\gamma}{(3+\sin^2\gamma)^{\frac{1}{2}}}\cdot\tan^{-1}\sqrt{\frac{2+\cos\gamma}{2-\cos\gamma}}\right] & \\ \qquad \text{if } \gamma \text{ belongs to } [\pi/2, \pi) \ . \end{cases}$$

Distribution for the angle $\angle X_1 X_2 X_3$.

$$P\left\{\angle X_1 X_2 X_3 > \gamma\right\}$$

$$= \frac{1}{\pi}\cdot\left[\pi - \gamma - \frac{2\sin\gamma}{(3+\sin^2\gamma)^{\frac{1}{2}}}\cdot\tan^{-1}\sqrt{\frac{2+\cos\gamma}{2-\cos\gamma}}\right]$$

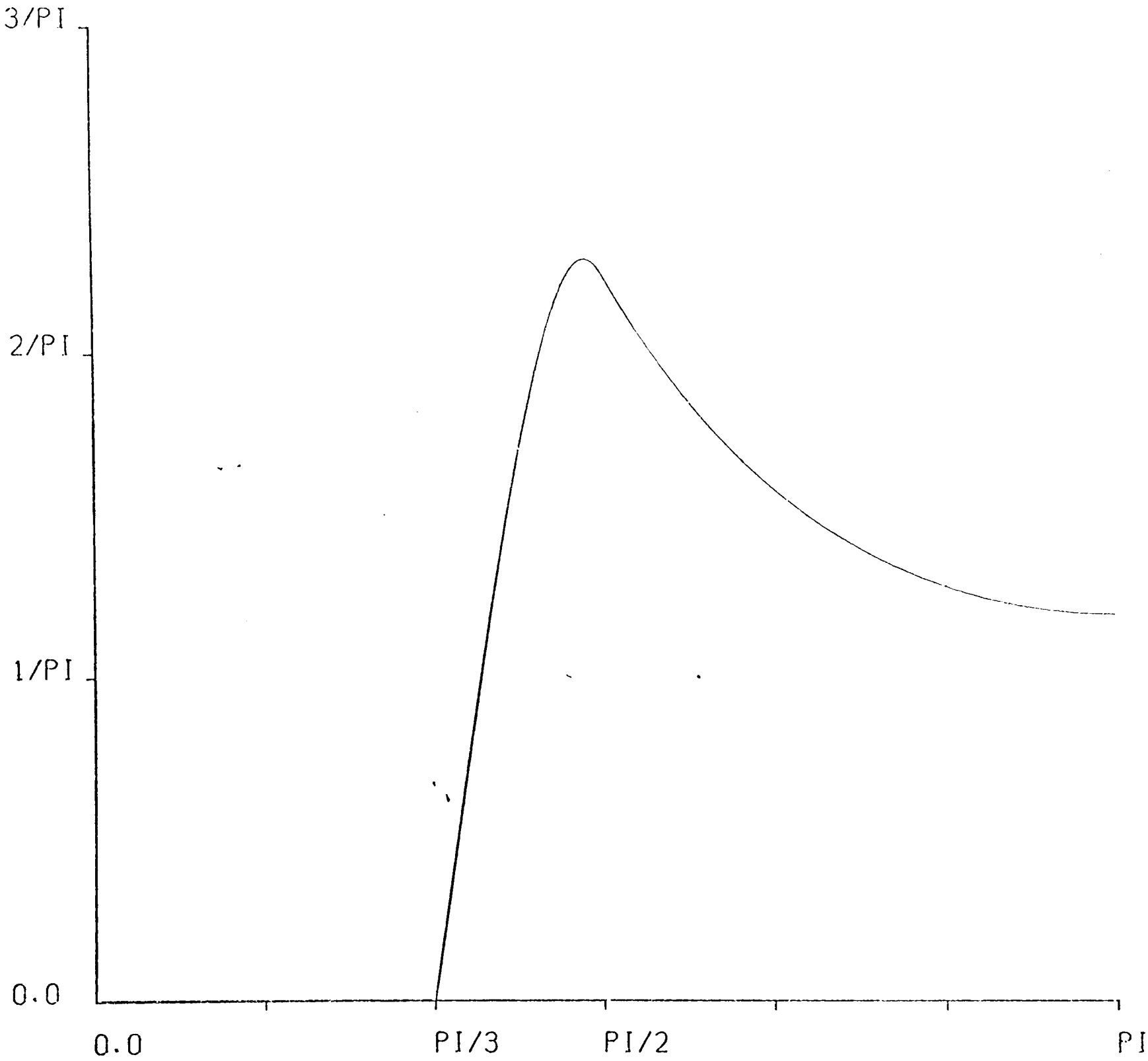
for γ belonging to $(0, \pi)$.

It is clear that the two distributions given above are both at least piece-wise once differentiable and thus possess density functions.

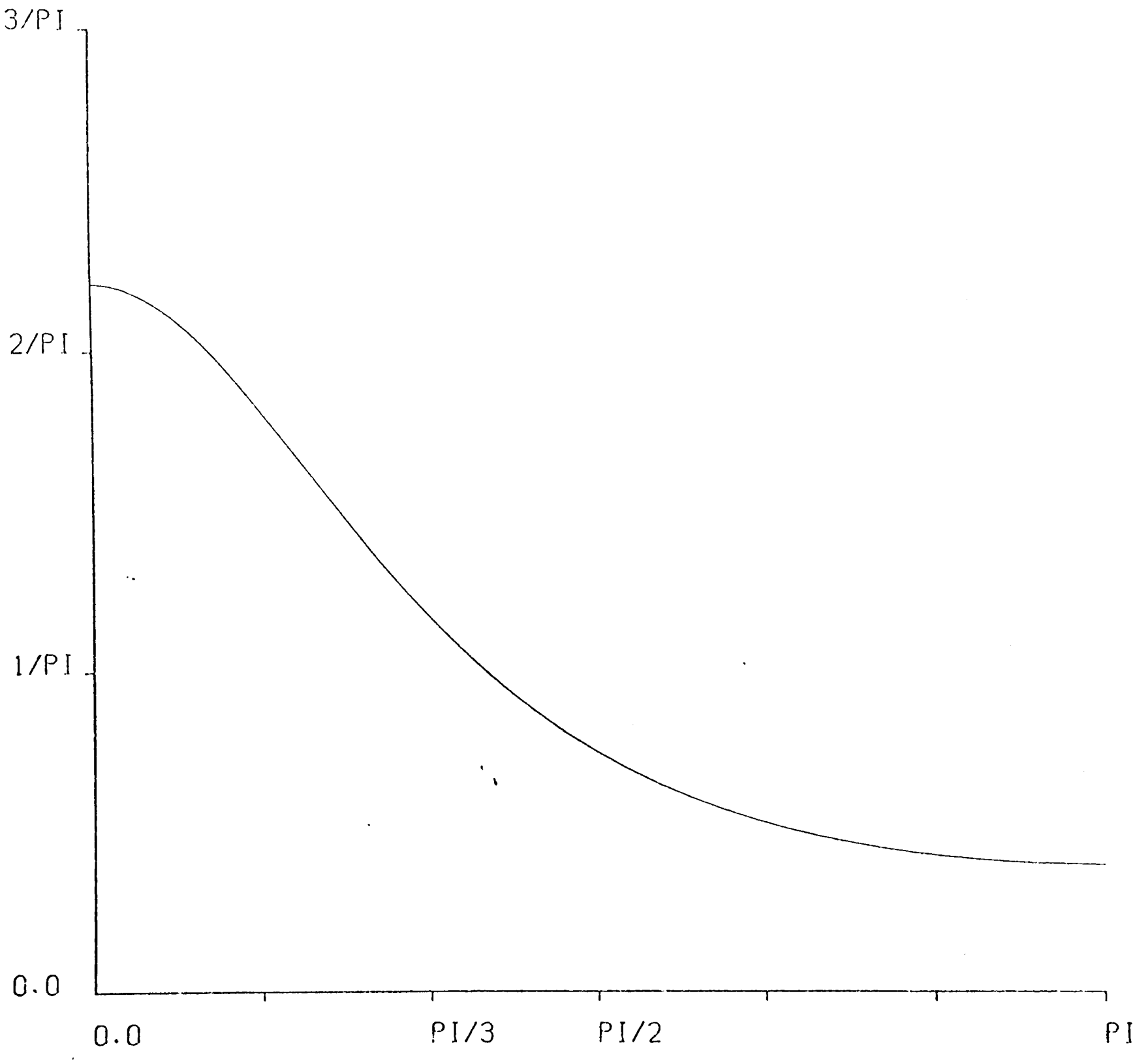
A straightforward computation shows that the density for the maximum angle angle of the triangle is the function of γ given by

$$f(\gamma) = \begin{cases} 0 & \text{if } \gamma \text{ belongs to } [0, \pi/3) \ ; \\ \dfrac{9}{\pi}\cdot\dfrac{1}{(3+\sin^2\gamma)^{\frac{1}{2}}}\cdot\left[\dfrac{1-4\cos^2\gamma}{1+2\cos^2\gamma} + \dfrac{2\cos\gamma}{(3+\sin^2\gamma)^{\frac{1}{2}}}\tan^{-1}\left(\dfrac{1-2\cos\gamma}{1+2\cos\gamma}\sqrt{\dfrac{2+\cos\gamma}{2-\cos\gamma}}\right)\right] & \\ & \text{if } \gamma \text{ belongs to } [\pi/3, \pi/2) \ ; \\ \dfrac{9}{\pi}\cdot\dfrac{1}{(3+\sin^2\gamma)^{\frac{1}{2}}}\cdot\left[1 + \dfrac{2\cos\gamma}{(3+\sin^2\gamma)^{\frac{1}{2}}}\cdot\tan^{-1}\sqrt{\dfrac{2+\cos\gamma}{2-\cos\gamma}}\right] & \\ & \text{if } \gamma \text{ belongs to } [\pi/2, \pi) \ . \end{cases}$$

Attached figures graph this density as a function over the range $[0, \pi)$ and also the density of the angle $\angle X_1 X_2 X_3$ over the same range. As observation of the graph suggests, the change in functional form at $\pi/2$ of the density of the maximum angle occurs without imposing a discontinuity on either f or the derivative of f. However there is a discontinuity in the second derivative of the density, occurring at $\pi/2$. As γ increases through $\pi/2$ the second derivative $f''(\gamma)$ jumps by $90/\pi$.



THE DENSITY FUNCTION OF THE MAXIMUM ANGLE OF A RANDOM GAUSSIAN TRIANGLE (MEASURED IN RADIANS).



THE DENSITY FUNCTION OF AN ANGLE
OF A RANDOM GAUSSIAN TRIANGLE
(MEASURED IN RADIANS).

The graph of the density of the maximum angle strongly suggests that the distribution is unimodal with maximum value of the density being obtained at γ just less than $\frac{1}{2}\pi$. However this has not been proved analytically.

It is possible to show that the density f is monotonically decreasing over the range $[\frac{1}{2}\pi, \pi]$. For the derivatives f' and f'' can be computed and it can be shown that $f'(\pi) < 0$ and $f''(\gamma) < 0$ for γ in $[\frac{1}{2}\pi, \pi]$. Consequently $f'(\gamma) < 0$ for all γ in $[\frac{1}{2}\pi, \pi]$.

The above suggests that $f(\pi)$ might be a good approximation for $f(\gamma)$ for γ near to π . This suggestion is borne out by an examination of the graph of f . In fact the approximation

$$P\{\text{maximum angle is greater than } \pi - \varepsilon\}$$
$$\simeq f(\pi).\ \varepsilon/\pi = (3/\pi - 1/\sqrt{3}).\ \varepsilon/\pi$$

is valid for quite large values of ε . The error in the approximation is less than 1% of the true value for ε as large as 10^{o} expressed in degrees (this follows from the monotonicity of f over $[\frac{1}{2}\pi, \pi]$ and a computation of $f(10.\pi/180)$.).

This approximation appears to hold good for random

triangles drawn from distributions other than the Gaussian distribution of circular symmetry. In the case of the distribution uniform on a square this was noticed by Broadbent (1979) and Heaton (1976) when considering simulation results. In section 1.3 the case of a Gaussian distribution of elliptical symmetry will be considered and from a graph of the corresponding density for the maximum angle a similar conclusion will be drawn. Clearly the value of the density of the maximum angle at π deserves a name; the first collinearity constant.

1.3 Stretching of parent distributions.

The application of the results of the preceding section to the collinearity problem is clear. For if the planar points $X_1, X_2, \ldots, X_n$ are independent and drawn from the same Gaussian distribution of circular symmetry then, if N_ε is the number of ε-blunt triangles formed by these points,

$$E(N_\varepsilon) = \binom{n}{3} P\{X_1 X_2 X_3 \text{ is } \varepsilon\text{-blunt}\}$$

$$= \binom{n}{3} . (3/\pi - 1/\sqrt{3}) . \varepsilon + o(\varepsilon)$$

as ε tends to zero. The closing remarks of the last section make it plain that the approximation will be accurate to 2 significant figures for ε as large as the equivalent of 10^0 in radians.

This section and the following section study the way in which the expected number of ε-blunt triangles changes when one alters the parent distribution for the sample of planar points $X_1, X_2, \ldots, X_n$. This will be done by considering the collinearity constant for the appropriate distribution, a number defined analogously to the collinearity constant for the Gaussian distribution. The observations of Broadbent and Heaton, and

graphs made of the density of the maximum angle for the random elliptical Gaussian triangle derived in this section, make it plausible that the limiting value which is to be the collinearity constant does provide a good approximation for the behaviour of $E(N_\varepsilon)$.

The first part of this section is concerned with the way in which the collinearity constant changes when the parent distribution undergoes "stretching" . In the Gaussian case this corresponds to studying the elliptic Gaussian distribution. In general suppose a sample of points $X_1, X_2, \ldots , X_n$ are drawn independently from a distribution on the plane of density the function $g(x) = g(x_1, x_2)$. As in the Gaussian case the expectation $E(N_\varepsilon)$ of the number N_ε of ε-blunt triangles is of interest, and in particular we define the (first) collinearity constant $\lambda(g)$ of the parent density g to be the limit

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} E(N_\varepsilon) \Big/ \binom{n}{3} = \lambda (g) .$$

if it exists.(Of course this definition is consistent with that made at the end of section 1.2 for the Gaussian case.) The action of stretching the parent distribution by a factor of s in the direction of the unit vector $v = (v_1 , v_2)$ is that of replacing the old density by

a new density function

$$g_s : x \longmapsto s.\, g(\,(s(x.v) - x.v)v + x\,) \quad .$$

Such a change of measure is equivalent to retaining the old measure but considering the sample

$$S_s X_1 \,,\, S_s X_2 \,,\, \ldots \,,\, S_s X_n$$

where S_s is the linear dilation that multiplies vectors by s if they are parallel to v and that leaves vectors unchanged if they are perpendicular to v . The following discussion will consider the behaviour of $\lambda(\, g_s \,)$ as a function of s and of course the axis of dilation v .

In the work referred to above Broadbent and Heaton have made an empirical study of this dependence on s for the case when g is uniform over a square. Their simulation results suggested that

$$\lambda\,(\, g_s \,) \quad = \quad \tfrac{1}{2}(\, s + 1/s \,).\, \lambda(\, g \,) \quad ,$$

a result subsequently proved by D.G.Kendall (1979) who has generalised it to a formula giving the dependence for any density uniform over a compact convex region. The theorem below provides a yet more general result that includes the case of the Gaussian density.

Theorem

Suppose the density g is bounded, continuous almost everywhere and such that the integral

$$\int_0^\infty \int_0^\infty (x_1^2 + x_2^2) \cdot g(x_1, x_2) \, dx_1 \, dx_2$$

is finite.

Then the limit $\lambda (g_s)$ exists for $s > 0$, moreover $\lambda (g_s) = \frac{1}{2}(A/s + Bs) \lambda(g)$ for two constants A and B such that $\frac{1}{2}(A + B) = 1$.

In certain cases $A = B$. Sufficient conditions for this equality to hold are

either SC 1 : the density is invariant under rotation through $\frac{1}{2}\pi$ around a certain point;

or SC 2 : the density is invariant under reflection through an axis making an angle of $\pi/4$ to the axis of dilation.

Remark

Note that the first sufficient condition is not dependent on the choice of axis of dilation. Therefore when the density is invariant under rotation through $\frac{1}{2}\pi$, thus having the symmetry of a square, the dilation factor

is $\frac{1}{2}(s + 1/s)$ for any axis of dilation. Note also one may assume the axis of dilation to be parallel to $x_2 = 0$.

Proof of the theorem.

Suppose that X_1, X_2, X_3 are independent points on the plane distributed with common density g_s . Then

$$\lambda(g_s) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} P\left\{ \begin{array}{l} \text{the largest angle of } X_1 X_2 X_3 \\ \text{is within } \varepsilon \text{ of } \pi \end{array} \right\}$$

in the sense that if one side exists then the other side exists and the two sides are equal.

But for our purposes it suffices to consider the distribution of a single angle;

$$\lambda(g_s) = \lim_{\varepsilon \to 0} \frac{3}{\varepsilon} P\{ \angle X_1 X_2 X_3 \text{ is within } \varepsilon \text{ of } \pi \}.$$

Note too that because X_1, X_2, X_3 have densities the line $X_1 X_3$ is almost surely well-defined .

Let the perpendicular distance from X_2 to the line $X_1 X_3$ be a and let the distance between $\frac{1}{2}(X_1 + X_3)$ and the perpendicular projection of X_2 onto $X_1 X_3$ be equal to

$$\| \tfrac{1}{2}(X_1 - X_3) \| \, . b \quad .$$

Then the angle $\angle X_1 X_2 X_3$ is within ε of π exactly when

$$b \leq 1 \quad \text{and} \quad a < (\sqrt{\operatorname{cosec}^2 \varepsilon - b^2} - \cot \varepsilon) . \| \tfrac{1}{2}(X_1 - X_3) \| \, .$$

$$\text{Writing} \quad t(b,\varepsilon) = \sqrt{(\operatorname{cosec}^2\varepsilon - b^2)} - \cot\varepsilon$$
$$= (\sqrt{(1 - b^2\sin^2\varepsilon)} - \cos\varepsilon)/\sin\varepsilon$$

it follows that

$$P\{\angle X_1 X_2 X_3 \text{ is within } \varepsilon \text{ of } \pi\}$$
$$= C_\varepsilon$$
$$= E\left\{\int_{-1}^{1}\int_{-t(b,\varepsilon)}^{t(b,\varepsilon)} \cdot \|\tfrac{1}{2}(X_1 - X_3)\|^2 \cdot \right.$$
$$\left. \cdot g_s\left[b\cdot\tfrac{1}{2}(X_1 - X_3) + \tfrac{1}{2}(X_1 + X_3) + \|\tfrac{1}{2}(X_1 - X_3)\|(u\cdot Y)\right] \cdot du\, db\right\}$$

where Y is the random vector of unit length making an angle of $\frac{1}{2}\pi$ with $\frac{1}{2}(X_1 - X_3)$ when the angle is measured counterclockwise. This follows from considering the region of the plane in which X_2 must fall if $\angle X_1 X_2 X_3$ is to be within ε of π and if X_1, X_3 are given.

In this investigation the above formula is of interest for small ε and for such ε an approximation can be found for $t(b,\varepsilon)$. By the binomial theorem and the fact that $\lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\sin\varepsilon = 1$ it follows that

$$t(b,\varepsilon)/\varepsilon \to \tfrac{1}{2}(1 - b^2) \quad \text{uniformly for } |b| \leq 1.$$

Using this limit, the boundedness of g_s and the fact that

the integral $\int_0^\infty \int_0^\infty (x_1^2 + x_2^2) \cdot g(x_1, x_2) \, dx_1 \, dx_2$

is finite (as stated in the theorem) imply that

$$\lim_{\varepsilon \to 0} C_\varepsilon/\varepsilon = \lim_{\varepsilon \to 0} \tfrac{1}{2} E \left\{ \int_{-1}^{1} \int_{-\varepsilon(1-b^2)/2}^{\varepsilon(1-b^2)/2} \cdot \| \tfrac{1}{2}(X_1 - X_3) \|^2 \cdot \right.$$

$$\left. \cdot g_s \left[b . \tfrac{1}{2}(X_1 - X_3) + \tfrac{1}{2}(X_1 + X_3) + \| \tfrac{1}{2}(X_1 - X_3) \| u . Y \right] \cdot du \, db \right\}$$

By the continuity almost everywhere of the density g_s and by the fact that X_1 and X_2 are independent and have densities it may be deduced that

$$g_s \left[b . \tfrac{1}{2}(X_1 - X_3) + \tfrac{1}{2}(X_1 + X_3) + \| \tfrac{1}{2}(X_1 - X_3) \| u . Y \right]$$

$$\longrightarrow g_s \left[b . \tfrac{1}{2}(X_1 - X_3) + \tfrac{1}{2}(X_1 + X_3) \right]$$

almost surely as u tends to zero. Hence almost surely

$$\frac{1}{\varepsilon} \int_{-\varepsilon(1-b^2)/2}^{\varepsilon(1-b^2)/2} \cdot g_s \left[b . \tfrac{1}{2}(X_1 - X_3) + \tfrac{1}{2}(X_1 + X_3) + \| \tfrac{1}{2}(X_1 - X_3) \| u . Y \right] du$$

$$\longrightarrow (1 - b^2) \cdot g_s \left[b . \tfrac{1}{2}(X_1 - X_3) + \tfrac{1}{2}(X_1 + X_3) \right] .$$

The boundedness of g_s and the finiteness of the integral mentioned in the theorem's statement permit the Dominated Convergence Theorem to be employed. This means that the existence of the limit $\lambda(g_s)$ can be shown and moreover an explicit integral formula can be given for it. For application

of the Dominated Convergence Theorem shows that

$$\lambda(g_s)/3 = \lim_{\varepsilon \to 0} C_\varepsilon/\varepsilon$$

$$= E\left\{\int_{-1}^{1} .(1-b^2).g_s\left[b.\tfrac{1}{2}(X_1-X_3) + (X_1+X_3)\right] . \|\tfrac{1}{2}(X_1-X_3)\|^2 .db\right\}.$$

The other conclusions of the theorem can be verified easily from this integral formula for $\lambda(g_s)$. It is clear that

$$\lambda(g_s) = 3.\lim_{\varepsilon \to 0} C_\varepsilon/\varepsilon = \tfrac{1}{2}(A/s + Bs)\lambda(g)$$

where A and B are given by somewhat involved integrals;

$$A = (2/\lambda(g)) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-1}^{1} .db\, du_1\, du_2\, dv_1\, dv_2.$$

$$.(1-b^2).g\left[b.\tfrac{1}{2}(u_1-v_1) + \tfrac{1}{2}(u_1+v_1),\ b.\tfrac{1}{2}(u_2-v_2) + \tfrac{1}{2}(u_2+v_2)\right].$$

$$. \tfrac{1}{4}\|(u_1 - v_1)\|^2 .g(u_1,u_2).g(v_1,v_2)$$

and B likewise only replacing the factor $\tfrac{1}{4}\|(u_1-v_1)\|^2$ by $\tfrac{1}{4}\|(u_2-v_2)\|^2$ in the integrand. These integrals for A and B are immediate from the formula $g_s(x_1, x_2) = s.g(sx_1, x_2)$.

It is clear from these formulae that A = B if either $g(x_1, x_2) = g(x_2, x_1)$ or $g(x_1, x_2) = g(-x_2,-x_1)$ or $g(x_1, x_2) = g(x_2,-x_1)$ or $g(x_1, x_2) = g(-x_2, x_1)$.

The symmetry condition SC 2 is equivalent to one of the first two equalities holding. The condition SC 1 is equivalent to one of the second two holding. Thus the sufficiency of either of the two conditions SC 1 or SC 2 has been established.

□

One should note that in the case of uniform densities over convex figures the formula given above for $\lim_{\varepsilon \to 0} C_\varepsilon / \varepsilon$ will simplify to the extent that direct calculations of $\lambda(g)$ and $\lambda(g_s)$ become feasible. As has been mentioned above, such work has been carried out by D.G.Kendall (1979).

The theorem applies to the Gaussian densities on the plane and in the case of circular symmetry the condition SC 1 applies. Thus when $X_1, X_2, \ldots, X_n$ are independent points picked from a Gaussian distribution of elliptic symmetry of density $g_s(x_1, x_2) = s/(2\pi) \cdot \exp\{-\frac{1}{2}(s^2 x_1^2 + x_2^2)\}$ the collinearity constant is

$$\lambda(g_s) = \tfrac{1}{2}(s + 1/s).(3/\pi - 1/\sqrt{3}) .$$

A rather more exotic case is that of the parent distribution being a mixture of uniform distributions over concentric squares. For such a density g the collinearity

constant for the stretched density g_s satisfies

$$\lambda(g_s) = \tfrac{1}{2}(s + 1/s).\lambda(g)$$

and this formula holds whatever the axis of dilation may be.

In the case where the parent density is a stretched Gaussian density it is possible to find out more information about the density of the maximum angle. The remainder of this section is concerned with obtaining an integral for the densities of a given angle and , for part of its range only, the maximum angle. From graphs that are displayed elsewhere it is clear that the approximation

$$P\{ S_s X_1\ S_s X_2\ S_s X_3 \text{ is } \varepsilon\text{-blunt} \} \simeq \lambda(g_s).\varepsilon/\pi$$

is good for reasonable values of ε and reasonable values of s when g is the Gaussian density of circular symmetry.

The key to the calculations is to exploit the circular symmetry of the density g . Suppose that X_1, X_2, X_3 are random points as in section 1.2 with circular Gaussian distributions. Supposing that S_s is a dilation corresponding to a stretch factor s then as noted before

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} P\left(S_s X_1\ S_s X_2\ S_s X_3 \text{ is } \varepsilon\text{-blunt} \right) = \lambda(g_s) .$$

But this relation is true whatever the axis of dilation of S_s

may be. In particular the angle θ of the axis of dilation to the axis $x_2 = 0$ may be taken to be random, uniformly distributed over $[0, 2\pi)$ and independent of the points X_1, X_2, X_3.

The density of the maximum angle shall be obtained only over the subrange $[\frac{1}{2}\pi, \pi]$, the part most relevant for the collinearity investigation. As in 1.2 elementary geometry shows that for $\alpha > \frac{1}{2}\pi$

$$P\{\text{maximum angle of } S_sX_1\ S_sX_2\ S_sX_3 > \alpha\} = 3.P\{\angle S_sX_1\ S_sX_2\ S_sX_3 > \alpha\}$$

and so it suffices to consider the density of a particular angle of the triangle. It is possible to obtain an integral formula for the density of the maximum angle over its whole range, and indeed D.G.Kendall has done this using different techniques (D.G.Kendall (1979)) . However just as in the case of the symmetric Gaussian triangle the formula becomes even more complicated over the range $(\pi/3, \frac{1}{2}\pi)$.

Computation of the formula.

Suppose that θ^* is the angle that X_2X_3 makes to the axis of $x_2 = 0$. Because θ is independent of

the points X_1, X_2, X_3 and is uniform it is the case that $\varphi = \theta^* + \Theta$ is independent of X_1, X_2, X_3 and uniform. Such a conclusion is an interesting property of uniform angular variables and can easily be verified directly.

Because of this it is useful to consider the effect that $S_s = S$ has on the angle between $X_1 X_2$ and the axis of stretch and on the angle between $X_3 X_2$ and the axis of stretch while conditioning on the value of φ. If before stretching the angle made by a line on the axis of stretch is ψ then after stretching the angle will be ψ' where ψ' is in the same quadrant as ψ and satisfies the trigonometrical relations

$$\tan \psi = s \cdot \tan \psi' \quad ;$$

$$\cos \psi = \frac{\cos \psi'}{s \sqrt{[(\cos^2 \psi')/s^2 + \sin^2 \psi']}} \quad ;$$

$$\sin \psi = \frac{\sin \psi'}{\sqrt{[(\cos^2 \psi')/s^2 + \sin^2 \psi']}}$$

If the angle $\angle X_1 X_2 X_3$ equals γ and if the corresponding stretched angle is γ' then the above relations enable us to compute the relationship between γ and γ'. For $\gamma + \phi$ is

the angle between $X_1 X_2$ and the axis of stretch and so the above relations can be applied to it as well as to the angle ϕ . With an obvious notation for the stretched angle corresponding to ϕ it follows that

$$\cos\gamma = \cos(\gamma + \phi - \phi)$$

$$= \cos(\gamma+\phi).\cos\phi + \sin(\gamma+\phi).\sin\phi$$

$$= \frac{s^{-1}.\cos\phi.\cos(\gamma'+\phi') + \sin\phi.\sin(\gamma'+\phi')}{\sqrt{\left[s^{-2}.\cos^2(\gamma'+\phi') + \sin^2(\gamma'+\phi')\right]}}$$

$$= \frac{\left[\begin{array}{l}\cos\gamma'.\cos^2\phi - s^{-1}.\cos\phi.\sin\phi.\sin\gamma' \\ + s.\cos\phi.\sin\phi.\sin\gamma' + \cos\gamma'.\sin^2\phi\end{array}\right]}{\sqrt{\left[\begin{array}{l}\cos^2\gamma' + 2(s - s^{-1}).\cos\phi.\sin\phi.\cos\gamma'.\sin\gamma' \\ + (s^2.\cos^2\phi + s^{-2}.\sin^2\phi).\sin^2\gamma'\end{array}\right]}}$$

and from this

$$\cos\gamma = \eta(\gamma';\phi,s)$$

$$= \frac{\cos\gamma' + (s - s^{-1}).\cos\phi.\sin\phi.\sin\gamma'}{\sqrt{\left[\begin{array}{l}\cos^2\gamma' + 2(s - s^{-1})\cos\phi.\sin\phi.\cos\gamma'.\sin\gamma' \\ + (s^2\cos^2\phi + s^{-2}\sin^2\phi).\sin^2\gamma'\end{array}\right]}}$$

So conditional on the value ϕ of the angle between the angle of stretch and $X_2 X_3$ the angle γ' must have derived from an angle γ with cosine $\zeta(\gamma' : \phi, s)$. Hence ζ defines a map from γ' to $\cos\gamma$. In order to produce a formula for the density of the stretched angle it is necessary to compute the derivative of ζ, firstly to provide a Jacobian and secondly to be able to check that the map is one-one as a map from $(0,\pi)$ to $[-1, 1]$. A rather detailed differentiation reveals that

$$\frac{d}{d\gamma'}\,\zeta(\gamma' : \phi, s)$$

$$= -\frac{(s.\cos^2\phi + s^{-1}.\sin^2\phi)^2 . \sin\gamma'}{\left[\cos^2\gamma' + 2(s - s^{-1}).\cos\phi.\sin\phi.\cos\gamma'.\sin\gamma' + (s^2.\cos^2\phi + s^{-2}.\sin^2\phi).\sin^2\gamma'\right]^{3/2}}$$

which is always negative for γ' in $(0,\pi)$.

Consequently the map is strictly decreasing and thus one-one.

From the work of section 1.2 the density of the angle $\angle X_1 X_2 X_3$ is given by

$$h(\gamma) = \frac{3}{\pi}.\left\{\frac{1}{(3+\sin^2\gamma)} . \left[1 + \frac{2.\cos\gamma}{\sqrt{(3+\sin^2\gamma)}}.\tan^{-1}\sqrt{\frac{2+\cos\gamma}{2-\cos\gamma}}\right]\right\}$$

when γ belongs to $(0,\pi)$.

Consequently the density of the cosine of $\angle X_1\, X_2\, X_3$ is $h(\cos^{-1} x) \cdot (1 - x^2)^{-\frac{1}{2}}$ and thus the formula for the density of $\angle SX_1\, SX_2\, SX_3$ conditional on ϕ is

$$\frac{h(\cos^{-1}[\flat(\gamma' : \phi, s)]) \cdot [-d/d\gamma' \cdot \flat(\gamma' : \phi, s)]}{\sqrt{[1 - \flat^2(\gamma' : \phi, s)]}} .$$

When the appropriate substitutions are made and the ϕ variable is integrated out (thus removing the conditioning) the formula for the (unconditional) density for the stretched angle stands as

$$\frac{1}{2\pi} \cdot \int_0^{2\pi} d\phi \cdot \frac{h(\cos^{-1}[\flat(\gamma' : \phi, s)]) \cdot (s\cos^2\phi + s^{-1}\sin^2\phi)}{\left[\begin{array}{c}\cos^2\gamma' + 2(s - s^{-1})\sin\phi\cos\phi\sin\gamma'\cos\gamma' \\ + (s^2\cos^2\phi + s^{-2}\sin^2\phi)\sin^2\gamma'\end{array}\right]}$$

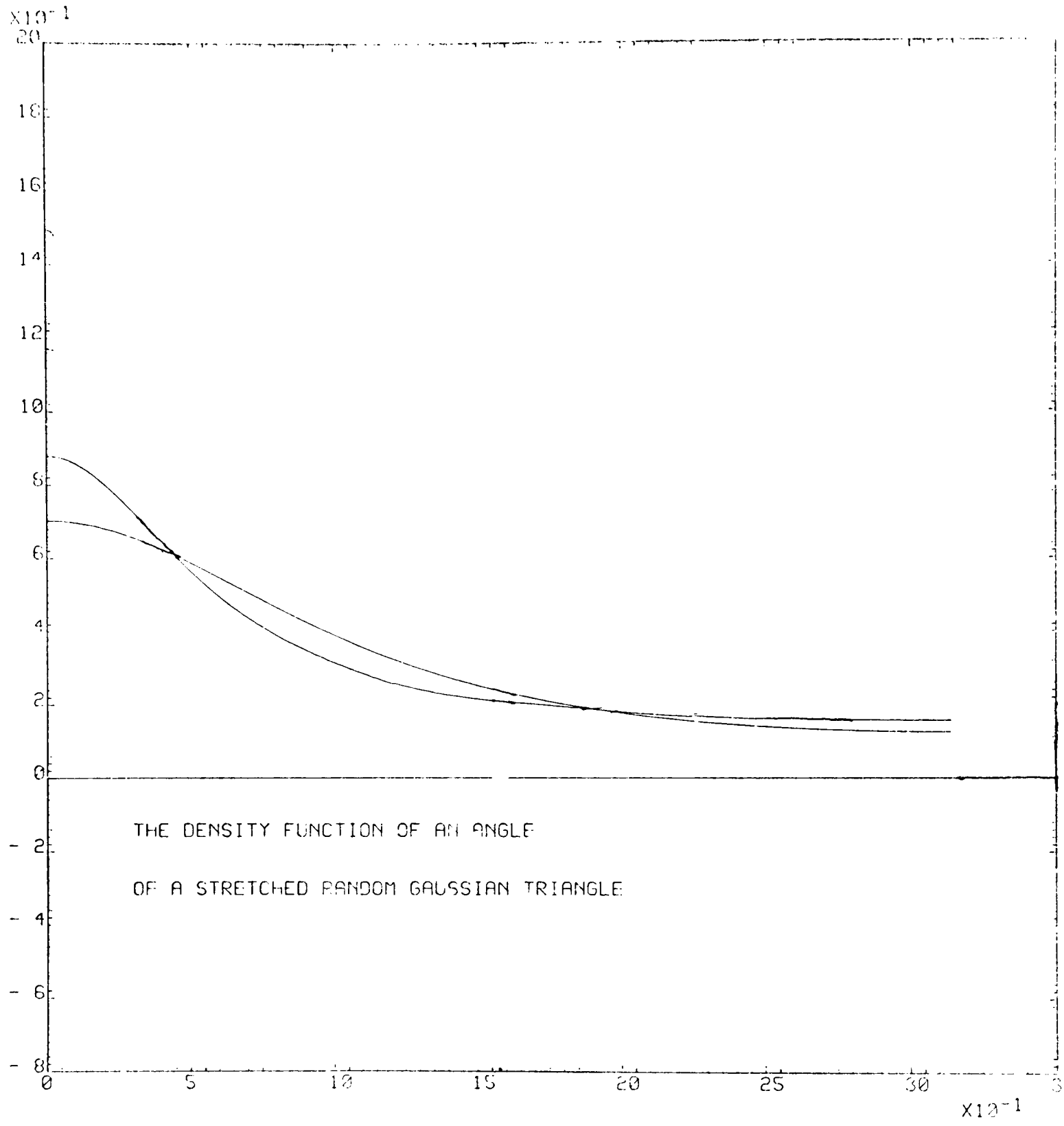
Because this formula involves an integral the computational effort involved in obtaining an accurate graph is immense. Nevertheless such a graph has been obtained by D.G.Kendall for various values of s, both using this formula and also using a geometric method still involving one integration for each data point. (Both methods agree !)

Indeed D.G.Kendall's method has the advantage over the formula given here in that it may be used to derive the density of the maximum angle of a stretched random Gaussian triangle over the whole of its range, from $\pi/3$ to π. The density of the maximum angle over the range $\pi/2$ to π may be obtained from the formula given above merely by multiplying by 3, as straightforward geometric reasoning shows. However obtaining the formula over the whole range by the method of this section, while a possibility, would

require an analysis involving the "wedges" of section 1.2 and would be yet more intricate.

A graph obtained by D.G.Kendall of the density of the density of the maximum angle for values of s ranging between 1 and 5 in steps of 0.5 may be found in D.G.Kendall and W.S.Kendall (1979) . As remarked above, this provides a further illustration of the validity of the approximation given when the collinearity constant is employed.

In fact for quite large values of s the density function of a particular angle is close to the density in the unstretched case. This point is illustrated by the following graph, which displays the densities for the cases of $s = 1$ and $s = 2$. The case $s = 2$ corresponds to the curve reaching a higher value at the right-hand side of the graph.



Graphed for stretch factor s = 1 , 2 . The graph of stretch factor 2 is higher at the right hand side. To graph the density for higher values of s requires much time and effort on the computer.

1.4 Bounds on the collinearity constant.

The results of section 1.3 go some way to describing how the collinearity constant changes as the parent distribution alters. This section produces bounds on the values that the collinearity constant may take given some rather general restrictions on the parent distribution.

For the parent distribution uniform over a compact convex region D.G.Kendall has shown that the collinearity constant must be no less than $1/\pi$ (D.G.Kendall (1979)). Moreover this bound is attained when the parent distribution is uniform on a disc. If the restriction on the parent distribution is relaxed then the collinearity constant may be zero (consider the case of a distribution uniform over a circle !). The following establishes the same lower bound of $1/\pi$ for the case when the parent distribution is restricted to be of circular symmetry (centred at the origin, say) and such that the density $f(x_1, x_2)$ is a decreasing function of the radius of symmetry $\sqrt{(x_1^2 + x_2^2)}$. Such a distribution is a mixture of uniform distributions on concentric discs and it will be convenient to call it "a mixture of discs".

It is necessary to establish that the collinearity constant of mixtures of discs is finite. This has already been established in the case when the mixture has bounded second moments, by the work of the theorem of section 1.3 . In fact it is also the case for general mixtures of discs but finiteness need not hold in general as the following example shows.

Example: a distribution of bounded density continuous almost everywhere but with collinearity constant infinite.

Consider a sequence of rectangles with sides parallel to the x_1- and x_2-axes and such that

the centroids of the rectangles coincide;

the sides parallel to the x_1-axis all have length one;

the sides of the n-th rectangle parallel to the x_2-axis have length 2^{3n} .

Then the distribution which is a mixture of uniform distributions on these rectangles giving weight 2^{-n} to the n-th rectangle is the required example.

Proof

The collinearity constant of the distribution uniform

on the zero-th rectangle is finite, by the work of the theorem of section 1.3 . This work also shows that the collinearity constant of the n-th restangle is

$$\tfrac{1}{2}.(2^{3n} + 2^{-3n}). C$$

where C is the collinearity constant of the zero-th rectangle. (The actual value of C is given in D.G.Kendall (1979))

Because of this if X_1 , X_2 , X_3 are independent points drawn from the mixture of the uniform distributions then for any N for sufficiently small ε .

$$\frac{1}{\varepsilon} \cdot P \{ X_1 X_2 X_3 \text{ is } \varepsilon\text{-blunt} \}$$

$$\geqslant \sum_{n=0}^{N} (2^{-n})^3 \cdot \tfrac{1}{2}.(2^{3n} + 2^{-3n}).C$$

$$\geqslant \tfrac{1}{2} N.C .$$

So the collinearity constant of the mixed distribution is infinite.

That the mixed distribution has a bounded density that is almost everywhere continuous is easy to verify. Clearly a straightforward modification produces an example with a continuous bounded density.

□

The following lemma substantiates the assertion made above that mixtures of discs have finite collinearity constants.

Lemma

If a distribution is a mixture of discs then whether or not its second moments exist it has a finite collinearity constant.

Proof

Let R_1, R_2, R_3 be independent positive random variables with the same distribution. Let X_1, X_2, X_3 be independent random points with distributions uniform on the discs centred at the origin of radii R_1, R_2, R_3 when conditioned on R_1, R_2, R_3. Clearly it suffices to prove that

$$\frac{1}{\varepsilon} \cdot P\{X_1\ X_2\ X_3 \text{ is } \varepsilon\text{-blunt}\}$$

tends to a finite limit.

That the limit

$$C(R_1, R_2, R_3) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \cdot P\{X_1\ X_2\ X_3 \text{ is } \varepsilon\text{-blunt} \mid R_1, R_2, R_3\}$$

exists and is finite is a straightforward modification of an application of the theorem of 1.3 to the case of a

mixture of three discs. If $C(R_1, R_2, R_3)$ can be bounded then the Dominated Convergence theorem can be applied to show that

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \cdot P\{X_1 X_2 X_3 \text{ is } \varepsilon\text{-blunt}\} = E[C(R_1, R_2, R_3)]$$

exists and is finite. But such a bound is a consequence of the inequality

$$P\{X_1 X_2 X_3 \text{ is } \varepsilon\text{-blunt} \mid X_1, X_2, (R_1 < R_2, R_3)\} \leq 2.4\varepsilon/\pi$$

which can be established by simple geometrical arguments. From this inequality

$$C(R_1, R_2, R_3) \leq 8/\pi .$$

□

This argument not only proves the existence of the collinearity constant for mixtures of discs but yields an upper bound of $8/\pi$. The following theorem will not only provide a lower bound but also as a corollary a best possible upper bound.

Theorem

Suppose X_1, X_2, X_3 are independent points on a plane with the same distribution of density $f(x_1, x_2)$, and such that the distribution is a mixture of discs as described above. Then $\lambda(f) =$

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \cdot P \{ X_1 X_2 X_3 \text{ is } \varepsilon\text{-blunt} \} \geq 1/\pi .$$

Proof

Because the distribution is a mixture of discs centred at the origin it suffices to establish that the limit

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \cdot P \{ Y_1 Y_2 Y_3 \text{ is } \varepsilon\text{-blunt} \}$$

is bounded below by $1/\pi$ when Y_1, Y_2, Y_3 are independent and uniformly distributed over concentric discs of radii r, s, 1 respectively, where $r \leq s \leq 1$

For this purpose three new limits are defined;

$$K_1(r, s) = \lim (1/\varepsilon) P\{ \angle Y_1 Y_2 Y_3 > \pi - \varepsilon \}$$

$$K_2(r, s) = \lim (1/\varepsilon) P\{ \angle Y_2 Y_3 Y_1 > \pi - \varepsilon \}$$

$$K_3(r, s) = \lim (1/\varepsilon) P\{ \angle Y_3 Y_1 Y_2 > \pi - \varepsilon \}$$

where the limits are taken as ε tends to zero. Clearly

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \cdot P \{ Y_1 Y_2 Y_3 \text{ is } \varepsilon\text{-blunt} \} =$$

$$K_1(r, s) + K_2(r, s) + K_3(r, s) .$$

Since $\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \cdot P\{ \angle Y_2\, Y_3\, Y_1 > \pi - \varepsilon \mid \| Y_3 \| > s \} = 0$

it follows that

$$K_2(r, s) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} P\{ \angle Y_2\, Y_3\, Y_1 > \pi - \varepsilon \text{ and } \| Y_3 \| \leq s \} .$$

Moreover the angle $\angle Y_2\, Y_3\, Y_1$ is greater than $\pi - \varepsilon$ exactly when Y_3 belongs to the intersection of the two discs of radius $\frac{1}{2} . \| Y_2 - Y_1 \| . \operatorname{cosec} \varepsilon$ with secant the line segment $Y_2\, Y_1$. Given Y_2, Y_1 for sufficiently small ε the intersection will always be contained in the disc of radius s centre the origin. So by the Dominated Convergence theorem

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \cdot P\{ \angle Y_2\, Y_3\, Y_1 > \pi - \varepsilon \text{ and } \| Y_3 \| \leq s \}$$

$$= E\left[\| Y_1 - Y_2 \|^2 \cdot \lim_{\varepsilon \to 0} \frac{1}{2\pi} \left(\operatorname{cosec}^2 \varepsilon - \frac{\cot \varepsilon}{\varepsilon} \right) \right]$$

$$= E \| Y_1 - Y_2 \|^2 \cdot 1/(3\pi)$$

$$= \tfrac{1}{2} . (r^2 + s^2) . 1/(3\pi) .$$

This completes the evaluation of $K_2(r, s)$. The evaluations of the other two limits will be more complicated.

Consider $K_3(r, s)$. This can be approached by evaluating $P\{ \angle Y_3\, Y_1\, Y_2 > \pi - \varepsilon \mid Y_1, Y_3 \}$ which is

equal to

$$\frac{\varepsilon}{\pi s^2} \cdot \left[\sqrt{(s^2 - \|Y_1\|^2 . \sin^2 \theta)} - \|Y_1\| . \cos \theta \right]^2 + o(\varepsilon)$$

where θ is the angle subtended on Y_1 by Y_3 and the origin (see figure). It is important to note that $o(\varepsilon)/\varepsilon$ is uniformly bounded and thus behaves when Y_1 , Y_3 are integrated out.

Integrating out Y_3 yields

$$P \left\{ \angle Y_3 Y_1 Y_2 > \pi - \varepsilon \;\middle|\; \|Y_1\| = t \right\} =$$

$$\frac{\varepsilon}{\pi s^2} \cdot \int_0^{2\pi} \left[\sqrt{(s^2 - t^2 \sin^2 \theta)} - t \cos \theta \right]^2 . \left[\sqrt{(1 - t^2 \sin^2 \theta)} + t \cos \theta \right]^2 . \frac{d\theta}{2\pi} + o(\varepsilon)$$

$$= \frac{\varepsilon}{2\pi^2 s^2} \cdot \left\{ \int_0^{2\pi} \cdot \left[s^2 - t^2 \sin^2 \theta + t^2 \cos^2 \theta \right] \left[1 - t^2 \sin^2 \theta + t^2 \cos^2 \theta \right] . d\theta \right.$$

$$\left. - 4 \int_0^{2\pi} \sqrt{(s^2 - t^2 \sin^2 \theta)} \sqrt{(1 - t^2 \sin^2 \theta)} . t^2 \cos^2 \theta \; d\theta \right\}$$

$$+ o(\Sigma)$$

(exploiting the oddness of $\cos \theta$ about $\frac{1}{2}\pi$ to eliminate crossterms.)

$$= \frac{\Sigma}{\pi s^2} \cdot \left\{ s^2 + \tfrac{1}{2} t^4 - \frac{8}{\pi} \cdot \int_0^{\frac{1}{2}\pi} \sqrt{(s^2 - t^2 \sin^2 \theta)} . \sqrt{(1 - t^2 \sin^2 \theta)} . t^2 \cos^2 \theta \; d\theta \right\}$$

$$+ o(\Sigma) \; .$$

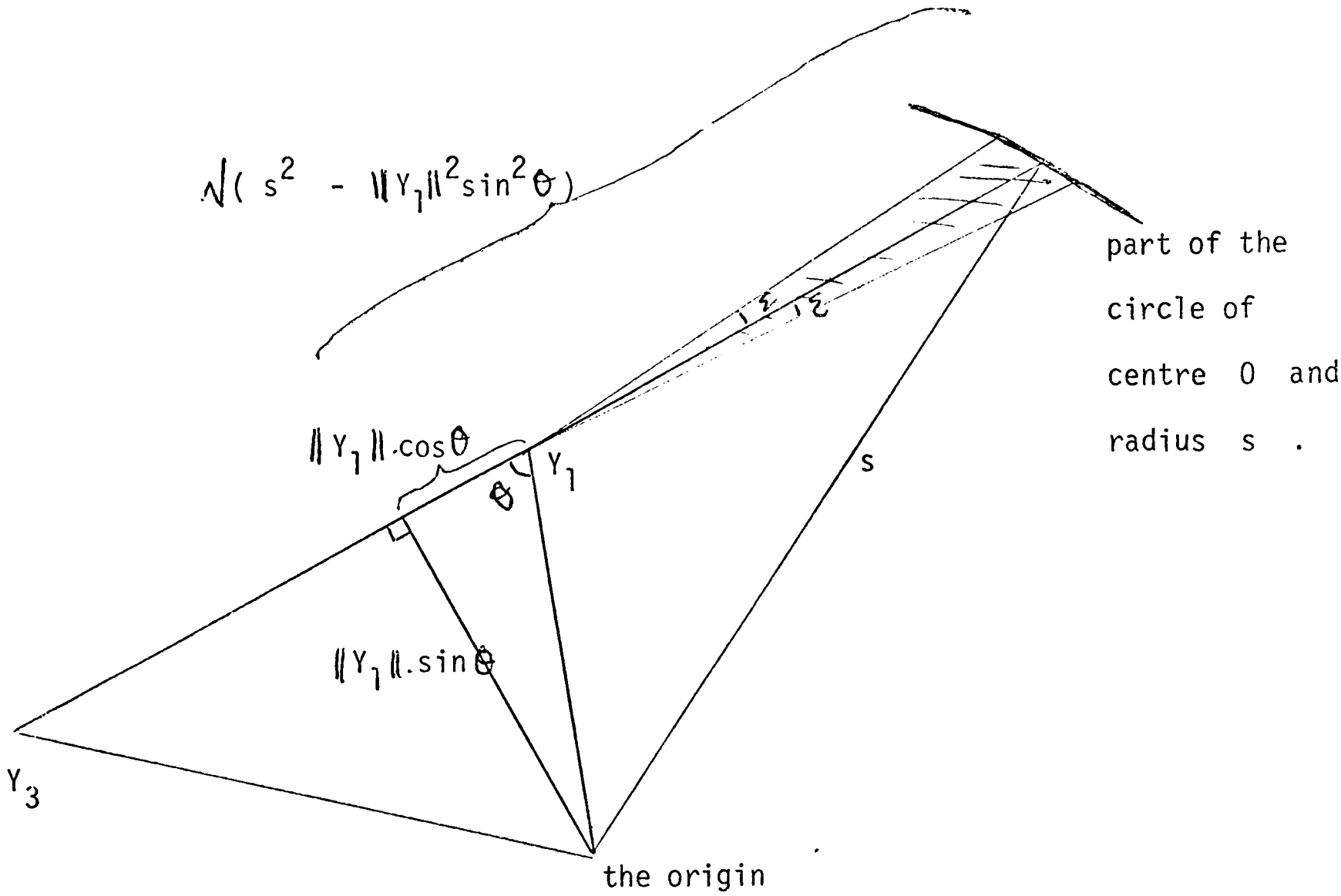


Figure to be attached after page 50 , section 1.4 .

The hatched region has area

$$\varepsilon \cdot \left[\sqrt{(s^2 - \|Y_1\|^2 \sin^2\theta)} - \|Y_1\| \cdot \cos\theta\right]^2 + o(\varepsilon)$$

where

$$o(\varepsilon) < \varepsilon(2s)^2$$

and

$$o(\varepsilon)/\varepsilon \longrightarrow 0 \quad \text{as} \quad \varepsilon \to 0 \ .$$

Making the substitution $t.\sin\theta = w$ and integrating out the dependence on $\| Y_1 \| = t$ leaves

$$P\{ \angle Y_3\, Y_2\, Y_1 > \pi - \varepsilon \}$$

$$= \frac{\varepsilon}{\pi s^2}\left\{ s^2 + r^4/6 - \frac{8}{\pi}\cdot\int_0^r \frac{2t\,dt}{r^2}\cdot \right.$$

$$\left. \cdot\int_0^t \sqrt{(s^2-w^2)}\sqrt{(1-w^2)}\sqrt{(t^2-w^2)}\;dw \right\} + o(\varepsilon)$$

$$= \frac{\varepsilon}{\pi s^2}\left\{ s^2 + r^4/6 - \frac{16}{3\pi r^2}\int_0^r \sqrt{(s^2-w^2)}\sqrt{(1-w^2)}\cdot(r^2-w^2)^{3/2}\,dw \right\}$$

$$+ o(\varepsilon)\ .$$

Putting $w = r.v$ leads to the deduction

$$K_3(r, s) =$$

$$\frac{1}{\pi}\cdot\left[1 + \frac{r^4}{6s^2} - \frac{16r^2}{3\pi s^2}\int_0^1 \sqrt{(s^2-r^2v^2)}\sqrt{(1-r^2v^2)}\cdot(1-v^2)^{3/2}dv \right]$$

The evaluation of $K_1(r, s)$ follows a similar line of attack. In fact by obvious geometric reasoning

$$\frac{1}{\varepsilon}.P\{ \angle Y_1\, Y_2\, Y_3 > \pi - \varepsilon \text{ and } \| Y_2 \| \le r \}$$

$$\rightarrow\ r^2/s^2\ .\ K_3(r, r)$$

$$= \frac{r^2}{\pi s^2}\left[1 + \frac{r^2}{6} - \frac{16}{3\pi r^4}\cdot\int_0^r \sqrt{(1-w^2)}\cdot(r^2-w^2)^2\,dw \right]\ .$$

Although the integral can be completely evaluated it transpires that some cancellation occurs between this expression and the expression for the limit of

$$\frac{1}{\varepsilon} \cdot P\{ \angle Y_1 Y_2 Y_3 > \pi - \varepsilon \text{ and } \|Y_2\| > r \}.$$

For conditioning on $\|Y_2\| = t > r$ and Y_3, with φ equal to the angle subtended on Y_2 by Y_3 and the origin, the same geometric arguments as were employed for $K_3(r, s)$ lead to

$$P\{ \angle Y_1 Y_2 Y_3 > \pi - \varepsilon \mid Y_2, Y_3 \}$$

$$= \frac{\varepsilon}{\pi r^2} \cdot \{ [\sqrt{(r^2 - \|Y_2\|^2 \sin^2\varphi)} + \|Y_2\| \cos\varphi]^2 - [\sqrt{(r^2 - \|Y_2\|^2 \sin^2\varphi)} - \|Y_2\| \cos\varphi]^2 \} + o(\varepsilon)$$

when $|\sin\varphi| < r/\|Y_2\|$ and $\cos\varphi$ is positive.

Integrating out Y_3 yields

$$P\{ \angle Y_1 Y_2 Y_3 > \pi - \varepsilon \mid \|Y_2\| = t \}$$

$$= \frac{\varepsilon}{\pi r^2} \cdot \int_{-\sin^{-1} r/t}^{\sin^{-1} r/t} 4 \cdot \sqrt{(r^2 - t^2 \sin^2\varphi)} \cdot t \cos\varphi \cdot [\sqrt{(1 - t^2 \sin^2\varphi)} - t \cos\varphi]^2 \frac{d\varphi}{2\pi} + o(\varepsilon)$$

$$= \frac{2\varepsilon}{\pi^2} \cdot \int_{-1}^{1} \sqrt{(1-w^2)} \cdot \left[\sqrt{(1-r^2w^2)} - \sqrt{(t^2-r^2w^2)} \right]^2 dw$$

$$+ \; o(\varepsilon)$$

(on putting $w = \frac{t}{r} \sin\varphi$)

$$= \frac{4\varepsilon}{\pi^2} \int_{0}^{1} \sqrt{(1-w^2)} \cdot \left[1 - r^2w^2 + t^2 - r^2w^2 - 2\sqrt{(1-r^2w^2)}\sqrt{(t^2-r^2w^2)} \right] dw$$

$$+ \; o(\varepsilon)$$

$$= \frac{4\varepsilon}{\pi^2} \cdot \tfrac{1}{4}\pi \cdot (1 + t^2 - \tfrac{1}{2}r^2) - \frac{8\varepsilon}{\pi^2} \cdot \int_{0}^{1} \sqrt{(1-w^2)}\sqrt{(1-r^2w^2)}\sqrt{(t^2-r^2w^2)} \; dw$$

$$+ \; o(\varepsilon) \; .$$

Finally integrating out $\| Y_2 \| = t$ yields

$$P \{ \angle Y_1 Y_2 Y_3 > \pi - \varepsilon \text{ and } \| Y_2 \| > r \}$$

$$= \frac{\varepsilon}{\pi s^2} \cdot \left[(1-\tfrac{1}{2}r^2) \cdot (s^2-r^2) + \tfrac{1}{2}(s^4-r^4) \right] - \frac{16\varepsilon}{\pi^2} \cdot \int_{0}^{1} \sqrt{(1-w^2)}\sqrt{(1-r^2w^2)} \cdot \left[(s^2-r^2w^2)^{3/2} - r^3(1-w^2)^{3/2} \right] dw$$

$$+ \; o(\varepsilon) \; .$$

From the above it follows that

$$K_1(r, s) =$$

$$\frac{1}{\pi} \cdot \left\{ 1 + \tfrac{1}{2}(s^2-r^2) + \frac{r^4}{6s^2} - \frac{16}{3\pi r^2 s^2} \int_0^r \sqrt{(1-u^2)\cdot(r^2-u^2)^2}\ du \right.$$

$$- \frac{16}{3\pi s^2} \cdot \int_0^1 \sqrt{(1-w^2)}\sqrt{(1-r^2w^2)}\ \cdot$$

$$\left. \cdot \left[(s^2-r^2w^2)^{3/2} - r^3(1-w^2)^{3/2}\right]\ dw \right\}$$

Putting $u = rw$ it can be seen that the first integral cancels with part of the second integral. Thus $K_1(r, s) =$

$$\frac{1}{\pi} \cdot \left\{ 1 + \tfrac{1}{2}(s^2-r^2) + \frac{r^4}{6s^2} \right.$$

$$\left. - \frac{16}{3\pi s^2} \cdot \int_0^1 \sqrt{(1-w^2)}\sqrt{(1-r^2w^2)}\cdot(s^2-r^2w^2)^{3/2}\ dw \right\}.$$

Consequently

$$K_1(r, s) + K_2(r, s) + K_3(r, s) =$$

$$\frac{1}{\pi} \cdot \left\{ 2 + (2s^2 - r^2)/3 + \frac{r^4}{3s^2} \right.$$

$$\left. - \frac{16}{3\pi s^2} \int_0^1 \sqrt{(1-w^2)}\sqrt{(1-r^2w^2)}\sqrt{(s^2-r^2w^2)}\cdot (s^2 + r^2 - 2r^2w^2)\ dw \right\}$$

---

and it is this expression that must be bounded from below.

The strategy to be used is to show that the expression is bounded below by $1/\pi$ when $s = 1$ and then to show that the expression decreases in value when s is increased with the ratio of r to s held constant. It is straightforward to show that

$$K_1(r, 1) + K_2(r, 1) + K_3(r, 1)$$
$$= \frac{1}{\pi} \cdot \left[1 + (r^2 - 1)^2/3\right] \geq 1/\pi .$$

So it suffices to show that

$$K_1(\lambda s, s) + K_2(\lambda s, s) + K_3(\lambda s, s)$$
$$= L(\lambda, s)$$

is a decreasing function of s when λ is held fixed. Clearly it may be useful to compute derivatives of L.

$$\pi \frac{\partial}{\partial s} L(\lambda, s) = \frac{2}{3}(2 - \lambda^2 + \lambda^4)\, s$$
$$- \frac{16}{3\pi} \int_0^1 (1 + \lambda^2(1-2w^2)) \sqrt{(1-w^2)} \cdot \sqrt{(1-\lambda^2 s^2 w^2)}\, \sqrt{(1-\lambda^2 w^2)} \cdot \frac{(1 - 2\lambda^2 s^2 w^2)}{(1 - \lambda^2 s^2 w^2)}\, dw$$

and

$$\pi \frac{\partial^2}{\partial s^2} L(\lambda, s) = \frac{2}{3}(2 - \lambda^2 + \lambda^4)$$
$$+ \frac{16}{3\pi} \int_0^1 (1 + \lambda^2(1-2w^2)) \sqrt{(1-w^2)} \cdot \sqrt{(1-\lambda^2 s^2 w^2)}\, \sqrt{(1-\lambda^2 w^2)} \cdot \frac{(3\lambda^2 s^2 w^2 - 2\lambda^4 s^4 w^4)}{(1 - \lambda^2 s^2 w^2)^2}\, dw .$$

Since $r < s$ the constant λ is less than 1. Thus the product $\lambda^2 s^2 w^2$ is less than 1 in the integral for $\frac{\partial^2}{\partial s^2} L(\lambda, s)$ and so the factor

$$\frac{(3\lambda^2 s^2 w^2 - 2\lambda^4 s^4 w^4)}{(1 - \lambda^2 s^2 w^2)}$$

is positive always. Consequently the second partial derivative is always positive, being a sum of positive terms. As a result the largest value of the first partial derivative, for a fixed ratio λ, must be attained when $s = 1$. But

$$\pi \frac{\partial}{\partial s} L(\lambda, s) \Big|_{s=1} = (\lambda^4 - \lambda^2)/3$$

which is non-positive for all λ in $[0, 1]$.

Thus $L(\lambda, s)$ is a decreasing function of s when λ is held fixed. So the lower bound on

$$K_1(r, 1) + K_2(r, 1) + K_3(r, 1)$$

is valid for general $L(r/s, s)$ and the theorem is proved.

□

In conclusion a corollary will be established that gives strict upper bounds on the collinearity constant of a mixture of discs.

Corollary

Suppose that X_1, X_2, X_3 are independent points on a plane with the same distribution. Suppose further that the distribution is a mixture of discs. Then

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \cdot P\{ X_1\ X_2\ X_3 \text{ is } \varepsilon\text{-blunt} \} \leq 2/\pi$$

and this upper bound is the least possible.

Proof

Using the methods of the proof of the theorem let Y_1, Y_2, Y_3 be uniformly distributed over concentric discs of radii $\lambda_1\lambda_2\lambda_3$, $\lambda_2\lambda_3$, λ_3 respectively. Then using the notation of the theorem proof

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \cdot P\{ X_1\ X_2\ X_3 \text{ is } \varepsilon\text{-blunt} \} = L(\lambda_1, \lambda_2)$$

$$\longrightarrow 2/\pi \text{ as } \lambda_2 \longrightarrow 0 \text{ and } \lambda_1 \text{ stays fixed}$$

as can be easily verified using the formulae given in the proof of the theorem. But it has been established that for fixed λ_1 as λ_2 decreases so $L(\lambda_1, \lambda_2)$ increases. Consequently the limit $2/\pi$ must also be the lower bound.

The above argument establishes the upper bound for the collinearity constant of a mixture of discs. To show that this upper bound of $2/\pi$ is also strict consider the parent distribution that is the mixture of n distributions uniform on concentric discs of radii

$$r_1, \ldots, r_n$$

with the ratio $r_m/r_{m+1} = \lambda$, taking the mixture to give equal weights to each uniform disc distribution.

By considering the various possible mixtures it can be seen that

$$\lim_{\lambda \to 0} \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \cdot P\{ X_1 X_2 X_3 \text{ is } \varepsilon\text{-blunt} \}$$

must be larger than

$$\frac{n(n-1)(n-2)}{n^3} \cdot \frac{2}{\pi} .$$

So collinearity constants may be found as close to $2/\pi$ as desired by chosing λ small enough and n large enough. □

The results of this section lend a certain perspective to the results of section 1.2 . For a Gaussian distribution of circular symmetry is clearly a mixture of discs and so its collinearity constant

may be compared with the upper and lower bounds attained on the collinearity constant in the above work.

The Gaussian constant is $3/\pi - 1/\sqrt{3} = 0.3776$ while the upper and lower bounds for the class of mixture of discs are 0·63662 and 0·31831 respectively. It is interesting that the Gaussian collinearity constant should be close to the smallest possible constant for the mixture of discs case. Indeed this suggests the possibility that a tighter upper bound may be obtained by further restricting the class of distributions while still retaining the Gaussian distribution among those considered. While this would be of relevance to practical problems, as indicated in section 1.5 , it is hard to think of a further restriction that would lead to an analytically tractable problem.

1.5 Conclusion.

The paper of D.G.Kendall and W.S.Kendall (1979) reports on the results obtained both in this chapter and also elsewhere concerning the statistical analysis of the archaeological problem described in the introduction. The results obtained elsewhere include in particular those concerning the "pontogram" developed by D.G.Kendall which provides the best and most flexible approach to the problem and the reader is referred to the above paper for a description of the use of this device. However the results obtained in this chapter enable a simple preliminary analysis to be carried out on collinearity problems.

For example suppose that an analysis is being carried out on a dataset composed of 48 archaeological sites. (Such a dataset, the Cumbria dataset, has been obtained by a map-survey of Cumbria carried out by Stewart and others (1978)) Assuming that these sites can be taken as being distributed independently with common Gaussian distribution of circular symmetry, then the expected number of ε-blunt triples formed is

28.495 (using the formula displayed at the start of 1.3)

when $\varepsilon = (0.25).\pi / 180$. (So that the blunt triples

are those triples with largest interior angle greater than 179.75^{o}). This expected number can be compared with the actual number of blunt triples observed, thus obtaining some feel for whether the actual number is at all unusual. In the case of the Cumbria dataset the actual number of triples with this level of bluntness is 32 which does not appear to be very different from the expectation. Of course to make this approach complete some knowledge is required of the variance of the random variable the number of blunt triples. An asymptotic expression can be found for the variance but it is in the form of a very complicated integral and is not given here.

The results of section 1.3 show the extent to which the analysis is affected if the underlying distribution is "stretched". Such considerations are relevant if the dataset is drawn from a region which appears to be more 'rectangular'in shape than'square' . The Land's End dataset dealt with in Broadbent (1979) appears to have this feature.

Bounds on the collinearity constant are of use in so far as they indicate the sensitivity of the statistical analysis to the assumptions concerning the underlying distribution. The results of section 1.4 show that if

the underlying distribution is a mixture of discs then the collinearity constant is at least $1/\pi$. In the case of the Cumbria dataset this would mean that the expected number of Σ-blunt triples would be at least 24.022 , a figure within 16 % of the corresponding figure for the Gaussian case. As mentioned at the end of section 1.4 , it seems possible that a further restriction on the class of distributions would yield an upper bound rather tighter than that corresponding to the case of mixtures of discs.

To finish this chapter reference will be made to similar results obtained by other workers. D.G.Kendall has found a formula for the collinearity constant for distributions uniform over compact convex regions, and has also evaluated the asymptotic variance corresponding to the number of Σ-blunt triples observed when points are drawn at random and uniformly from a square or rectangle. These results are announced in D.G.Kendall and W.S.Kendall (1979) . Silverman and Brown (1978) consider the random process N_u = "the number of u-blunt triples" and produce convergence theorems for this process suitably scaled. R.Edwards (1978) has obtained the joint density of the angles of the triangle formed by three points each of Gaussian

distribution but centred at the three points of an equilateral triangle. Other workers (Gečiauskas (1971) and(1972) and Sukiasyai (1978)) have considered the rather different random triangles defined using ideas of integral geometry.

Chapter 2 : The knotting of Brownian motion.

2.1 Introduction.

The trajectory of Brownian motion in 3-space has been used to model the spatial configuration of a long polymer molecule. The question of whether or not the path knots has some physical relevance. This is because the existence or otherwise of a knot in a polymer molecule places a topological constraint on the phase space of the molecule since (if the finite length of the molecule is disregarded) it is not possible for the polymer to unknot without passing through itself. Edwards and Deam, in Edwards (1968) and Deam and Edwards(1976), use integral invariants of knots in computation of bulk properties of conglomerates of polymer molecules.

In Edwards (1968) the conjecture is made that Brownian motion in 3-space knots "infinitely often" in each period of time, however small. Of course such behaviour does not occur for a polymer molecule; the approximation by the trajectory of a Brownian motion

only has validity as a large-scale approximation. The following work is an investigation of Edwards' conjecture, which, as it arose in a physical context, has some indirect physical interest.

By the work of Dvoretsky, Erdös and Kakutani (1950) Brownian motion in 3-space is known to have infinitely many double points, these double points forming a set dense along the Brownian path. Indeed, according to Fristedt (1967), the Hausdorff dimension of the set of double points equals one. So Brownian motion in 3-space is by no means a simple curve; it intersects itself very frequently. Since the topological definition of a knot refers only to simple closed curves in 3-space, a modification of the concept of knotting is needed.

An appropiate modification is suggested by the following result, Paul Lévy's so-called "Forgery Theorem" for Brownian motion;

Theorem

Suppose B to be a Brownian motion in n-space begun at the origin, and let f be any continuous map from $[0, 1]$ to $\mathbb{R}^n$ with $f(0) = 0$. Then for every positive ε the event

$$\{ \|B(t) - f(t)\| < \varepsilon \quad \text{for all} \quad t \quad \text{in} \quad [0, 1] \}$$

has positive probability.

For a proof of this result the book by Freedman (1971) may be consulted. Its name comes from the corollary attributed to Whittle;

"Almost surely every arc of the Brownian path contains the complete works of Shakespeare in the handwriting of Bacon.".

This corollary is a consequence of the result, scaling properties and the Strong Markov property.

Section 2.2 presents the modification of the concept of knotting to be used and discusses properties of curves knotting in this modified sense. Section 2.3 applies this concept to the Brownian path in 3-space. Finally section 2.4 discusses the topological character of the Brownian path in higher dimensional space; in 5-space and above the answer is straightforward and is given in Milnor (1964) while in 4-space it is only possible to give a partial answer at present.

2.2 Knots and implication.

The modified notion of knotting to be presented will be called <u>implication in a knot-tube.</u>

Suppose that d is the usual Euclidean metric for $\mathbb{R}^3$. Let C be the unit cube for $\mathbb{R}^3$, that is the set $\{(x_1,x_2,x_3) \in \mathbb{R}^3 : |x_i| \leq 1 \text{ for } i = 1,2,3\}$. Let ∂C be the boundary of C. Then a <u>knot-pair</u> (k, ε) is a pair consisting of a piecewise-linear map

$$k : [0, 1] \longrightarrow C$$

and a positive number ε with (k, ε) satisfying the conditions (i) - (iv) below.

The idea of the conditions is to ensure that if all points that are closer than ε to the image of k are removed from C then the surface of the remaining set is topologically a knotted torus. Clearer it would be possible to construct a more sophisticated definition of greater generality. However the definition contained in the conditions below is descriptive and adequate for the purposes of this work.

Condition (i) is that k must not intersect itself and that $|\varepsilon| < \frac{1}{2}$.

Condition (ii) requires that the endpoints of k are fixed at $k(0) = -(1,0,0)$ and $k(1) = (1,0,0)$, moreover that $k(t) \notin \partial C$ for $t \in (0, 1)$.

Condition (iii) is the following; if

$$0 = r_0 < r_1 < \ldots < r_s = 1$$

is the coarsest partition of $[0, 1]$ such that the restriction k_p of k to $[r_p, r_{p+1}]$ is linear then the images of k_p and k_q must be more than 2ε apart whenever $|p - q| \geq 2$. Moreover the image of k_p must be more than ε away from ∂C whenever $p \neq 0$ and $p \neq s-1$.

Condition (iv) is that if $N = \{x \in C : d(x, \text{Im. } k) < \varepsilon\}$ then $N \cap \partial C$ has precisely two connected components, the <u>entrance window</u> containing $-(1,0,0)$ and the <u>exit window</u> containing $(1,0,0)$.

For future reference define $K = \partial C \cup \partial N$.

A <u>knot-tube</u> based on the knot-pair (k, ε) is a dilation of a translation of K. The <u>entrance</u> and <u>exit windows</u> of the knot-tube are the dilated translates of the corresponding subsets of K. The dilated translate of N is the <u>core</u> of K.

A diagram of a simple knot-tube can be found on page 70, with entrance and exit windows and core labelled.

Let K be a knot-tube and let $h : \mathbb{R} \to \mathbb{R}^3$ be a

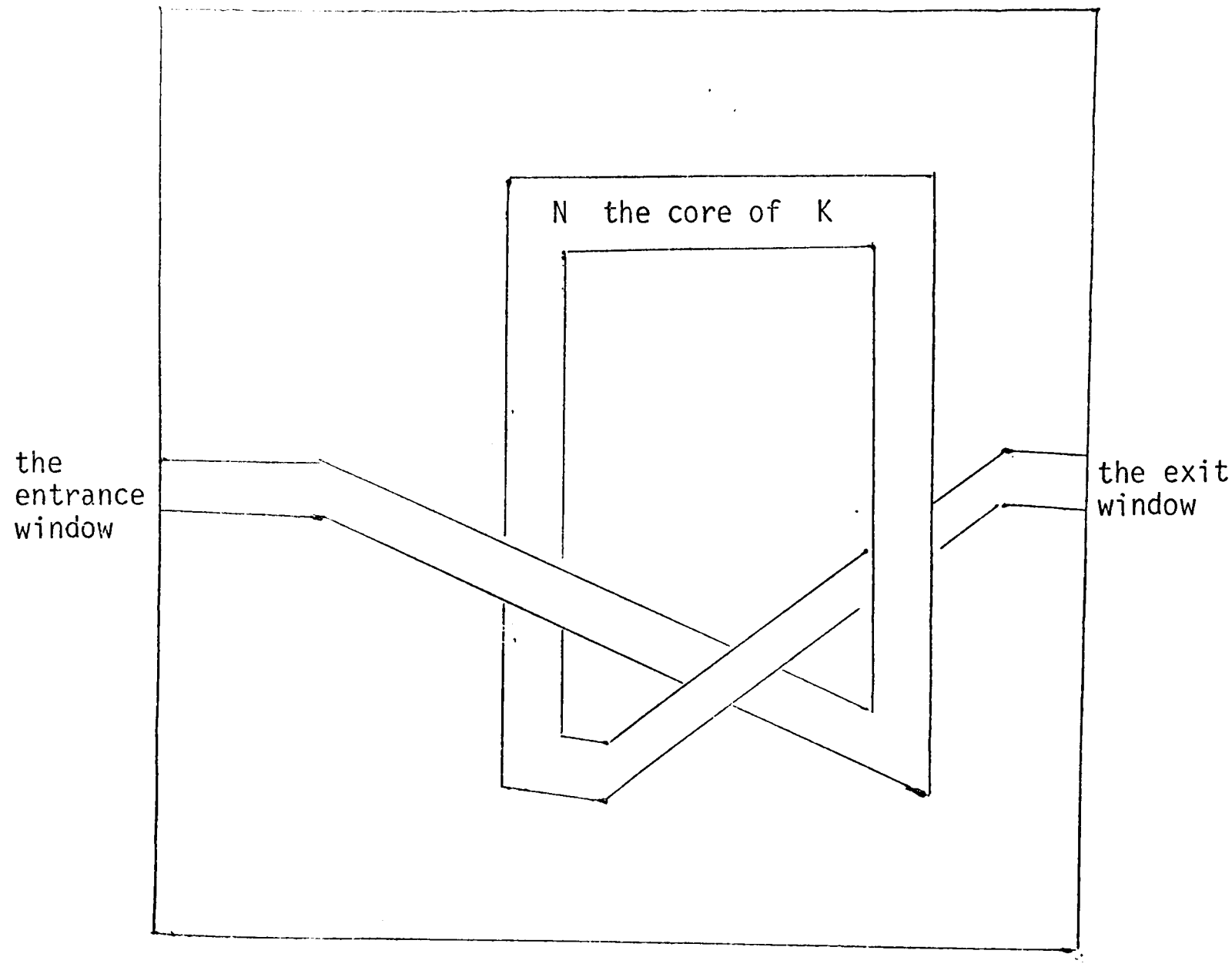


Figure to be attached after page 69 , section 2.2

Diagram of a knot-tube based on a map k corresponding to a trefoil knot.

continuous map (possibly with multiple points) such that $d(h(u), 0)$ tends to infinity as the absolute value of u tends to infinity. The map h is said to be <u>implicated</u> in the knot-tube K if there exists t in $\mathbb{R}$ such that

(i) $h(t)$ belongs to the core of K ;

(ii) $h(s)$ does not meet the entrance window for $s > t$;

(iii) $h(s)$ does not meet the exit window for $s < t$;

(iv) h never intersects K save in the window subsets.

These four conditions make precise the idea of h winding once through the knotted torus. The reasons for (i) and (iv) are clear. Conditions (ii) and (iii) ensure that h cannot unwind itself having once wound through K and cannot wind itself through K more than once. Note that the path will intersect each window if it is implicated and it may do so infinitely often.

To justify the claim that implication bears some resemblance to knotting some topological results concerning

implication follow.

Result

If h is implicated in a knot-tube based on $(k, \mathcal{E})$ and piecewise-linear and if h is simple, and intersects at least one of the windows once only then h may be 'combed out' by an isotopy so as to contain the knot k in the sense of Schubert's factorisation of knots (See Fox (1962) p 140 f).

Proof

Suppose that h intersects the entrance window once only. (If it is the exit window that is intersected once only then the proof is easily modified.) The crux of the proof is to show that the conditions on the knot-tube allow the core to be "slid out" of the knot-tube.

Recall the definition of k_p in condition (iii) of the definition of a knot-tube. If N_p is that part of the core that is closer to the image of k_p than to any image of k_q for $q \neq p$, then the conditions for a knot-tube ensure that $\overline{N}_p$ is the bounded section of a cylinder cut off by two planar slices. Moreover by condition (iii) the intersections of the slices with the

cylinder are disjoint. This enables one to define a piecewise-linear homeomorphism between $\bar{N}$ and the cylinder $\{(x,y) : x^2+y^2 \leq 1\} \times [0, 1]$ by using affine maps

$$H_p : \{(x,y) : x^2+y^2 \leq 1\} \times [p/s, (p+1)/s] \longrightarrow \bar{N}_p$$

that map the disc-shaped ends of the cylinder onto the parts of $\bar{N}_p$ that are the sections by planes, and such that the affine maps agree on the intersections of $\bar{N}_p$ with $\bar{N}_{p+1}$ for $p = 0, 1, \ldots, s-1$.

Let the piecewise-linear homeomorphism be denoted by H . Then H can be used to construct a family of isotopies I_t that squeeze most of the core up against the exit window but leave everything outside the core fixed. Take I_t to be the map that sends the point $H(x,y, v)$ to the point $H(x,y,v.(1+t.r)/(1+t.r.v))$ where $r = 1 - x^2 - y^2$, and leaves all other points fixed. Then the I_t do provide a family of isotopies and for large t the I_t "slide" the knot h right over to the exit window leaving behind a piecewise-linear curve that is a slight perturbation of k .

By this means the curve h can be seen to contain the knot of k . Note it is essential that one of the windows

is intersected only once. For otherwise the sliding maps leave behind a number of nearly parallel copies of k in the larger part of the core, and in general it will not be possible to unknot these copies. So in general the intersection requirement is necessary to ensure that k is contained in the Schubert factorisation of h .

□

This result shows how in a simple case the idea of implication relates to the usual definitions of knot theory. The next result gives a more general property of implication, describing its effect on a self-intersecting curve.

Result

If h is any curve implicated in a knot-tube based on (k, Σ) then the homomorphism of knot-groups

$$\pi_1(C \setminus \mathrm{Im}\, k) \longrightarrow \pi_1(\mathbb{R}^3 \setminus \mathrm{Im}\, h)$$

induced by the inclusion map (for A the knot-tube minus the core)

$$A \longrightarrow \mathbb{R}^3 \setminus \mathrm{Im}\, h$$

is an injection. Thus the "knot" of h in $\mathbb{R}^3$ inherits complexity from the knot of k in C .

Proof

The proof uses Van Kampen's theorem, a proof of which may be found in Crowell and Fox (1963) .

If A is the knot-tube minus the core then it is straightforward to prove that

$\pi_1(A) = \pi_1(C \setminus \mathrm{Im}\ k)$. The idea of the proof is to use the decomposition

$$\mathbb{R}^3 \setminus \mathrm{Im}\ h = A \cup (\mathbb{R}^3 \setminus A \setminus \mathrm{Im}\ h)$$

and apply Van Kampen's theorem to it. The decomposition does not quite satisfy the conditions of the theorem but a little topological manipulation allows for this; the problem is that A is not open but the topological argument still holds if A is enlarged slightly and its interior used. The conclusion of the theorem is that

$$\pi_1(\mathbb{R}^3 \setminus \mathrm{Im}\ h)$$

is the free product with amalgamated subgroup of

$$\pi_1(A) \quad \text{with} \quad \pi_1(\mathbb{R}^3 \setminus A \setminus \mathrm{Im}\ h)$$

amalgamating the subgroup the fundamental group of the intersection $\pi_1(\partial A)$ where ∂A is the boundary of the set A .

This would settle the theorem if only it can be shown that the amalgamation of $\pi_1(\partial A)$ in the product

does not lead to nontrivial elements of $\pi_1(A)$ being made equal to the identity in the product. Only elements in $\pi_1(\partial A)$ can suffer this fate and it can be shown that any such element will be trivial in $\pi_1(A)$ or else in $\pi_1(\mathbb{R}^3 \setminus A \setminus \mathrm{Im}\, h)$. This can be shown as indicated below.

Suppose that α and β are loops in ∂A that generate the fundamental group, chosen so that β winds once round h and is contained in the entrance window while α winds once round a small perturbation of β lying inside A and does not wind round h at all. The loops are sketched on page 77 in the trivial case when the knot-tube is built on a trivial knot.

Using compactness and approximation it can be shown that if β is null-homotopic in $\mathbb{R}^3 \setminus \mathrm{Im}\, h$ then it is null-homotopic in $\mathbb{R}^3 \setminus \mathrm{Im}\, h^*$ for some closed simple piecewise-linear curve h^* winding once through the torus ∂A. But this can be shown to be a contradiction using the idea of the linking number of β with h^*. Likewise if $n \neq 0$ then β^n can be shown to be nontrivial.

Similar ideas of linking number can be applied to the loop α^n to show that it is not null-homotopic

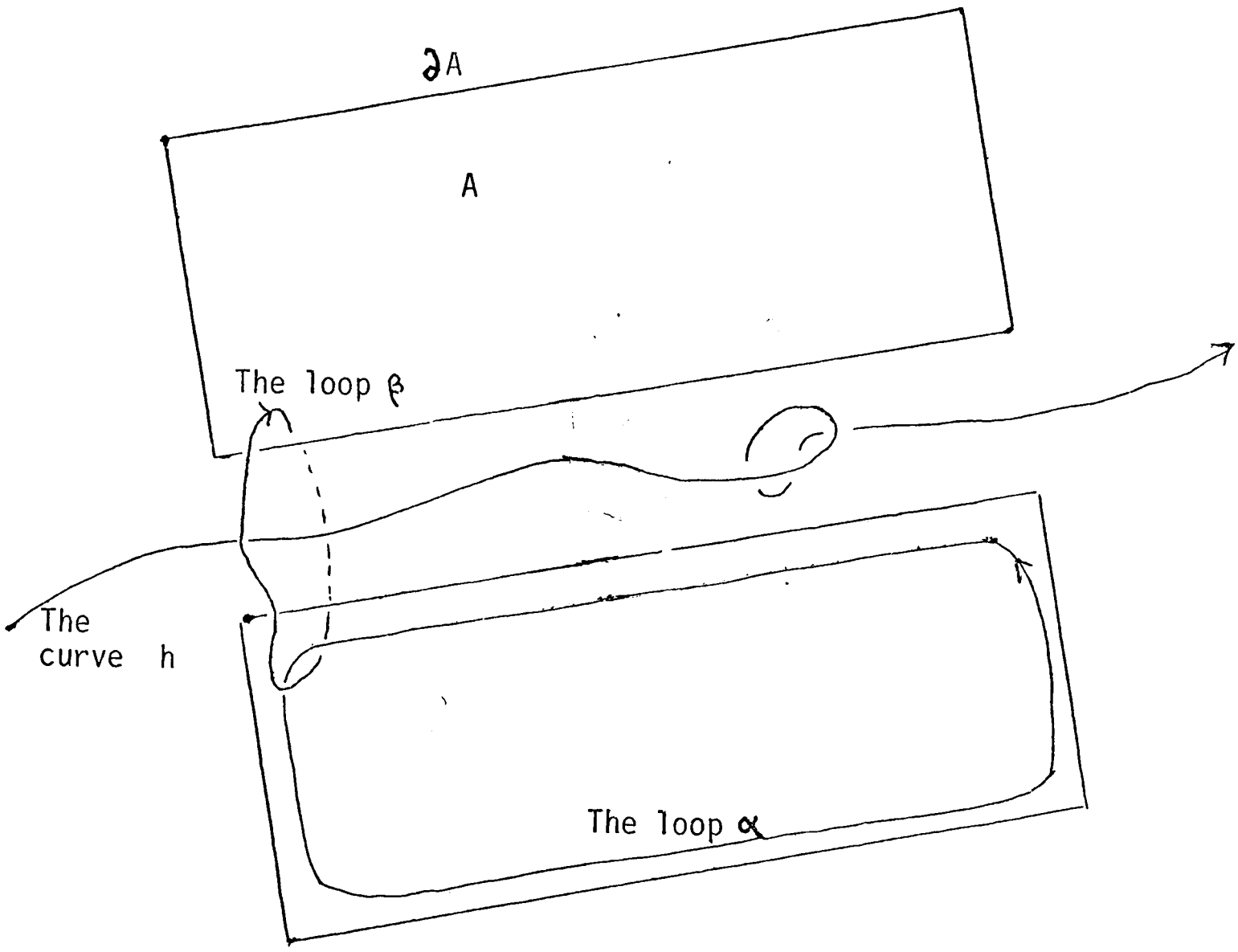


Figure to be attached after page 76 , section 2.2

Sectional diagram to illustrate the loops generating the fundamental group of the torus ∂A .

Note that intuitively if h winds through the torus then β should not be trivial in $\pi_1(\mathbb{R}^3 \setminus \text{Im } h)$. There is no loss of generality in taking the torus to be unknotted except that in general α may not be trivial in A .

in $\mathbb{R}^3 \setminus A \setminus \text{Im } h$ unless $n = 0$. For unless $n = 0$ the curve α^n will have a non-zero linking number with the small perturbation of β mentioned.

It follows that $\pi_1(\mathbb{R}^3 \setminus \text{Im } h)$ contains a copy of $\pi_1(C \setminus \text{Im } k)$.

□

The above result shows that if k is a nontrivial knot in C then h cannot be an unknotted map of $\mathbb{R}$ into $\mathbb{R}^3$ For if that were the case then $\pi_1(\mathbb{R}^3 \setminus \text{Im } h)$ would be trivial, while $\pi_1(C \setminus \text{Im } k)$ could not be trivial being the knot group of a nontrivial tame knot. (Crowell and Fox (1963) p 164)

Clearly the above can be extended to show that implications of disjoint knot-tubes imply appropriate complexity in the knot group.

The final result is not strictly to do with the topological properties of implication but will be of use when Brownian motion is discussed.

Result

There is a countable family of knot-tubes such that if a curve is implicated in a general knot-tube based on (k, Σ)

then the curve is also implicated in a knot-tube of this family based on a knot-pair of which the knot has the same knot-type as that of k .

Proof

Suppose h is implicated in a knot-tube based on (k, ε) then it will also be implicated in one based on (k, ε') for some rational ε' smaller than ε . Moreover a knot k' can be chosen piecewise-linear with respect to a partition $0 = r'_0 < r'_1 < \ldots < r'_s = 1$ of rational numbers such that $k'(r'_1), \ldots, k'(r'_{s-1})$ are all vectors of rational numbers , such that k' has the same knot type as k and such that h is implicated in (k', ε'). All this is straightforward from the density of the rationals in $\mathbb{R}$. Thus the family of knot-tubes mentioned in the result may be taken to be the family of all rational dilations and translations of knot-tubes based on such (k', ε').

□

2.3 The Brownian path.

Implication has been defined for curves in 3-space that extend out to infinity. To discuss the implication properties of Brownian motion, therefore, the process obtained by tying two Brownian motions head-to-head will be considered. To be precise, let B be a Gaussian process with state space $\mathbb{R}^3$, time space the whole of $\mathbb{R}$, defined on a probability space $(\Omega, \mathcal{F}, P)$ and satisfying

(i) $B(0) = 0$ and $E[B(t)] = 0$ for all t;

(ii) if B_i is the i^{th} coordinate process then the processes B_1, B_2, B_3 are independent and

$$E[B_i(t)\, B_i(t)] = \begin{cases} |t| \wedge |s| & \text{if } t, s \\ & \text{are of the same sign;} \\ 0 & \text{otherwise;} \end{cases}$$

(iii) with probability one the map $t \mapsto B(t)$ is continuous.

The random process B restricted to $[0, \infty)$ in time space is a standard Brownian motion which is independent of $B(t)$ for $t \leq 0$. The process B restricted to $(-\infty, 0]$ has similar properties.

As shown in Itô and McKean (1965) the Brownian motion in 3-space wanders out to infinity. Consequently the curve B satisfies the conditions necessary to be a candidate for implication.

If the implication of B in knot-tubes is to be discussed then it must be shown that implication in a knot-tube is a measurable property of B. The event that B is implicated in a knot-tube K is the intersection of the events

$$\left\{\begin{array}{l} B \text{ hits both entrance and exit windows but} \\ \text{nothing else of } K \text{ .} \end{array}\right\}$$

and

$$\left\{\begin{array}{l} \text{the greatest } u \text{ with } B(u) \text{ in the entrance} \\ \text{window is less than the least } v \text{ with } B(v) \\ \text{in the exit window and between these two times} \\ \text{the path } B \text{ belongs to the core of the knot-tube.} \end{array}\right\}.$$

These events are measurable since last exit and first hitting times are measurable. So the event of implication is measurable.

From the last result of section 2.2 it can be deduced that the implication of B in a general knot-tube based on a knot which is nontrivial defines a measurable event.

Given a fixed knot-pair (k, Σ) let K be the knot-tube based on (k, Σ) centred at the origin and of side length 2 .

Theorem

The curve B is implicated in infinitely many disjoint knot-tubes which are translates of K .

Proof(*)

Suppose that $T_0 = 0$ and that T_n is defined by

$$T_n = \inf \{ t > T_0 : B_1(t) = 4n \}$$

for $n \geq 1$. Let K_n be the knot-tube derived from K by translation through the vector

$$(1, 0, 0) + B(T_n) .$$

Then K_n and the origin are separated by the plane $x_1 = 4n-1$. Moreover the arc of B restricted to $[0, T_n)$ lies outside K_n , being on the same side of the plane $x_1 = 4n$ as the origin.

Let A_n be the event

$$\left\{ \begin{array}{ll} d(B(u+T_n) , k(u) + (1,0,0) + B(T_n)) < \Sigma & \text{when } u \in [0,1] \\ d(B(u+T_n) , (u+1,0,0) + B(T_n)) < \Sigma & \text{when } u \in [1,3] \\ d(B(u+T_n) , (1,0,0) + B(T_n)) > \sqrt{3} & \text{when } u \geq 3 \end{array} \right\} .$$

(*) The later stages of this proof were originally more involved. Simplifications were suggested by D.Aldous and H.Kesten and are incorporated here.

By the Forgery Theorem, the wandering out to infinity of 3-space Brownian motion and the Strong Markov property it follows that

$$P(A_n) = P(A_1) > 0 .$$

That the event A_n holds is not sufficient to ensure that B is implicated in K_n since B restricted to the negative part of the time axis is not constrained from hitting K_n by A_n. So define

$$C_n = \left\{ d(B(s), B(T_n) + (1,0,0)) > \sqrt{3} \text{ when } s \leq 0 \right\} .$$

Then B will be implicated in K_n if $A_n \cap C_n$ occurs. By scaling and the fact that 3-space Brownian motion wanders out to infinity it is clear that

$$P(C_n) \geq \rho > 0$$

where ρ is the probability that B restricted to the negative part of the time axis does not hit the sphere centred at $(5,0,0)$ and of radius $\sqrt{3}$.

By the Strong Markov property A_n and C_n are independent conditional on $B(T_n)$ and so

$$P(A_n \cap C_n \mid B(T_n)) = P(A_n \mid B(T_n)).P(C_n \mid B(T_n)) .$$

Therefore by scaling

$$P(A_n \cap C_n) \geq P(A_n)\rho .$$

So by Fatou's lemma

$$P\left(\limsup A_n \cap C_n \right) \geq \limsup P(A_n \cap C_n) \geq P(A_1) \cdot \rho > 0 \quad .$$

If $\mathcal{G} = \bigcap_{u \geq 0} \sigma\{ B(s) : |s| > u \}$, so that $\mathcal{G}$ is the 'double-tail σ-field' of the process B , then $P(F)$ is zero or one for all F in $\mathcal{G}$. This is a consequence of the 0-1 principle of Blumenthal (Itô and McKean (1965)) applied to the six-dimensional Brownian motion defined for positive u by

$$u \mapsto u \cdot (B_1(1/u), B_2(1/u), B_3(1/u), B_1(-1/u), B_2(-1/u), B_3(-1/u)) \quad .$$

Thus the theorem will be proved if it can be shown that $A = \limsup A_n \cap C_n$ belongs to $\mathcal{G}$, since it has been shown above that $P(A) > 0$.

No generality is lost if Ω is taken to be the path space $C_3(\mathbb{R})$ of continuous functions on taking values in $\mathbb{R}^3$. Given $u > 0$ then suppose that f is in A and that g in $C_3(\mathbb{R})$ is such that g agrees with f on $\mathbb{R} \setminus [-u, u]$. Since f

and g agree save on a compact set they must agree as to whether to belong to $A_n \cap C_n$ for sufficiently large n. For $A_n \cap C_n$ concerns the behaviour of the path in the region of space where $x_1 \geq 4n-1$.

Of course how large "sufficiently large" must be depends on f and g. However this does not matter in regard to A. So g must belong to A as well as f. Since whether or not g is in A depends only on the restriction of g to $\mathbb{R} \setminus [-u, u]$, it follows that $A \in \sigma\{B(s) : |s| \geq u\}$.

Since u was arbitrary it follows that A belongs to $\mathcal{G}$ and so the theorem is proved. The probability $P(A)$ is zero or one and must be positive, so $P(A) = 1$.

□

There follow three corollaries.

Corollary 1

Given a sphere containing the origin, B is implicated in infinitely many disjoint knot-tubes contained in the sphere and based on a given (k, ε).

Proof

Given $a > 1$ let $p(a)$ be the probability of the existence of a knot-tube based on (k, ε) implicating the curve B and located so as to be contained in the sphere of radius a and centre the origin but also so as to contain the sphere of radius 1 and centre the origin. Then by the theorem $\lim_{a \to \infty} p(a) = 1$. So there exists a monotonic increasing sequence a_n such that

$$p(a_n) > 1 - 2^n .$$

By the scaling properties of Brownian motion $p(a_n)$ is the probability of the existence of a knot-tube based on (k, ε) and implicating B, contained in the sphere of radius $(a_1 a_2 \ldots a_{n-1})^{-1}$ centred at the origin and lying outside the sphere of radius $(a_1 a_2 \ldots a_n)^{-1}$ centred at the origin. Consequently by the Borel-Cantelli lemma the probability of infinitely many disjoint knot-tubes implicating B inside any sphere centred at the origin must equal one. □

From this corollary it follows by translational invariance and the Markov property that the Brownian path

is implicated in infinitely many disjoint knot-tubes in every time interval. Moreover every segment of the path is knotted in this sense. Thus the corollary establishes a version of Edwards'conjecture.

Corollary 2

The path B is implicated in knot-tubes based on all possible knot-pairs.

Proof

This uses the countable family defined in the last result of section 2.2 . The proof of the theorem can be modified so as to show that the Brownian path will be implicated in dilated translations of all of these knot-tubes. An approximation argument shows that the path will be implicated in knot-tubes based on all possible knot-pairs.

□

Corollary 3

If B is implicated in a knot-tube K then there exist $u > \sup \{ t : B(t) \in K \}$ and $v < \inf \{ t : B(t) \in K \}$ such

that $B(u) = B(v)$. Consequently the motion may be said to complete each implication in that it achieves the implication in a closed curve.

Proof

It suffices to show that for any positive U there exist $u > U$ and $v < -U$ such that $B(u) = B(v)$. From the work of Frisedt (1967) the set

$$\{ B(w) : w < -U \}$$

will have positive capacity almost surely. Hence with positive probability $\{B(w) : w < -U\}$ will be hit by the process B restricted in time space to $[U, \infty)$. By scaling it is clear that the probability of such a hit being made does not depend on U . The event that such hits occur for arbitrarily large values of U is an event that belongs to the double tail σ-field $\mathcal{G}$ defined in the proof of the theorem of this section. So by Fatou's lemma and the degeneracy of the law P on $\mathcal{G}$ it follows that the probability of such a hit being made must be 1 .

□

2.4 Higher dimensions.

In dimensions higher than 3 piecewise-linear simple curves always unknot. However in any dimension higher than 2 it is not always the case that a simple curve is homeomorphic in placement to a piecewise-linear curve. A curve which is not amenable in this way is called wild .

Because of this although Dvoretsky, Erdös and Kakutani (1950) show that Brownian motion in 4-space and above is simple the only topological question to be asked is whether such random curves are wild. This question is partly settled by the work of Milnor (1964) who shows by straightforward techniques that with probability one the curve of Brownian motion in 5-space and above is not wild. The case of 4-space is not settled. The problem is that because the curve of Brownian motion in 3-space is not simple the projection technique used by Milnor no longer applies in the 4-space case.

This section contains one partial result on the case of Brownian motion in 4-space. To state this a topological definition must first be made. A set F in Euclidean space is said to be 1-ULCC (of complement that

is uniformly locally 1-connected) if for every ε there is a η such that for any x in the complement of F

if $B_x(\varepsilon)$, $B_x(\eta)$ are the balls centred on x of radii ε and η then for any closed loop γ in $B_x(\eta) \cap F$ the loop is contractible to a point in $B_x(\varepsilon) \cap F$.

Result

With probability one the curve of Brownian motion in 4-space is 1-ULCC .

Proof

This depends on the fact that Brownian motion in 4-space avoids any specified 2-plane with probability one. Let G be the union of all 2-planes in the 4-space that are defined by three non-collinear points each of rational coordinates. Then because G is a countable union of 2-planes it is avoided by Brownian motion with probability one , except at the origin.

For a given ε suppose that a point x is at a positive distance from the Brownian motion . For any loop in $B_x(\varepsilon)$ the loop is contractible within $B_x(\varepsilon)$. If the loop is missed by the motion then the loop and the motion are separated by a distance of δ , say. It will be convenient to reduce δ so that also the loop and the complement of $B_x(\varepsilon)$

are separated by a distance of at least δ .

The contraction of the loop is essentially a map from a disc to 4-space agreeing with the loop map on the boundary of the disc and lying entirely within $B_x(\varepsilon)$. By straightforward approximation this map can be approximated to within $\delta/2$ by a piecewise-linear map from the disc to the ball $B_x(\varepsilon)$. Moreover the piecewise-linear map may be taken to have range lying within G , and so its range will be avoided by the Brownian motion. But because the restriction of the approximating map to the boundary of the disc will be a $\delta/2$ -approximation to the loop and the Brownian motion comes no closer than δ to the loop it follows that by a linear interpolation the piecewise-linear map may be extended to form a homotopy of the loop to a point in $B_x(\varepsilon)$ that avoids the Brownian motion. This concludes the proof. □

Clearly the Brownian motion would be wild if it were not 1-ULCC , since if it were not 1-ULCC then it could not possibly be homeomorphic in placement to a piecewise-linear curve. Whether or not the converse follows does not appear to be known. The concept of 1-ULCC has been used in the

study of topological embeddings (see for example Hempel and McMillan (1966)) but the case of curves in 4-space seems to be intractable.

In conclusion to this section and the chapter it should be mentioned that the work of Milnor (1964) also treats the question of the Baire category of the ensembles of wild curves in the Euclidean spaces. Indeed this forms the main part of his paper. Baire category results are easier to obtain because in the sense of Baire category almost all curves in 3-space are simple.

Chapter 3 : The contours of Brownian processes with multidimensional times.

3.1 This chapter concerns itself with generalisations of Brownian motion to processes with several-dimensional times (also called random fields). Interest will be focussed on the topological nature of the level-sets of these processes. For the purpose of illustration consider the case of two-dimensional time. Then the random field can be thought of as assigning a scalar to each point of a two-dimensional space, thus providing a random map of the space to the real line. Indeed the process can provide a 'random landscape' if the value of the process at a particular point is thought of as the height of the landscape at that point. This chapter will consider the topological behaviour of the contours of such a random landscape.

The first process to be considered will be the Lévy generalisation of Brownian motion described in Lévy (1965 ch viii). For the sake of brevity this will be referred to subsequently as the Lévy process. It is defined as follows. Let $(\Omega , \mathcal{F} , P)$ be a complete probability space; then

$$X : \Omega \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

is said to be a Lévy process if

(i) the family $\{ X(u) : u \in \mathbb{R}^n \}$ is a Gaussian family of random variables;

(ii) $X(\underline{0}) = 0$, $E[X(u)] = 0$ and

$$E [X(u) - X(v)]^2 = \| u - v \|$$

for all u and v in $\mathbb{R}^n$ and for $\|\cdot\|$ the Euclidean norm on $\mathbb{R}^n$;

(iii) for almost all ω in Ω the realisation X_ω is a continuous map from $\mathbb{R}^n$ to $\mathbb{R}$.

That X can exist as prescribed is not trivial. A proof of existence including the continuity property (iii) may be found in Lévy (1965). The process is of interest in pure mathematics because of the remarkable isometric embedding of $\mathbb{R}^n$ into Hilbert space that it provides via property(ii), and also because of its relevance for some group representation theory (see Takenaka (1977)) . It is of use in applied mathematics insofar as it affords a simple example of a random field to compare with landscapes, turbulent behaviour etc. For employment of the Lévy process in such a role see Mandelbrot (1977) .

The other process to be considered is the Brownian sheet, first constructed by Chentsov (1956) . Take $(\Omega, \mathcal{F}, P)$ as above. Then the random field

$$W : \Omega \times [0,\infty)^n \longrightarrow \mathbb{R}$$

is a Brownian sheet with n-dimensional time if

(i) the family $\{W(s) : s \in [0,\infty)^n\}$ is a Gaussian family of random variables;

(ii) $W(s) = 0$ if one of the components of the vector

$s = (s_1 , \ldots , s_n)$ is zero, $E [W(s)] = 0$ and

$$E [W(s) . W(t)] = (s_1 \wedge t_1) \times \ldots \times (s_n \wedge t_n)$$

for all $s = (s_1 , \ldots , s_n)$ and $t = (t_1 , \ldots , t_n)$ in $[0, \infty)^n$;

(iii) for almost all ω in Ω the realisation W_ω is a continuous map from $[0, \infty)^n$ to $\mathbb{R}$.

Pyke (1973) indicates how W can arise as the weak limit of partial sums of matrix arrays and gives references to various constructions of W . He also suggests questions concerning the properties of the zeroset of W , that is the level-set on which W is zero. In response to this Adler (1978b) has shown that with probability one for all x the level-set of W on which W equals x has uniform Hausdorff dimension $n - \frac{1}{2}$. However this does not say much about the topological nature of the level-sets. For example, if 'contours' are defined to be the connected components of the level-sets then the question of the boundedness of the contours remains open. This question and the corresponding question for the Lévy process are both attacked in this chapter. Although a complete answer is not obtained it is shown that the level-sets and contours exhibit rather remarkable topological behaviour.

Before embarking on the analysis of the two processes it is worth remarking on their differences. The Brownian sheet has a

time space with a boundary. Moreover its covariance structure is clearly related to the cartesian axes. On the other hand the Lévy process not only has no boundary but is isotropic, being unaltered in law under rotation of the time space. This makes a difference in the analyses of the two processes and so they have to be treated seperately.

3.2 Analysis of the Lévy process.

Let X be a Lévy process as defined above. In the sequel the symmetry and scaling properties of X will be of great use and it is convenient to enumerate them as follows;

(i) The process X has 'stationary increments' . That is to say, for any y in $\mathbb{R}^n$ the process

$$x \longmapsto X(x+y) - X(y)$$

is a Lévy process as defined above;

(ii) if R is a rotation of $\mathbb{R}^n$ fixing the origin then

$$x \longmapsto X(Rx)$$

is once again a Lévy process;

(iii) for any r in $[0, \infty)$, the scaled process

$$x \longmapsto X(rx).r^{-\frac{1}{2}}$$

is a Lévy process;

(iv) the process produced by inversion

$$x \longmapsto \begin{cases} 0 & \text{if } x = 0 , \\ x \, . \, X(x/\|x\|^2) & \text{otherwise} \end{cases}$$

is a Lévy process.

Proofs of these properties follow from computation of the covariance structures. The only complication is in dealing with the inverted process of (iv). It is necessary

to consider the variance

$$V = E \left[\|x\| . X(x/\|x\|^2) - \|y\| . X(y/\|y\|^2) \right]^2 .$$

A rotation can be found that interchanges $x / \|x\|$ with $y / \|y\|$ (in the case $n = 2$ this may have to be a reflection) and use of property (ii) with this rotation shows that

$$V = E \left[\|x\| . X(y/\|x\|.\|y\|) - \|y\| . X(x/\|x\|.\|y\|) \right]^2 .$$

Expanding the square and using the scaling property (iii) ,

$$V = \|x\| + \|y\| - 2.E\left[X(x).X(y)\right]$$

$$= E\left[X(x) - Y(y)\right]^2 .$$

Continuity of the new process follows because it is continuous away from the origin and has the same law as the Lévy process away from the origin. Hence one may deduce

$$\lim_{x \to 0} \|x\| . X(x/\|x\|^2) = 0 .$$

These properties, clearly analogous to those of the classical Brownian motion, are well-known in the literature .

The next task is to define the level-sets and contours of the Lévy process and to state the problem which this section addresses. Given $x \in \mathbb{R}^n$ the level-set of X at x is defined to be the set

$$L(x) = \{ y \in \mathbb{R}^n : X(y) = X(x) \} .$$

The contour of X at x is defined to be the set $C(x)$ which

is the connected component of $L(x)$ containing x . By the continuity of X both $L(x)$ and $C(x)$ are closed sets.

When $n = 2$, so that X can be thought of as providing a random landscape, then the contours can be thought of as contours in the cartographical sense. In general the contours will be the maximal connected sets on which X is constant.

The result to be proved falls into two parts. Firstly it will be shown that with probability one every contour is bounded. Secondly an application of the inversion property will show that for any x in $\mathbb{R}^n$ the contour $C(x)$ is trivial. The second part does <u>not</u> imply that every contour is trivial - but this will be elaborated on in section 3.3 .

The idea of the proof involves "fencing in" contours by finding spheres on which X must be greater than a certain value. So the following preliminary lemma will be of assistance;

<u>Lemma</u>

Let S^{n-1} be the unit sphere centred at the origin of $\mathbb{R}^n$. Then the probability

$$P\{ X(u) > 0 \text{ for all } u \text{ in } S^{n-1} \}$$

is positive.

Proof of Lemma

Let Y be the average value of X taken over the surface of S^{n-1}. So

$$Y = \int_{S^{n-1}} X(u)\, dm(u)$$

where m is the surface measure for the sphere S^{n-1} such that $m(S^{n-1}) = 1$. The key to the proof is that Y is a non-degenerate Gaussian random variable that is independent of the random function

$$x \mapsto X(x) - Y \text{ . for } x \text{ in } S^{n-1} .$$

Because the random variable Y belongs to the Hilbert space of random variables generated by the Gaussian family $\{X(x) : x \in \mathbb{R}^n\}$ it is clear that it is Gaussian. Moreover because it is jointly Gaussian with X the independence follows if Y can be shown to be uncorrelated with $X(x) - Y$ when $x \in S^{n-1}$.

By the rotational invariance of the Levy process and of m the number $E[X(x).Y] = \int_{S^{n-1}} E[X(x).X(u)]\, dm(u)$ does not depend on x. Consequently

$$E(Y.Y) = \int_{S^{n-1}} \int_{S^{n-1}} E[X(x).X(u)]\, dm(u)\, dm(x)$$
$$= E[X(x).Y]$$

and so

$$E\{Y . [X(x) - Y]\} = 0 .$$

Indeed this rotational invariance also means Y must be non-degenerate for

$$E(Y.Y) = E[X(x).Y] = \int_{S^{n-1}} E[X(x).X(u)]\, dm(u)$$

and if $x \in S^{n-1}$ then

$$E\left[X(x).X(u)\right] = 1 - \tfrac{1}{2}.\|x - u\| \quad \text{is always positive.}$$

So Y is independent of $X(x) - Y$ when $x \in S^{n-1}$ and is a non-degenerate Gaussian random variable. But the random function $x \mapsto X(x) - Y$ is continuous and so has a minimum value M for x in the compact set S^{n-1}. Since Y and M are independent,

$$P\{ Y < z \mid M = z \} = P\{ Y < z \}$$

and because Y is non-degenerate and Gaussian this last is positive for all z. Consequently

$$P\{ Y < M \} > 0$$

and so with positive probability

$$Y < X(x) - Y \quad \text{for all } x \text{ in } S^{n-1}.$$

Subtracting Y from both sides yields the proof of the lemma.

□

Suppose that the event mentioned in the lemma occurs, so that

$$X(u) > 0 \quad \text{for all } u \text{ in } S^{n-1}.$$

Then the intersection

$$L(0) \cap \{ x : \|x\| < 1 \}$$

will be both relatively open and relatively closed as a subset of $L(0)$.

This means that $C(0)$ will be contained in $L(0) \cap \{ x : \|x\| < 1 \}$ by the definition of a connected component. So if the event occurs then $C(0)$ must be a bounded subset of $\mathbb{R}^n$.

Moreover, application of the scaling property (iii) shows that the probability

$$P \{ X(r.u) > 0 \text{ for all } u \text{ in } S^{n-1} \} = p$$

does not depend on the positive scalar r . This corollary sets the scene for the main result.

Theorem

With probability one all the contours of X are bounded. Moreover, given x in $\mathbb{R}^n$, with probability one the connected component $C(x)$ is trivial.

Proof

Let $r_k = \exp(2^k)$. Then as shown above

$$P \{ X(r_m.u) > 0 \text{ for all } u \text{ in } S^{n-1} \} = p > 0$$

for some positive p that does not depend on m . Indeed by Fatou's lemma if

$$A = \lim_{k \to \infty} \bigcup_{m=k}^{\infty} \{ X(r_m.u) > 0 \text{ for all } u \text{ in } S^{n-1} \}$$

then $P(A) \geq p$. Moreover, because $-X$ has the same law as X it follows that

$$P(\lim_{k \to \infty} \bigcup_{m=k}^{\infty} \{ X(r_m.u) < 0 \text{ for all } u \text{ in } S^{n-1} \})$$

$$= P(A) \geq p .$$

Arguments similar to those immediately preceding the theorem now show that if A occurs then all the contours of non-positive values of X must be bounded. And if the corresponding event (concerning X being negative on hyperspheres) occurs then all the contours of non-negative values of X must be bounded.

In fact the second part of the theorem is settled if p = 1 for then the inversion property (iv) and the property of stationary increments (i) may be invoked to show

$$\lim_{k \to \infty} \bigcup_{m=k}^{\infty} \{ X(x + r_m u) > X(x) \text{ for all } u \text{ in } S^{n-1} \}$$

must have probability one.

It is possible to prove that P(A)= 1 by appealing to a 0-1 law stated by McKean (1963) and proved by Molchan (1967) . This law says that P is degenerate on the tail σ-field

$$\mathcal{T} = \bigcap_{r>0} \sigma \{ X(y) : \|y\| > r \} .$$

Since A clearly belongs to $\mathcal{T}$ it follows that P(A) is zero or one. The above argument shows P(A) to be positive and so this proves the theorem. However the rest of the section will provide an alternative proof that does not appeal to this 0-1 law.

The alternative proof uses the idea exploited in McKean (1963) of expanding the process X using spherical harmonics, though it will not be necessary to use any results from the theory of spherical harmonics. Consider ϕ a function from the Hilbert space $L^2(S^{n-1}, m)$. This yields a new random process H_ϕ with

$$H_\phi(t) = e^{-t/2} \int_{S^{n-1}} X(e^t.u)\, \phi(u)\, dm(u)$$

for t in $\mathbb{R}$ and obviously the process H_ϕ is linear in ϕ.

These new processes can be used to define an operator, the covariance operator K_α, for each $\alpha > 0$ defined by

$$\begin{aligned} \langle \psi, K_\alpha \phi \rangle &= \int_{S^{n-1}} \psi . K_\alpha \phi \, dm \\ &= E\left[H_\psi(t-\alpha).H_\phi(t+\alpha)\right]. \end{aligned}$$

Clearly K_α is linear, and it does not depend on t because the H processes are stationary by the scaling property (iii) of X. In fact as is easily seen K_α is an integral operator with kernel

$$k_\alpha(u,v) = \cosh \alpha - \tfrac{1}{2} \| e^{-\alpha}.u - e^{\alpha}.v \|.$$

This kernel is a symmetric and positive function of u and v in S^{n-1}. Symmetry follows from the observation that a reflection interchanging u and v leaves the Euclidean distance $\| e^{-\alpha}.u - e^{\alpha}.v \|$, and hence the kernel, unchanged. Positivity is a consequence of the observation that the Euclidean distance attains its maximum possible value when u

equals $-v$. For this means that the minimum value of $k_\alpha(u,v)$ is $\cosh \alpha - \frac{1}{2} \| e^{-\alpha}.u + e^{\alpha}.u \| = 0$.

The family $\{ K_\alpha : \alpha \text{ in } \mathbb{R} \}$ commutes. For it is the case that

$$\int_{S^{n-1}} k_\alpha(u,w).k_\beta(w,v)\, dm(w) =$$
$$\int_{S^{n-1}} k_\alpha(u,R^{-1}w).k_\beta(R^{-1}w,v)\, dm(w)$$

for R a reflection permuting u and v , by the invariance of the surface measure m . But by the invariance of the Euclidean distance $k_\alpha(u,R^{-1}w) = k_\alpha(Ru,w) = k_\alpha(w,Ru)$
$= k_\alpha(w,v)$.

A similar computation for $k_\beta(R^{-1}w,v)$ shows that the integral must equal

$$\int_{S^{n-1}} k_\alpha(w,v).k_\beta(u,w)\, dm(w) \quad .$$

From these remarks it is clear that $K_\alpha K_\beta = K_\beta K_\alpha$ and so the family of integral operators must commute.

Because the family commutes it is possible to provide a complete orthonormal set of functions which are simultaneously eigenfunctions for all the K_α . This fact can be proved in a variety of ways- by appealing to the theory of spherical harmonics; by group representation theory; or by the theory of compact operators. This last being the most basic way is the way adopted in the following.

Since the kernel of K_α is continuous and S^{n-1} is compact the integral operator K_α is compact. It has already been noted that the operator is symmetric and so the spectral theory of compact self-adjoint operators may be applied. In essence this says that

the Hilbert space $L^2(S^{n-1},m)$ may be decomposed into an orthogonal sum of subspaces invariant under K_γ and such that K_α restricted to a summand acts simply as multiplication by a scalar. Moreover these subspaces may be chosen so that all but one are finite-dimensional and so that the possibly infinite-dimensional space is the kernel of K_α .

Because the K_β commute each of the subspaces of such a decomposion may be taken to be invariant under the action of the other K_β . What is desired is to show that the decomposition may be so refined that on each summand each K_β acts as multiplication by a scalar. This intuitive fact can be shown rigorously using Zorn's lemma. Consider the family of all orthogonal decompositions satisfying

(i) invariance under all K_β ;

(ii) having at most one infinite-dimensional summand;

(iii) for some β being a refinement of the eigen-decomposition for K_β described above.

If this family is ordered in the usual way under refinement then any totally ordered subfamily must be dominated by a more refined decomposition in the big family (property (ii) is essential here.). So by Zorn's lemma there is a maximal member of the big family. The action of any K_β on any summand of this maximal member must be one of multiplication for otherwise

the maximal member could be further decomposed by use of another eigendecomposition. If (ϕ_h) is an orthonormal basis for $L^2(S^{n-1},m)$ such that every ϕ_h belongs to one of the summands of this maximal decomposition then clearly

$$K_\beta \phi_h = \lambda_h(\beta)\phi_h \quad \text{for all } \beta .$$

Thus the possibility of providing a complete set of simultaneous eigenfunctions has been demonstrated.(*)

An exposition of the theory of compact self-adjoint operators can be found in Dieudonné (1969) chapter XI . Continuing to use the theory developed in that chapter, we see

$$\sup_h \lambda_h(\alpha)^2 \leq \sum_h \lambda_h(\alpha)^2$$

$$= \int_{S^{n-1}} \int_{S^{n-1}} |k_\alpha(u,v)|^2 \, dm(u)\, dm(v)$$

and this double integral is no more than

$$\min \left\{ e^{-2\alpha} , e^{2\alpha} \right\}$$

since the kernel $k_\alpha(u,v)$ is less than $\cosh(\alpha)$ and is positive.

These are the results necessary for the proof. For the fact that the ϕ_h are orthogonal eigenfunctions means that

$$E\left[H_{\phi_a}(s).H_{\phi_b}(t)\right] = \langle \phi_a, K_{\frac{1}{2}(s-t)}\phi_b \rangle$$

$$= \lambda_b(\tfrac{1}{2}(s-t))\langle \phi_a, \phi_b \rangle$$

$$= 0 \quad \text{if } a \neq b .$$

Consequently the Gaussian processes H_{ϕ_a} , H_{ϕ_b} are independent

---

(*) Arveson (1976) provides a general theorem containing this fact in his Theorem 1.4.4 .

when $a \neq b$. Moreover any one of these processes H_{ϕ_a} has the interesting property that

$$E\left[H_{\phi_a}(s).H_{\phi_a}(s+2\alpha)\right] = \langle \phi_a, K_\alpha \phi_a \rangle = \lambda_a(\alpha) \leq \min\left\{e^{-\alpha}, e^{\alpha}\right\}.$$

Consider then the continuous random function X restricted to the sphere S^{n-1} . Because the family (ϕ_h) is an orthonormal basis for $L^2(S^{n-1})$ it is possible to expand X as an L^2 convergent sum of these functions;

$$X = \sum_h H_{\phi_h}(1).\phi_h .$$

It follows that, for N a positive integer and ε a positive number, an integer j can be chosen and a subset G of $\mathbb{R}^j$ found so that

$$P\left(\{X(u) > 0 \text{ for all } u \text{ in } S^{n-1}\} \triangle \{(H_1(1),\ldots,H_j(1)) \in G\}\right) \leq \varepsilon/(2N+2) ,$$

where $\triangle$ is the set-theoretic operation of symmetric difference.

These results suggest that the behaviour of the process on the sphere of radius r_m will turn out to be only weakly dependent on the behaviour of the process on the sphere of radius r_{m-1}

This is indeed the case, and will be shown to be so using the work of Kolmogorov and Rozanov (1960) relating the maximum correlation coefficient between spaces of Gaussian random variables to Rosenblatt's coefficient of strong mixing. For set

$$A(j,N,a,b) = \bigcap_{h=a+1}^{a+b} \{ (H_1(2^h),\ldots,H_j(2^h)) \notin G \}$$

then the absolute difference

$$\left| P\left[\bigcap_{h=a+1}^{a+N} A(j,N,h-1,h)\right] - P\left[\bigcap_{h=a+1}^{a+N-1} A(j,N,h-1,h)\right] . P\left[A(j,N,a+N-1,a+N)\right] \right|$$

is shown by this work to be no greater than the maximum correlation coefficient between the two linear spaces of random variables spanned by $\{ H\phi_k(2^h) : k=1,\ldots,j \text{ and } h=a+1,\ldots,a+N-1 \}$ and $\{ H\phi_k(2^{a+N}) : k=1,\ldots,j \}$ respectively. But because the covariance of $H\phi_k(2^h)$ with $H\phi_{k'}(2^{a+N})$ is zero if $k \neq k'$ and bounded absolutely by

$$\min \{ \exp(2^{a+N} - 2^h) , \exp(2^h - 2^{a+N}) \}$$

otherwise, it is clear by elementary calculations that the maximum correlation coefficient will tend to zero as a tends to infinity.

A straightforward repetition of this kind of argument shows that for sufficiently large a

$$\left| P\left[A(j,N,a,N)\right] - \prod_{h=a+1}^{a+N} P\left[A(j,N,h-1,h)\right] \right| \leq \varepsilon/(N+1) .$$

Thus the suggestion of weak dependence has been made precise. To finish the proof of the theorem it suffices to recall that by the scaling property of the Lévy process the event $A(j,N,h-1,h)$ approximates to the event

$$\{ X(2^h.u) = 0 \text{ for some } u \text{ in } S^{n-1} \}.$$

For this together with a few manipulations of probability calculus shows that for sufficiently large a

$$P \left\{ \begin{array}{l} X(r_h.u) > 0 \text{ for all } u \text{ in } S^{n-1} \\ \text{for some } h \text{ from } a+1,a+2,\ldots,a+N \end{array} \right\}$$

is within

$$N\,\varepsilon/(2N+2) + \varepsilon/(N+1) + N\,\varepsilon/(2N+2) = \varepsilon$$

of the number

$$1 - \prod_{h=a+1}^{a+N} P \{ X(r_h.u) = 0 \text{ for some } u \text{ in } S^{n-1} \}$$

$$= 1 - (1-p)^N .$$

It follows that the probability of X being positive on at least one of the spheres of radii $r_1, r_2, \ldots$ wil be equal to 1, since $p > 0$ and ε is arbitrary. This finishes the proof of the theorem since it shows that the event A of the discussion at the beginning of the proof has probability one.

□

As is mentioned above, the theorem does not imply that all contours of X are trivial. However if U is the union of all non-trivial contours,

$$U \quad = \quad \{ x \in \mathbb{R}^n \ : \ C(x) \neq \{x\} \} ,$$

then almost surely U is a null-set. To state this as a corollary of the theorem,

Corollary

With probability one U has Lebesgue measure zero.

Proof

For any z in $\mathbb{R}^n$ the function

$$f(z) \quad = \quad P\{ z \in U \}$$

is zero, by the theorem. But then by Fubini's theorem

$$\int_{\mathbb{R}^n} 1_{\{z \in U\}} \, dz \quad = \quad \int_{\mathbb{R}^n} f(z) \, dz \quad = \quad 0 \ ,$$

thus proving the corollary.

□

3.3 A topological interlude.

The previous section has shown that the union of the non-trivial contours of X has Lebesgue measure zero. Naturally one asks whether it is possible for all the contours of X to be trivial. In fact such a phenomenon is impossible for topological reasons. For any continuous map X from $\mathbb{R}^n$ to $\mathbb{R}$ there must be nontrivial contours.

Take X a continuous map from $\mathbb{R}^n$ to $\mathbb{R}$. If X is constant then $\mathbb{R}^n$ is itself a connected level-set for X and thus a nontrivial contour. If X is not constant then there are u and v in $\mathbb{R}^n$ with $X(u) \not< X(v)$ and so by the continuity of X there is a w in $\mathbb{R}^n$ with

$$X(u) < X(w) < X(v) .$$

Moreover for such a w the level-set $L(w)$ must separate u from v. If it is impossible for totally disconnected closed sets (sets with their connected components all trivial) to separate points in Euclidean space then the above argument shows that $L(w)$ must have a nontrivial connected component.

To settle the question the problem is first reduced to one of considering <u>compact</u> totally disconnected sets. Let E_r be the closed ball of radius r and centred at the origin.

If $L(w)$ separates u from v in $\mathbb{R}^n$ then for any E_r containing both u and v the intersection $L(w) \cap E_r$ separates u from v in E_r. Moreover if $L(w)$ is totally disconnected then so is $L_r = L(w) \cap E_r$. Being bounded and closed, L_r must be compact. We consider such an L_r with $r > 1$.

The argument turns on showing that a totally disconnected compact L_r cannot separate u from v in E_r. Since L_r can have uncountably many components the compact nature of L_r must be exploited to replace it by a more amenable set.

Since neither u nor v belong to L_r there is a positive ε less than 1 such that u and v are further than ε from L_r. Given a particular x in L_r and an open ball U centred at x and of radius ε, a set V can be found, clopen (closed and open) relative to L_r, contained in $L_r \cap U$ and containing x. For if $y \in L_r \setminus U$ there is a clopen $V_y \subset L_r$ such that

$$x \in V_y \quad \text{and} \quad y \in L_r \setminus V_y .$$

The complements $L_r \setminus V_y$ of such V_y are also clopen in L_r and cover the compact set $L_r \setminus U$. So a finite number of complements

$$L_r \setminus V_{y_1} , \ldots , L_r \setminus V_{y_s}$$

can be found that together cover $L_r \setminus U$. But then the intersection

$$\bigcap_i V_{y_i} = V$$

is clopen relative to L_r and satisfies $x \in V$ and $V \subset L_r \cap U$.

A further application of compactness shows that L_r can be covered by a finite number of such clopen sets, say

$$V_1 , \ldots , V_m ,$$

such that each V_j can be contained in a ball of radius ε that itself contains neither u nor v .

This covering may be refined by considering the family of 2^m sets obtained by intersections of m sets where the i th set in the intersection is either V_i or $L_r \setminus V_i$ The member of the family which is $\bigcap_i L_r \setminus V_i$ is empty since the V_i cover L_r . The other members cover L_r , are all disjoint , are all clopen and are all contained in open balls of radius ε and centred on members of L_r .

Because the distances of u and v from L_r are greater than ε , neither u nor v belong to any open balls of radius ε and centered on members of L_r . Moreover such open balls cannot separate u from v in E_r (recall $r > 1$ and $\varepsilon < 1$) . A fortiori none of the constituents of the last covering of L_r can separate u from v . Consequently L_r is composed of a finite number of disjoint sets (clopen in the L_r topology and hence closed in the E_r topology) none of which separate u from v in E_r . Intuitively it is clear that L_r

itself cannot separate u from v .

To make this intuitive step rigorous is rather delicate. A proof follows which uses the theory of singular homology (for an exposition of this see Greenberg (1967)). A direct proof does not appear to be in the literature but the result is an almost immediate consequence of the homology theory used.

Clearly it suffices to show that if two disjoint closed subsets of E_r fail individually to separate u from v then so does their union. Suppose C_1 , C_2 are the two disjoint closed sets, with neither u nor v belonging to $C_1 \cup C_2$. Then the abelian groups $H_o\left[E_r \setminus (C_1 \cup C_2)\right]$, $H_o\,(E_r \setminus C_1)$, $H_o\,(E_r \setminus C_2)$ are the free $\mathbb{Z}$-modules generated by the path-connected components of $E_r \setminus (C_1 \cup C_2)$, $E_r \setminus C_1$, $E_r \setminus C_2$ respectively. Let $[u]$ and $[v]$ refer to the path-connected components containing u and v in the appropriate space (according to context) . An application of the Mayer-Vietoris sequence produces group homomorphisms such that

$$H_1\,(E_r) \longrightarrow H_o\left[E_r \setminus (C_1 \cup C_2)\right] \xrightarrow{\Phi} H_o(E_r \setminus C_1) \oplus H_o(E_r \setminus C_2)$$

is <u>exact</u> (that is, the image of the first map is precisely the kernel of the second) . The homology group $H_1\,(E_r)$ is rather

harder to define than the H_0 groups but in this case it does not matter; E_r is contractible and so $H_1(E_r) = 0$ the trivial abelian group.

The theory of the Mayer-Vietoris sequence describes the maps. The map ϕ is easy to describe; if $[z]$ is a path-connected component of $E_r \setminus (C_1 \cup C_2)$ then ϕ sends it to the pair $([z], -[z])$ where the first $[z]$ is the path-connected component in $E_r \setminus C_1$ containing z and the second $[z]$ is the corresponding component of $E_r \setminus C_2$. These evidently geometric definitions define the map because ϕ is a group homomorphism. On the other hand the first map is obvious in this example because $H_1(E_r)$ is trivial. The map is completely described by saying its range is 0 the zero element of $H_0[E_r \setminus (C_1 \cup C_2)]$.

Because of the exactness of the sequence, the triviality of the first map means that the kernel of ϕ must be trivial, so that ϕ is an injection. Consider the element $\phi([u] - [v])$. This equals $\phi[u] - \phi[v] = ([u] - [v], [v] - [u])$ which is zero because in either of $E_r \setminus C_1$ and $E_r \setminus C_2$ the elements u and v are in the same connected components. But because ϕ is an injection this must mean $[u] = [v]$ in

the space $E_r \setminus (C_1 \cup C_2)$. So u and v are in the same connected component of $E_r \setminus (C_1 \cup C_2)$, completing the proof that makes the intuitive step mentioned above rigorous.

It has been shown , then, that X has non-trivial contours. Indeed the argument for their existence also proves their union to be dense in $\mathbb{R}^n$. This follows because for every u and v in $\mathbb{R}^n$ either X is constant on the line segment joining u and v or because X varies a non-trivial contour must intersect the line segment. Either way the line segment contains part of a non-trivial contour. Consequently the union of non-trivial contours is dense in $\mathbb{R}^n$.

3.4 Analysis of the Brownian sheet.

Take W to be a Brownian sheet as defined in 3.1 . As remarked on in that section, the symmetries and scaling properties of W differ considerably from those of the Lévy process. Thus although the strategy of the analysis will be that of section 3.2 - to prove triviality of a contour by enclosing it in a curve on which W is larger or smaller than the value corresponding to the contour - the tactics must be rather different.

As with the Lévy process it is convenient to give a list of the symmetry and scaling properties that will be used.

(i) The process is invariant under permutation of the coordinates. That is, if π is a permutation of $\{1,\ldots,n\}$ then the process

$$t = (t_1,\ldots,t_n) \longmapsto W(t_{\pi 1},\ldots,t_{\pi n})$$

is a Brownian sheet;

(ii) for any r in $(0,\infty)$ the scaled process

$$t \longmapsto W(r.t_1,t_2,\ldots,t_n).r^{-\frac{1}{2}}$$

is a Brownian sheet;

(iii) the process produced by inversion

$$t \longmapsto \begin{cases} 0 & \text{if } t_1 = 0 \ , \\ t_1.W(t_1^{-1},t_2,\ldots,t_n) & \text{otherwise} \ , \end{cases}$$

is a Brownian sheet.

Proofs of these properties follow from computation of the covariance structures, just as in the case of the Levy process. Again the only complication is in dealing with the inverted process of (iii) and this can be dealt with as before.

Note that combinations of these invariance properties provide for such results as

$$t \longmapsto W(r.t_1, r.t_2, \ldots, r.t_n).r^{-n/2}$$

being a Brownian sheet for r in $(0,\infty)$.

An orthogonally scattered Gaussian random measure was used by Chentsov to construct the Brownian sheet in Chentsov (1956) and this means of presenting W will be of use in the proofs of the theorems of this section.(*) If the random measure η is defined as taking values in the Hilbert space $L^2(\mathcal{N}, \mathcal{F}, P)$ and such that

$$\eta\{ s : 0 \leq s_i \leq t_i \text{ for all } i \} = W(t)$$

for all $t = (t_1, \ldots, t_n)$ in $[0,\infty)^n$ then η takes values that are independent normal variables on disjoint open sets.

In order to investigate the level-sets of W definitions

(*) R.Pyke suggested the use of this measure in a personal communication. Without it it is very hard to describe the proof in the case of dimension higher than two.

are made as for the Lévy process but with a slight alteration to accomodate for the existence of a boundary. As before, the level-set is defined by

$$L(t) = \{ s \text{ in } [0,\infty)^n \text{ with } W(s) = W(t) \}.$$

However $C(t)$ is defined differently to make the case where t belongs to the boundary nontrivial. The contour through t is $C(t)$ the connected component of

$$(L(t) \setminus \partial\mathbb{T}) \cup \{t\}$$

that contains t, where $\mathbb{T} = [0,\infty)^n$ and $\partial\mathbb{T}$ is the boundary of $\mathbb{T}$, the set $\{ s : s_i = 0 \text{ for at least one } i \}$.

The results concerning $C(t)$ are not as complete as the corresponding results for the Lévy process. The behaviour of the set is not easy to establish when t belongs to a "corner" of $\partial\mathbb{T}$, that is when more than one of the t_i are zero. However when t is in $\partial\mathbb{T}$ but not in a "corner" the following result can be obtained;

Theorem

If precisely one of the coordinates of t is zero then there is an open neighborhood U of t in the boundary $\partial\mathbb{T}$ such that with probability one for all s in U the contour $C(s) = \{s\}$.

Proof

Because of property (i) there is no loss of generality in taking $t_1 = 0$, $t_2 > 0$, ... , $t_n > 0$.

With probability one

$W(1,t_2,\ldots,t_n) \neq 0$ since its variance is $t_2 \times \ldots \times t_n \neq 0$.

Consequently by the continuity of W there is a positive number p and an open neighborhood U of t in ∂T such that

$$P\left\{ W(1,s_2,\ldots,s_n) \neq 0 \text{ whenever } (0,s_2,\ldots,s_n) \in U \right\} \geq p \ .$$

By property (ii) the scaled process

$$s \longmapsto W(\lambda s_1, s_2, \ldots, s_n) \cdot \lambda^{-1/2}$$

has the same law as W and so if $\lambda \neq 0$ then

$$P\left\{ W(\lambda,s_2,\ldots,s_n) \neq 0 \text{ whenever } (0,s_2,\ldots,s_n) \in U \right\} \geq p \ .$$

By Fatou's lemma the event

$$\lim_{k\to\infty} \bigcup_{m=k}^{\infty} \left\{ W(1/m,s_2,\ldots,s_n) \neq 0 \text{ whenever } (0,s_2,\ldots,s_n) \in U \right\}$$

must have probability greater than p .

The last-mentioned event belongs to the σ-field

$$\bigcap_{\lambda > 0} \sigma\left\{ W(r,s_2,\ldots,s_n) \ : \ r < \lambda \text{ and } s_2,\ldots,s_n \geq 0 \right\}$$

and this "tail σ-field" appears a suitable candidate for a 0-1 law. Orey and Pruitt have proved a 0-1 law for such σ-fields, the proof combining an application of Kolmogorov's 0-1 law with an

exploitation of the random measure representation of W in terms of η (Orey and Pruitt (1973)). Consequently

$$P\left(\lim_{k\to\infty} \bigcup_{m=k}^{\infty} \left\{W(1/m, s_2, \ldots, s_n) \neq 0 \text{ whenever } (0, s_2, \ldots, s_n) \in U\right\}\right)$$

$$= 1 \quad .$$

This completes the proof that $\{s\}$ is a connected component of $(L(s)\setminus\partial\mathbb{I})\cup\{s\}$ for any s in U , and hence proves the theorem.

□

Corollary

No nontrivial contours intersect the "non-corner" portion of the boundary, the set $\left\{t \in \partial\mathbb{I} : \text{precisely one of the } t_i \text{ is zero .}\right\}$.

Proof

The "non-corner" portion of the boundary can be covered by open neighborhoods like U in the theorem above and a countable subcovering can be selected. With probability one the contours of points belonging to any one of this countable number of neighborhoods are trivial.

□

To establish a similar result for t in the interior of $\mathbb{T}$ a 0-1 law is required for events determined by the behaviour of W arbitrarily close to t. Although the proof is immediate from the paper of Orey and Pruitt mentioned above, a proof will be given here for the sake of completeness.

Lemma

For t in $\mathbb{T}$ the probability law P is degenerate on the "germ σ-field"

$$\bigcap_{\lambda>0} \sigma\left\{ W(s) - W(t) \ : \ \sup_i |s_i - t_i| < \lambda \right\} .$$

Proof

Let A_m be the subset of $\mathbb{T}$ defined by

$$A_m = \left\{ s \ : \ \begin{array}{c} s_i \leqslant t_i + 1/m \ \text{ for all } \ i \\ \text{and for at least one } \ j \\ s_j > t_j - 1/m \end{array} \right\} .$$

If the σ-field $\mathcal{F}_m$ is defined by

$$\mathcal{F}_m = \sigma\{ \eta(G) \ \text{ for } \ G \ \text{ a Borel set contained in } \ A_m \ \text{ but disjoint from } \ A_{m+1} \}$$

then clearly the $\mathcal{F}_m$ are independent, since η is an orthogonally scattered Gaussian measure. The germ σ-field of the statement of the lemma is contained in the σ-field $\bigvee_{k\geqslant m} \mathcal{F}_k$ generated by $\mathcal{F}_m$, $\mathcal{F}_{m+1}, \ldots$ for any m. This is because

$$W(s) - W(t) = \eta\left\{u : u_i \leq s_i \text{ for all } i \text{ and for at least one } j \; u_j \geqslant t_j\right\} - \eta\left\{u : u_i \leq t_i \text{ for all } i \text{ and for at least one } j \; u_j \geqslant s_j\right\} = \eta(F_1) - \eta(F_2)$$

which clearly belongs to the σ-field

$$\bigvee_{k \geqslant m} \mathcal{F}_m$$

when $\sup_i |s_i - t_i| < 1/m$. The attached figure (p 125)

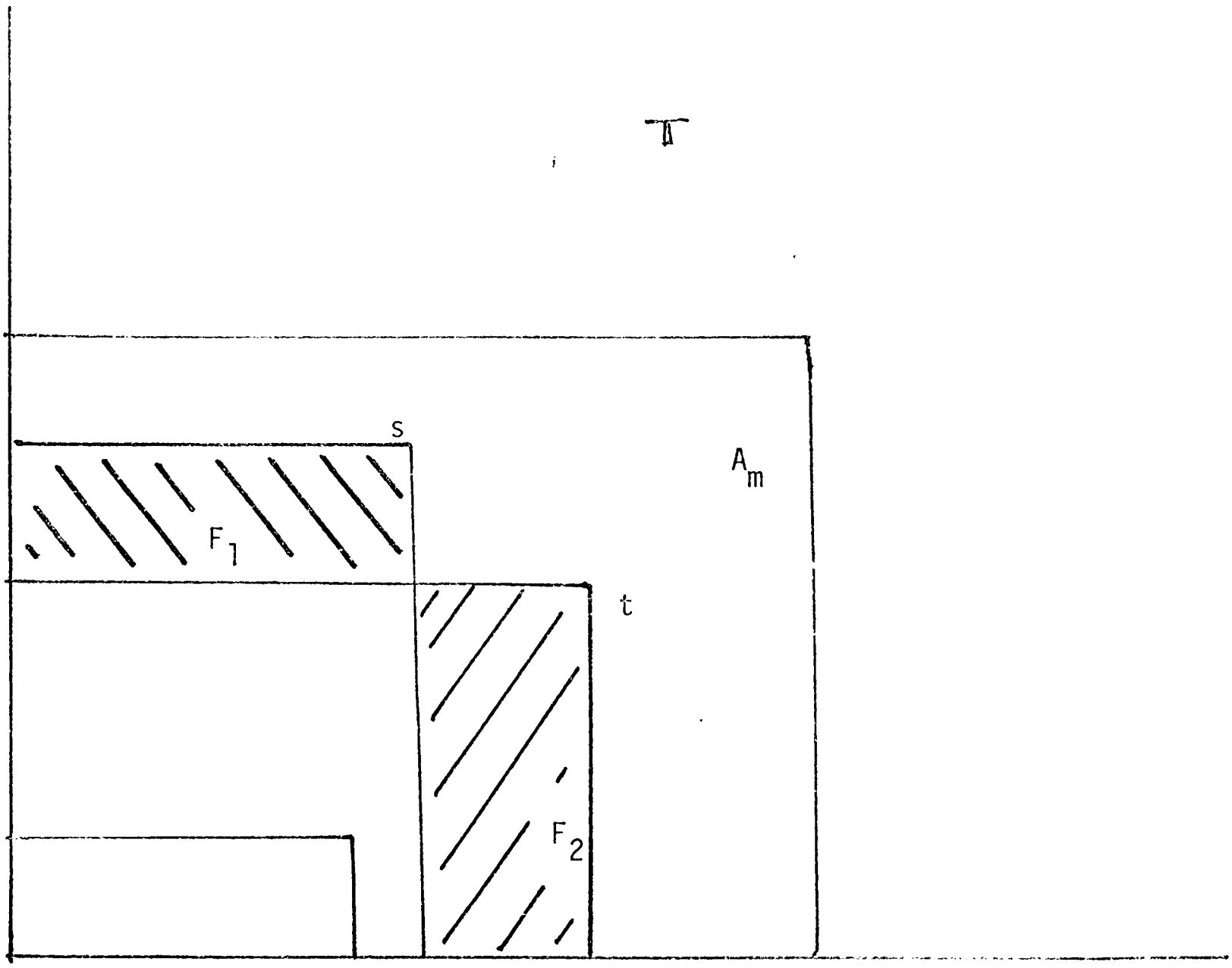
illustrates these remarks.

This argument shows that the germ σ-field is contained in $\bigvee_{k \geqslant m} \mathcal{F}_m$ for all m . Because the $\mathcal{F}_m$ are independent an application of Kolmogorov's 0-1 law completes the proof of the lemma. ÷

□

The 0-1 law mentioned in the proof of the previous theorem can be proved in a similar manner.

The machinery to deal with the contour at a point in the interior of $\mathbb{I}$ is now assembled.



$$W(s) - W(t) = \eta(F_1) - \eta(F_2)$$

Figure to be attached after page 124 section 3.4 .

This illustrates the proof of the lemma. The shaded regions correspond to the sets F_1 and F_2 . The region marked out in plain is the region A_m .

For obvious reasons this figure only deals with the case $n = 2$.

Theorem

Suppose t is in $\mathbb{T}\setminus\partial\mathbb{T}$, so that the coordinates of t are all positive. Then with probability one the contour $C(t)$ is trivial. That is to say,

$$C(t) = \{t\} \quad \text{with probability one.}$$

Proof

By the scaling property (ii) there is no loss in generality in taking $t = (1,1,\ldots,1) = \underline{1}$. The proof consists in establishing a lower bound on the probability of $W(s) - W(\underline{1})$ being positive on a box containing $\underline{1}$ and then using Fatou's lemma and the 0-1 law proved above.

Suppose $h \in (0,1)$ then we consider

$$\{W(\underline{1}+hs) - W(\underline{1}) \quad : \quad \sup_i |s_i| = 1\} \quad \text{which is to do with}$$

the process W restricted to the surface of a box of sides $2h$ centred on $\underline{1}$. The representation of W in terms of the random measure η is of great assistance for by it the process W can be decomposed;

$$\text{for } u \text{ in } \mathbb{I} \text{ such that } u_i \geq 1-h \text{ for all } i,$$

$$W(u) = W((1-h)\underline{1}) + \eta\{v : v_i \leq u_i \text{ for all } i \text{ and } v_j > 1-h \text{ for at least one } j\},$$

$$\text{So} \quad W(u) = W((1-h)\underline{1}) + \sum_{k=1}^{n} B_k(u_k - 1 + h)$$

$$+ \sum_{m=2}^{n} \sum_{\substack{\text{distinct} \\ j_1 < j_2 < \ldots < j_m \\ \text{in } \{1,\ldots,n\}}} W_{j_1,\ldots,j_m}(u_{j_1} - 1 + h, \ldots, u_{j_m} - 1 + h)$$

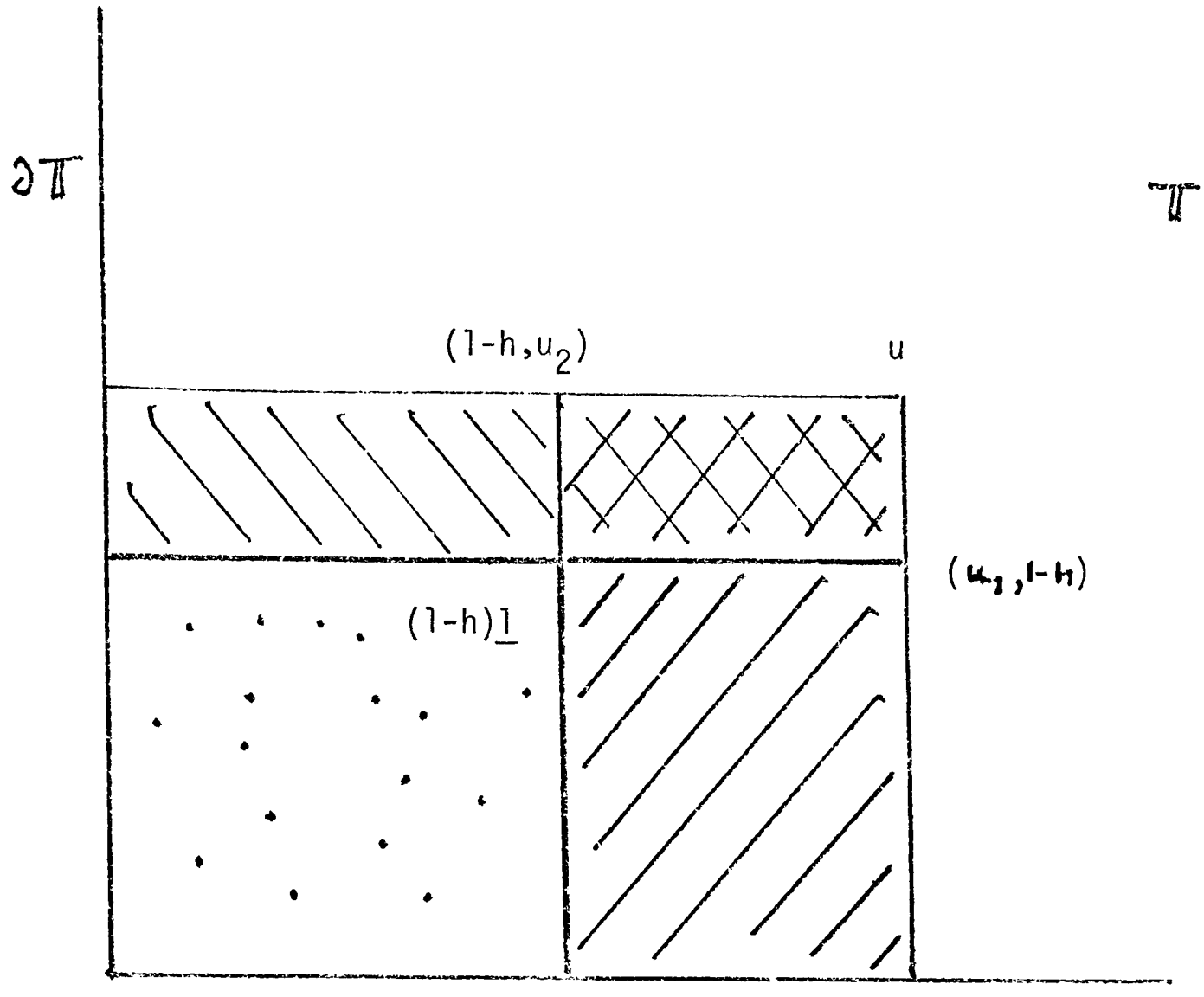
where the various B_k and $W_{j_1,\ldots,j_m}$ are Brownian motions and Brownian sheets that can be described in terms of η as follows;

$$B_k(r) = \eta\{v : v_k \leq r+1-h,\ v_k > 1-h \text{ and } v_q \leq 1-h \text{ for all } q \neq k\};$$

$$W_{j_1,\ldots,j_m}(r_1,\ldots,r_m)$$

$$= \eta\left\{v : v_{j_k} \leq r_k+1-h,\ v_{j_k} > 1-h \text{ for } k = 1,\ldots,m \text{ and } v_q \leq 1-h \text{ for all } q \neq j_1,\ldots,j_m\right\}.$$

To check this assertion is a somewhat detailed activity. Another figure is attached to help the reader visualise the decomposition in the case $n = 2$. It is important to note that the Brownian motions and Brownian sheets are all independent. This follows because they are defined using the values of η on disjoint sets. Moreover they are all independent of $W((1-h)\underline{1})$, for the same reason.



$$W(u) = W((1-h)\underline{1}) + B_1(u_1-1+h) + B_2(u_2-1+h) + W_{1,2}(u_1-1+h,u_2-1+h)$$

Figure to be attached after page 127 , section 3.4 .

This illustrates the decomposition of W in the second theorem .

Hatching /// for the area for B_1 ;

\\\ for the area for B_2 ;

⊠ for the area for $W_{1,2}$;

∴ for the area for $W((1-h)\underline{1})$.

The next step is to employ the so-called "Forgery Theorem" of Lévy (proved in Freedman (1971) p30, 1.3 Thm 38) and the scaling property (ii) to exploit this decomposition. By these two results there is a $p > 0$ such that

$$P\left\{\begin{array}{l} B_k(r) \leqslant h^{\frac{1}{2}} \quad \text{for} \quad r \in (0,2h) \\ B_k(0) \text{ and } B_k(2h) \leqslant h^{\frac{1}{2}}/4 \quad , \quad B_k(h) \geqslant (1-1/(4n)).h^{\frac{1}{2}} \end{array}\right\}$$
$$= \quad p$$

for $k = 1,\ldots,n$ and for h in $(0,1)$.

By the scaling property and by the continuity of the Brownian sheet there is an $M > 0$ and $p' > 0$ such that for any increasing sequence $j_1,\ldots,j_m$ taken from $1,\ldots,n$

$$P\left\{ |W_{j_1,\ldots,j_m}(r_1,\ldots,r_m)| \leqslant M.h^{m/2} \text{ when } r_k \in (0,h) \text{ for all } k \right\}$$
$$\geqslant \quad p' \quad .$$

Because all the processes concerned are independent it follows that the probability of all these bounds being valid at once is

$$p^n . \sum_{m=2}^{n} \binom{n}{m} (p')^m$$

a positive constant not depending on h . However if all the bounds are valid at once then one may deduce the following;

from the bounds on the Brownian sheets

$$\left| \sum_{m=2}^{n} \sum_{\substack{\text{distinct} \\ j_1 < j_2 < \ldots < j_m \\ \text{in } 1,\ldots,n}} W_{j_1,\ldots,j_m}(u_{j_1} - 1 + h, \ldots, u_{j_m} - 1 + h) \right|$$

$$\lesssim \sum_{m=2}^{n} \binom{n}{m} M.h^{m/2}$$

$$= M.((1+h^{\frac{1}{2}})^n - 1 - n.h^{\frac{1}{2}}) = O(h)$$

so that this part of the decomposition of W will be negligible relative to the rest when h is small;

from the restrictions on the Brownian motions

if $\sup_k |u_k - 1| = h$ then

$$\sum_{k=1}^{n} B_k(u_k - 1 + h) \lesssim (n-1)h^{\frac{1}{2}} + h^{\frac{1}{2}}/4 = (n-3/4)h^{\frac{1}{2}}$$

if $u_k - 1 + h = h$ for all k so that $u = \underline{1}$ then

$$\sum_{k=1}^{n} B_k(u_k - 1 + h) \gtrsim n.(1-1/(4n))h^{\frac{1}{2}} = (n-1/4)h^{\frac{1}{2}} .$$

These inequalities mean that for sufficiently small h, so that the $O(h)$ bound on the Brownian sheets is smaller than $h^{\frac{1}{2}}/4$,

there are inequalities

$$\sup\left\{ W(hs + \underline{1}) - W((1-h)\underline{1}) \;:\; \sup_i |s_i| = 1 \right\}$$

$$< \quad (n-\tfrac{1}{2}).h^{\frac{1}{2}} \;,$$

$$W(\underline{1}) \;-\; W((1-h)\underline{1}) \quad > \quad (n-\tfrac{1}{2}).h^{\frac{1}{2}} \quad ,$$

holding with probability greater than

$$p^n . \sum_{m=2}^{n} \binom{n}{m} (p')^m \quad > \quad 0 \quad .$$

But if these inequalities hold then the contour $C(\underline{1})$ must be confined within the box

$$\left\{ t \;:\; \sup_i |t_i - 1| = h \right\} \quad .$$

By Fatou's lemma the probability of such pairs of inequalities holding for arbitrarily small h must be greater than 0 . So by the 0-1 law of the lemma it must be almost certain that such pairs of inequalities hold for arbitrarily small h . Consequently the contour $C(\underline{1})$ is contained within arbitrarily small boxes and so must be trivial with probability one. So the theorem is proven .

□

Note that the proof will only work for h becoming very small. If h is considered close to 1, which is as large as it can be, then the dominant term in the expansion of W becomes the last term of the last summation and this itself is a Brownian sheet of n-dimensional time. So the decomposition of W is not as illuminating in that case.

Of course the arguments of section 3.3 still apply to the Brownian sheet. Although as a corollary to the above theorem it can be shown that the union of all nontrivial contours of W must have zero Lebesgue measure in $\mathbb{T}$, nontrivial contours must exist and indeed will be in union dense in $\mathbb{T}$.

In contrast to the Levy process several questions remain unsettled about contours of the Brownian sheet. These are;

(i) are there unbounded contours ?

(ii) is it possible for a nontrivial contour to intersect $\{t \in \partial\mathbb{T} : \text{two or more of } t_i \text{ are zero}\}$?

(iii) is it possible for an unbounded contour to intersect $\{t \in \partial\mathbb{T} : \text{two or more of } t_i \text{ are zero}\}$?

Using the inversion property of the Brownian sheet it is easy to establish that an affirmative answer to (ii) implies an affirmative answer to (i). Since an affirmative answer

to (iii) quite obviously implies an affirmative answer to (ii) the questions are in a natural order. However they all seem to be rather hard.

3.5 Conclusion.

The results of this chapter have an interesting bearing on the simulations of the Brownian sheet by Adler (1978a) and others. These workers have obtained simulations of the sheet for the case $n = 2$, $T = [0,\infty)^2$ and in particular simulations of the zero level-set. Naturally the pathological structure shown to exist above cannot be directly observed. However in Adler's simulation three drawings of the level-set near zero are displayed, at progressively greater magnification. The reader may care to refer to these simulations and to see if they help him to imagine that mathematical idealisation, the level-set of the Brownian sheet together with all its trivial contours.

References for the introduction.

Edwards,S.F. (1968) "Statistical Mechanics with Topological Constraints II " J.Phys.A Gen.Phys. 1 pp 15-28 .

Kendall,D.G and W.S.Kendall (1979) "Alignments in two-dimensional random sets of points." To appear.

Kendall,W.S (1979) "The knotting of Brownian motion in 3-space." Journal of the L.M.S., forthcoming.

Milnor,J (1964) "Most knots are wild." Fund.Math. 54 pp 335 - 338 .

Williams,D. (1979) "Diffusions, Markov processes, and Martingales. Volume I. Foundations." Wiley and Sons, London.

References for chapter one.

Behrend,D.M. (1978)

Broadbent,S.R. (1979) "Simulating the ley hunter."

Edwards,R. (1978) "Normal triangles."
unpublished.

Gečiauskas,E. (1971) "The distribution of the size of an angle of a random triangle inside an oval,"
Litovsk. Mat. Sb. XI 2 p447 .

Gečiauskas,E. (1972) "The distribution of the size of an angle of a triangle in case of random side."
Litovsk. Mat. Sb. XII 2 pp 193 - 94 .

Heaton,N. (= N.Lavery) (1976) "Ley lines."
University of London M.Sc. dissertation.

Kendall,D.G. (1979) "Shape-manifolds, procrustean metrics, and complex projective spaces."

References for chapter one, continued.

Kendall,D.G. and W.S.Kendall (1979) "Alignments in two-dimensional random sets of points."

Silverman,B.W. and T.C.Brown (1978) "Short distances, flat triangles and Poisson limits."
J. Appl. Prob. 15 pp 815 - 825 .

Stewart,P.A.A.,D.Wylie,P.Bestley and P.D.Hodgson (1978) "A search for ley lines in Cumbria."

Sukiasyai,G.S. (1978) "On random triangles in the plane."
Dokl. Akad. Nauk Armen. S.S.R. LXVI 3 pp 150-15

References to chapter two.

Crowell,R.H. and R.H.Fox (1963) "Introduction to knot theory."
Ginn and Co. , New York.

Deam,R.T. and S.F.Edwards (1976) "The theory of Rubber Elasticity."
Phil. Trans. Roy. Soc. (A) 280 pp 317 - 353.

Dvoretsky,A.,P.Erdös and S.Kakutani (1950) "Double points of
Brownian motion in n-space."
Acta Sci, Mat. (Szeged.) 12 B pp 75 - 81 .

Edwards,S.F. (1968) "Statistical Mechanics with Topological
Constraints, II ."
J.Phys. A Gen. Phys. 1 pp 15 - 28 .

Fox,R.H. (1962) "A quick trip through knot theory." pp 120 - 167
in "Topology of 3-manifolds." (Proc. Top. Inst.
of Univ. of Georgia, 1961) Prentice-Hall, New Jersey.

Freedman,D. (1971) "Brownian motion and Diffusion."
Holden-Day, San Francisco.

Fristedt,Bert. (1967) "An extension of a theorem of S.J.Taylor concerning
the multiple points of the symmetric stable process."
Z. Wahrscheinlichkeitstheorie, 9 pp 62 - 64 .

References for chapter two, continued.

Hempel,J.P. and D.R.McMillan Jr. (1966) "Locally nice embeddings of manifolds."
Am. J. of Maths. 88 pp 1 - 19 .

Itô,K. and H.P.McKean,Jr. (1965) "Diffusion Processes and their Sample Paths."
Springer Verlag , Berlin.

Milnor,J. (1964) "Most knots are wild."
Fund. Math. 54 pp 335 - 338

References for chapter three.

Adler,R.J. (1978a) "Some erratic patterns generated by the planar Wiener process."
Supp. Adv. Appl. Prob. 10 pp 22 - 27 .

Adler,R.J. (1978b) "The uniform dimension of the level sets of a Brownian sheet."
Ann. Prob. 6 pp 509 - 515 .

Arveson,W. (1976) "An invitation to C*-algebra."
Springer, New York.

Chentsov,N.N. (1956) "Wiener random fields depending on several parameters."
Dokl. Akad. Nauk. S.S.S.R. N.S. 106 pp 607 - 609 .

Dieudonné,J. (1969) "Foundations of Modern Analysis."
Academic Press, New York.

Freedman,D. (1971) "Brownian motion and diffusion."
Holden-Day,

Greenberg,M.J. (1967) "Lectures on Algebraic Topology."
Benjamin, New York.

References for chapter three, continued.

Kolmogorov,A.N. and Y^{u}.A.Rozanov (1960) "On strong mixing conditions for stationary Gaussian processes."
Theory of Prob. and its Applic. V pp 204 - 208 .

Lévy,P. (1965) "Processus stochastiques et mouvement Brownien."
Gauthiers-Villars, Paris.

Mandelbrot,B.B. (1977) "FRACTALS form, chance and dimension."
Freeman, San Francisco.

McKean,H.P.Jr. (1963) "Brownian motion with a several-dimensional time."
Theory of Prob. and its Applic. VIII pp 335 - 354.

Molchan,G.M. (1967) "On some problems concerning Brownian motion in Lévy's sense."
Theory of Prob. and its Applic. XII pp 682 - 690 .

Orey,S and W.E.Pruitt (1973) "Sample functions of the N-parameter Wiener process."
Ann. Prob. 1 pp 138 - 163 .

Pyke,R. (1973) "Partial sums of matrix arrays, and Brownian sheets.
pp 331 - 348 of "Stochastic Analysis", edited by D.G.Kendall and E.F.Harding, publ. Wiley and Sons, London.

References for chapter three,continued.

Takenaka,S. (1977) "On projective invariance of multi-parameter Brownian motion."
Nagoya Math. Journal Vol 67 pp 89 - 120 .