The Jet-Set Travelling Salesman (I): Triangular Tours.

by

Antonia J. Jones and Graham Benyon-Tinker

Abstract. This paper is the first in a series concerned with the distribution of tour lengths on the surface of a sphere (hence the title), and treats the simplest case, where only $n = 3$ cities are visited. As a byproduct of the investigation we obtain a formula for the area of a 'spherical' ellipse and a curious pseudo-elliptic integral.

Keywords: Geometrical probability, spherical triangle, elliptic integral, pseudo-elliptic integral.

DEPARTMENT OF COMPUTING
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE
UNIVERSITY OF LONDON
180 Queen's Gate, London, SW7 2BZ

Telephone: 071-589 5111 Ext: 5017/5096
Telefax: 071-581 8024
E-Mail: ajj@doc.ic.ac.uk

Date/version: 14 October 2000

The Jet-Set Travelling Salesman (I): Triangular Tours.

CONTENTS

List of figures

The Jet-Set Travelling Salesman (I): Triangular Tours.

Introduction.

The best *exact solution* methods for the Travelling Salesman Problem (TSP) are capable of solving problems of several hundred cities [Grötschel 1991], but unfortunately excessive amounts of computer time are used in the process and, as the number of cities n increases, any exact solution method rapidly becomes impractical. In this paper we are interested in the geometric TSP in which the cities are confined to a closed bounded region and the distances are symmetric and satisfy the triangle inequality. For large problems we have no way of knowing the exact solution, but in order to gauge the solution quality of any algorithm we need a reasonably accurate estimate of the minimal tour length. This is usually provided in one of two ways.

For a uniform distribution of cities the classic work by Beardwood, Halton and Hammersley (BHH) [Beardwood 1959] obtains an asymptotic best possible upper bound for the minimum tour length for large n. Let $\{X_i\}$, $1 \le i < \infty$, be independent random variables uniformly distributed over the unit square, and let L_n denote the shortest closed path which connects all the elements of $\{X_1,...,X_n\}$. In the case of the unit square they proved, for example, that there is a constant $c > 0$ such that, with probability 1,

$$\lim_{n \to \infty} L_n \, n^{-1/2} = c \qquad (1)$$

where $c > 0$ is a constant. In general c depends on the geometry of the region considered.

One can use the estimate provided by the BHH theorem in the following form: the expected length L_n^* of a minimal tour for an n-city problem, in which the cities are uniformly distributed in a square region of the Euclidean plane, is given by

$$L_n^* \approx 0.765\sqrt{nR} \qquad (2)$$

where R is the area of the square and the constant 0.765 has been determined empirically [Stein 1977].

A second possibility would be to use a problem specific estimate of the minimal tour length which gives a very accurate estimate: the *Held-Karp lower bound* [Held 1970], [Held 1971]. This involves finding the Minimal Spanning Tree for n-1 cities of the TSP and a *Lagrangean relaxation*; altogether at least a $O(n^2)$ operation, although there are intelligent short cuts.

In addition to an accurate estimate for the minimal tour length it would also be useful, in evaluating an heuristic algorithm, to have either an exact formula for, or some efficient method of calculating, the general distribution of tour lengths (not necessarily minimal) as a function of the number of points. It is this question which has prompted the present work. Specifically, we have considered the following problem:

> Let n points be independently selected from a uniform distribution within a bounded region and joined in the order of selection, with the last point being connected to the first, thus giving a closed tour of length r. Over the population of all possible such constructions what is the distribution $T_n(r)$ of r?

Whilst $T_n(r)$ is not the same as the (discrete) distribution of the ½(n-1)! possible tour lengths for a *fixed* set of randomly placed points, it gives considerable insight into the averaged behaviour of the discrete distribution over *all possible* problems, which is precisely what is of interest in algorithm design (specifically the lower tail of T_n). In particular, our research suggests that for large n, T_n differs quite substantially from a Normal distribution in this region.

Conventionally the geometric TSP has been studied with reference to points uniformly distributed in the unit square. A difficulty with this approach lies in the nature of the probability density function for the distance between any two arbitrarily chosen points.

Theorem 1. The probability density function for the distance r between two independent uniformly distributed points in the unit square is given by

$$p(r) = \begin{cases} 2\pi r - 8r^2 + 2r^3 & (0 \le r \le 1) \\ 4r\left(\sin^{-1}\frac{1}{r} - \cos^{-1}\frac{1}{r} + 2\sqrt{r^2 - 1} - \frac{1}{2}r^2 - 1\right) & (1 \le r \le \sqrt{2}) \end{cases} \tag{3}$$

and is zero ouside [0, $\sqrt{2}$].

Corollary. This distribution has mean τ and standard deviation σ given by

$$\tau = \frac{1}{15}(2+\sqrt{2}) + \frac{1}{3}\log(1+\sqrt{2}) = 0.521\ 405\ 433$$
$$\sigma = \sqrt{\frac{1}{3} - \tau^2} = 0.247\ 930\ 852 \tag{4}$$

It would appear that this result was first proved in [Borel 1925] who also obtained the corresponding result for triangles, polygons in general, and the circle. The distribution for the square is also discussed in [Gečiauskas 1966] who deduced a formulation for the distribution of the distance between two points independently distributed uniformly over a convex region symmetrical about a fixed point (he calls this an ’oval’), and hence the result for a square as a special case. These results can also be obtained using Crofton's theorem, see [Kendall 1963].

The division of the probability density function into two distinct forms makes it awkward for use in any analysis of the distribution of the sum of n edges, derived by convolution: this convolution is not, of course, T_n but is closely related to it. The asymmetries of the distance distribution are related to the imposition of a boundary around the region of investigation and will occur for any planar bounded region, e.g. the unit circle.

- These technical difficulties can be significantly reduced by considering the distribution of tour lengths over points scattered uniformly on the surface of a *sphere*.

We choose a sphere because it is the simplest closed bounded region having no boundary but other surfaces such as an ellipsoid or torus would also be suitable. We denote by $d(r)$ the probability density function for the spherical distance r between two points. We shall shortly see that $d(r)$ has an extremely simple form, and for this reason we were able to make analytic progress. For obvious reasons we have named this the *Jet-Set* TSP.

Our initial aim was to obtain the distribution of the n-fold convolution of $d(r)$, which we shall refer to as $\sigma_n(r)$. Of course, by the Central Limit Theorem $\sigma_n(r)$ is asymptotic to a Normal distribution. However, this observation is not so helpful, because it is precisely the tail of the distribution which is of interest and there the approximation to Normality turns out to be rather inaccurate. The function $\sigma_n(r)$ is essentially the probability density function of the length of an *open* tour on n+1 points joined up in random order. We have obtained an exact formula for $\sigma_n(r)$, *viz.*

Theorem 2. For $n \geq 3$ the probability density function $\sigma_n(r)$ for the surface of a sphere of radius $1/\pi$ is

$$\left(\frac{1}{2}\right)^n \sum_{j \leq r} \binom{n}{j} \left(\sum_{k=1}^{n} M_k (-1)^k \, i^{\,k} \left(e^{\,i\pi(r-j)} + (-1)^k e^{\,-i\pi(r-j)} \right) (r-j)^{k-1} \right) \tag{5}$$

for $0 \leq r \leq n$, where

$$M_k = \frac{(2n - k - 1)! \; \pi^k}{(n - 1)!(k - 1)!(n - k)! \; 2^{2n-k}} \tag{6}$$

This distribution has mean $n/2$ and variance

$$\sigma_n^{\,2} = n\left(\frac{1}{4} - \frac{2}{\pi^2}\right) = 0.047\ 357\ 632\ n \tag{7}$$

The derivation of this result will be left for a future paper. We believe (although we have not yet proved) that not only is $T_n(r)$ asymptotic to $\sigma_n(r)$ as n -> ∞ but that the difference is of sufficiently small order to offer an excellent approximation to $T_n(r)$.[1] Certainly, experimental results indicate that the two functions are indistinguishable for $n > 10$.

However, when n is small, there is a systematic difference between $T_n(r)$ and $\sigma_n(r)$. This is most apparent in the case $n = 3$, which exhibits some interesting features, and the work of the present paper is directed towards the study of $T_3(r)$. This also serves to illustrate some of the key issues in the determination of T_n, our original goal. Of course when $n = 3$ there is essentially only one tour (and its reverse), so in this case knowing the distribution of tour lengths also tells us the distribution of *minimal* tour lengths, which for general n is unknown.

Spherical triangles.

Working on the surface of a sphere the geodesic is the shortest great circle distance and we shall use α, β, γ to denote the three sides of a spherical triangle. In this paper the sphere is normalised to have radius 1.

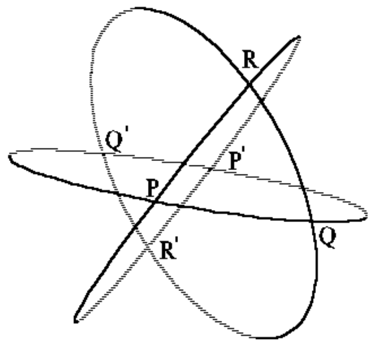

Figure 1 Spherical triangle PQR and its dual.

We have remarked that, due to the lack of boundaries, working on the surface of a sphere confers technical benefits. However, it also introduces some undesired complications, especially when n is small. For example, suppose we fix one side of the triangle so that $\gamma = \pi$, then *wherever* the third point is placed we find that $\alpha + \beta + \gamma = 2\pi$. Not suprisingly it turns out that the probability density function $T_3(r)$ for $r = \alpha + \beta + \gamma$ has a singularity at 2π. These effects do not occur, for example, on the unit square; given one side of a triangle, for any fixed total side length, the possible third points are restricted to the intersection of an ellipse with the interior of the closed square, consequently the set of such third points has zero measure with respect to the square.

There are very beautiful symmetries associated with spherical triangles. Let PQR be a spherical triangle with sides α, β, γ and angles A, B, C. Then by extending each of the great circle arcs joining PQ, QR, RP and considering the three other points of intersection we can construct a dual triangle P´Q´R´ with corresponding sides and angles α', β', γ' and A´, B´, C´ respectively. In texts on spherical trigonometry it is shown that

$$
\begin{gathered}
A' = \pi - \alpha, \quad B' = \pi - \beta, \quad C' = \pi - \gamma \\
A = \pi - \alpha', \quad B = \pi - \beta', \quad C = \pi - \gamma' \\
\text{Area}\triangle(\text{PQR}) = A + B + C - \pi
\end{gathered}
\tag{8}
$$

[1] This, as it stands, is not as helpful as might at first appear because the form of the function in (5) turns out to be rather difficult to compute accurately for large n.

These symmetries yield the rather elegant observation that

- Knowing the distribution of the sum of the side lengths $\alpha + \beta + \gamma$ also tells us the distribution of the sum of the angles and of the area.

The cosine rule for the spherical triangle in Figure 2 on a sphere of radius 1, gives

$$\cos\alpha = \cos\gamma\,\cos\beta + \sin\beta\,\sin\gamma\,\cos\omega \tag{9}$$

If we introduce coordinates (r, ω), as in Figure 2, since the point R = (r, ω) is uniformly distributed over the surface of the sphere the joint probability of (r, ω) will be proportional to the element of surface area, that is, to $\sin r\, dr\, d\omega$. We can, without loss of generality, suppose that P and Q are on the equator and so the upper and lower hemispheres are identical. Therefore we can take the joint distribution of (r, ω) to be given by the frequency function

$$f(r,\ \omega)\ dr d\omega = \frac{\sin r}{2\pi}\ dr d\omega \qquad (0 \le r \le \pi,\ 0 \le \omega \le \pi) \tag{10}$$

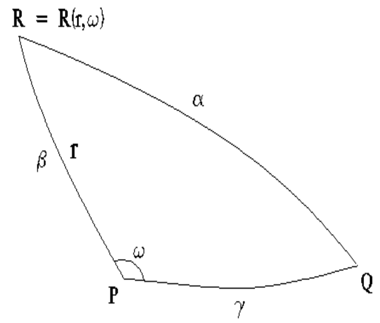

Figure 2 Viewed from above the surface of the sphere.

Integrating over ω leads immediately to the probability density function for r

$$d(r) = \begin{cases} \frac{1}{2}\sin r\ , & 0 \le r \le \pi \\ 0\ , & \text{otherwise} \end{cases} \tag{11}$$

which is analogous to (3).

For future reference we note that the 2-fold convolution of *d(r)* is given by

Theorem 3. Let $\sigma_2(r)$ be the two fold convolution of *d(r)* ($0 \le r \le 2\pi$) then

$$\sigma_2(r) = \begin{cases} \frac{\pi}{8}\left(\sin r - r\cos r\right) & (0 \le r \le \pi) \\ \frac{\pi}{8}\left(\sin(2\pi - r) - (2\pi - r)\cos(2\pi - r)\right) & (\pi \le r \le 2\pi) \end{cases} \tag{12}$$

Remark. That the integral has two functional forms arises from the fact that *d(r)* has compact support (vanishes outside a bounded region), and this has a fundamental impact on the analysis. The probability density function $\sigma_2(r)$ is continuous at $r = \pi$ as are its first and all *even* derivatives, but the third and all higher *odd* derivatives are discontinuous at $r = \pi$. In fact, this is generally true for $n \ge 2$.

Elliptic Integral notation.

The subsequent analysis will involve elliptic integrals. Since there are variations in the way these are defined and of notation, we give here the notation which we shall use and quote the important theorem of Legendre. We have used Byrd and Friedman [Byrd 1971] as a comprehensive reference to the properties of elliptic integrals.

Elliptic Integrals of the First kind.

$$F(\phi, m) \equiv \int_0^\phi \frac{d\phi}{\sqrt{1 - m\sin^2\phi}} \qquad 0 \le m \le 1 \tag{13}$$

Elliptic Integrals of the Second kind.

$$E(\phi, m) \equiv \int_0^\phi \sqrt{1 - m\sin^2\phi}\; d\phi \qquad 0 \le m \le 1 \tag{14}$$

We shall often write $m = k^2$, $k' = \sqrt{(1 - k^2)}$ and $m' = 1 - m$. It is also convenient to write

$$E_m \equiv E\left(\frac{\pi}{2}, m\right) \qquad F_m \equiv F\left(\frac{\pi}{2}, m\right) \tag{15}$$

Elliptic Integrals of the Third kind.

$$\Pi(\phi, p, m) \equiv \int_0^\phi \frac{d\phi}{(1 - p\sin^2\phi)\sqrt{1 - m\sin^2\phi}} \qquad 0 \le m \le p \le 1 \tag{16}$$

Legendre (1826) proved that a complete elliptic integral of the third kind can be expressed in terms of elliptic integrals of the first and second kind:

$$\Pi(\phi,\ p,\ m)\ =\ \frac{1}{Q(p,\ m)}\Big((E_m\ -\ F_m)F(\mu,\ m')\ +\ F_mE(\mu,\ m')\Big) \tag{17}$$

where

$$Q(p,\ m)\ =\ \sqrt{\frac{(1\ -\ p)(p\ -\ m)}{p}}$$

$$\mu\ =\ \sin^{-1}\sqrt{\frac{(p\ -\ m)}{pm'}} \tag{18}$$

Heuman's Lambda function [Heumann 1941] can then be defined from (17) as

$$\Lambda_0(\mu,\ m)\ \equiv\ \frac{2}{\pi}\Big((E_m\ -\ F_m)F(\mu,\ m')\ +\ F_mE(\mu,\ m')\Big)\ =\ \frac{2}{\pi}Q(p,\ m)\ \Pi(\tfrac{\pi}{2},\ p,\ m) \tag{19}$$

In the case $\mu = \pi/2$, $p = 1$, Legendre also proved the identity

$$E_mF_{m'}\ +\ F_mE_{m'}\ -\ F_mF_{m'}\ \equiv\ \frac{\pi}{2} \tag{20}$$

which implies that $\Lambda_0(\pi/2,\ m) = 1$ for all m. It can be readily verified that $\Lambda_0(\mu,\ \mathrm{m})$ is an odd function of μ and increases monotonically from 0 to 1 as μ varies from 0 to $\pi/2$.

The spherical ellipse.

In the Euclidean plane an ellipse is usually defined as the locus of points such that the sum of the distances from two fixed points, the *foci*, is constant. We define an analogous figure on the surface of a sphere which we call a *spherical ellipse* and which turns out to be useful in our analysis of T_3.

Given any two points placed at random on a sphere, we can always choose a system of coordinates where the great circle through the points defines an equatorial plane and the origin of longitude *O* lies halfway between the points. Place two focal points P and Q on the equator at longitudes $-\gamma/2$ and $\gamma/2$ respectively. Let R be any other point such that the sum of the arc lengths RP + RQ = λ, where $\lambda \geq \gamma$ is constant. The spherical ellipse is the locus of R in coordinates (ρ, θ), where ρ is the arc length *O*R and θ is the angle between the great circle through *O*R and the equatorial plane.

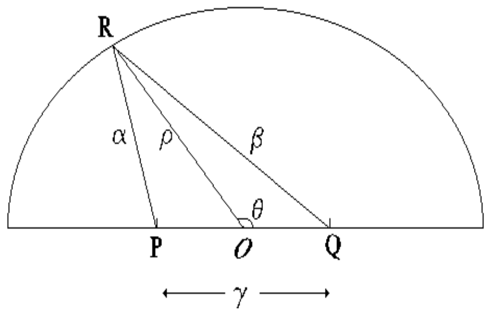

Figure 3 Coordinate system for the spherical ellipse.

Theorem 4. The spherical ellipse is defined by the equation

$$\cos^2\rho = \frac{\sin^2\frac{\lambda}{2}.\cos^2\frac{\lambda}{2} - \sin^2\frac{\gamma}{2}.\cos^2\frac{\lambda}{2}.\cos^2\theta}{\sin^2\frac{\lambda}{2}.\cos^2\frac{\gamma}{2} - \sin^2\frac{\gamma}{2}.\cos^2\frac{\lambda}{2}.\cos^2\theta} \tag{21}$$

Proof. Let the distances RP and RQ be α and β respectively. Then applying (9) to spherical triangles *O*RQ and *O*RP we have

$$\begin{aligned} \cos\alpha &= \cos\rho.\cos\frac{\gamma}{2} + \sin\rho.\sin\frac{\gamma}{2}.\cos\theta \\ \cos\beta &= \cos\rho.\cos\frac{\gamma}{2} - \sin\rho.\sin\frac{\gamma}{2}.\cos\theta \end{aligned} \tag{22}$$

On applying the condition $\cos(\alpha + \beta) = \cos\lambda$ the statement of the theorem follows after some straightforward, if rather tedious, manipulations.

Remark. Another interpretation for λ is that it is the arc length of the major axis. If we let μ be the arc length of the minor axis then from (21) with $\theta = \pi/2$ we have

$$\cos\frac{\mu}{2} = \frac{\cos\frac{\lambda}{2}}{\cos\frac{\gamma}{2}} \tag{23}$$

from which follows

$$\sin^2\frac{\lambda}{2} - \sin^2\frac{\mu}{2} = \sin^2\frac{\lambda}{2}.\frac{\tan^2\frac{\gamma}{2}}{\tan^2\frac{\lambda}{2}} \tag{24}$$

which is analogous to the relation $a^2 - b^2 = a^2e^2$ for the plane ellipse.

We next consider the problem of determining the area of the spherical ellipse. For this we need

Lemmma 1. Let

$$S(u,\ \mathrm{v}) \equiv \int_0^{\frac{\pi}{2}} \frac{\mathrm{v}}{\mathrm{u}} \sqrt{\frac{1 - u^2\cos^2\theta}{1 - \mathrm{v}^2\cos^2\theta}}\ d\theta \tag{25}$$

Then S is a complete elliptic integral of the third kind and

$$S(u,\ \mathrm{v}) = Q(p,\ m)\Pi(\frac{\pi}{2},\ p,\ m) = \frac{\pi}{2}\Lambda_0(\mu,\ m) \tag{26}$$

where Q(p, m) and μ are defined by (18) and

$$p = u^2$$
$$m = \frac{u^2 - \mathrm{v}^2}{1 - \mathrm{v}^2} \tag{27}$$

Proof. In (25) make the transformation from θ to φ given by

$$\tan\theta = \tan\phi\ \sqrt{1 - u^2} \tag{28}$$

Then the limits of integration remain unchanged and the integral transforms to

$$S(u,\ \mathrm{v}) = Q(p,\ m) \int_0^{\frac{\pi}{2}} \frac{d\phi}{(1 - p\ \sin^2\phi)\sqrt{1 - m\ \sin^2\phi}} \tag{29}$$

where p, m are defined in terms of u, v by (27).

Apart from the factor Q(p, m) the righthand side of (29) is in the normal form for a complete elliptic integral of the third kind. Now (26) follows from (19).

Theorem 5. Consider a spherical ellipse with major axis λ and focal separation γ, so that γ < λ < 2π - γ. Let G(λ, γ) be the ratio of the area of the ellipse to the area of the whole sphere. Then

$$G(\lambda,\ \gamma) = \frac{1}{2} - \frac{1}{\pi}\left((E_m - F_m)F(\frac{\nu}{2},\ m') + F_mE(\frac{\nu}{2},\ m')\right) = \frac{1}{2} - \frac{1}{2}\Lambda_0(\frac{\nu}{2},\ m) \tag{30}$$

where $\nu = \pi - \mu$, $m = \sin^2(\gamma/2)$, $m' = \cos^2(\gamma/2)$ and from (23)

$$\cos\frac{\lambda}{2} = \cos\frac{\mu}{2}.\cos\frac{\gamma}{2} = \sin\frac{\nu}{2}.\cos\frac{\gamma}{2} \tag{31}$$

Remark. As λ increases from γ to 2π - γ, the minor axis μ increases from 0 to 2π and ν/2 decreases from π/2 to -π/2.

Proof. In considering the rotation of the radius vector as point R traverses the perimeter of the spherical ellipse, the area swept out between θ and θ + dθ is just

$$dA = (1 - \cos\rho)d\theta \tag{32}$$

Hence

$$A = 4\int_0^{\frac{\pi}{2}}(1 - \cos\rho)d\theta \tag{33}$$

so that

$$G(\lambda, \gamma) = \frac{A}{4\pi} = \frac{1}{2} - \frac{1}{\pi}\int_0^{\frac{\pi}{2}}\cos\rho \; d\theta \tag{34}$$

From Theorem 4

$$\cos\rho = \frac{v}{u}\sqrt{\frac{1 - u^2\cos^2\theta}{1 - v^2\cos^2\theta}} \tag{35}$$

where

$$u = \frac{\sin\frac{\gamma}{2}}{\sin\frac{\lambda}{2}} \quad \text{and} \quad v = \frac{\tan\frac{\gamma}{2}}{\tan\frac{\lambda}{2}} \tag{36}$$

We see by Lemma 1 that the integral on the righthand side of (34) is now a complete elliptic integral of the third kind as given by (26). Substituting (17) into (26), and (36) into (27), yields the statement of the theorem.

The conditional probability density function for λ = α + β.

We are interested in the distribution of r = α + β + γ. One way to approach this is by asking: if the length of the first side PQ is γ what is the probability density function for λ = α + β conditional upon γ?

Theorem 6. For 0 < γ < π and γ < λ < 2π - γ the probability density function for λ conditional upon γ is given by

$$g(\lambda|\gamma) = \frac{1}{2\pi}\cdot\frac{E_m - F_m\cos^2\frac{\lambda}{2}}{\sqrt{\cos^2\frac{\gamma}{2} - \cos^2\frac{\lambda}{2}}} \tag{37}$$

where $m = \sin^2(\gamma/2)$.

Remark: The function is plainly symmetric about $\lambda = \pi$ for any fixed γ, $0 < \gamma < \pi$. Moreover it has singularities at both endpoints. Figure 4 illustrates g as a surface plot. The poles at the endpoints of the λ range are more easily seen in the logarithmic plot of Figure 5.

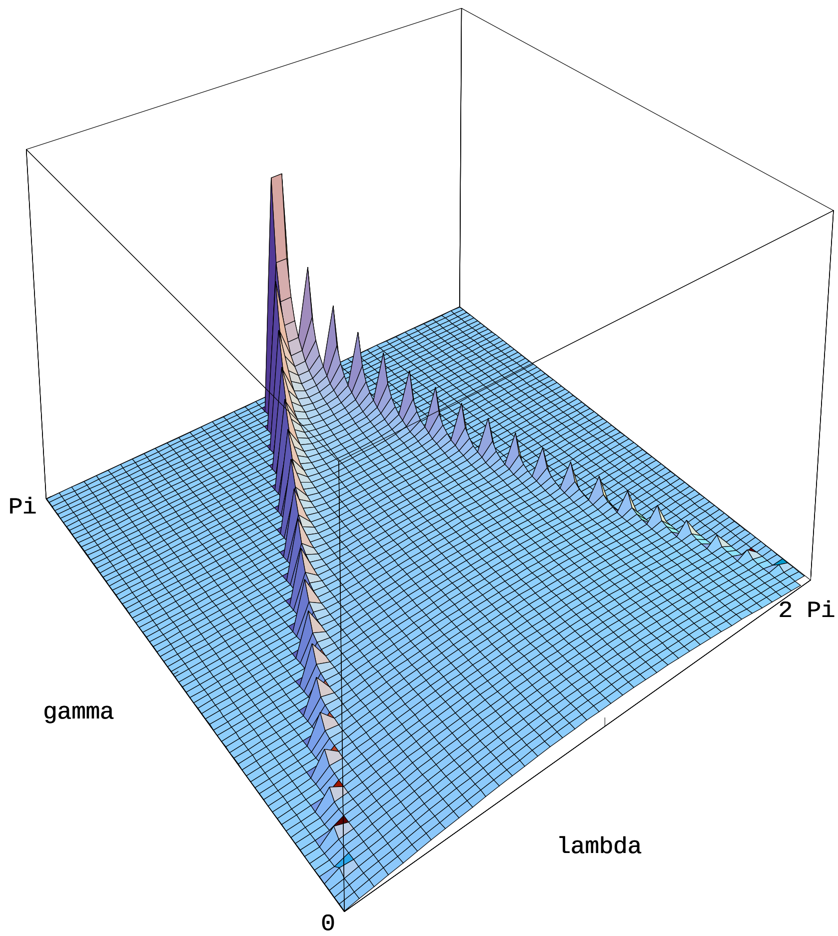

Figure 4 The function g(λ | γ) as a surface plot.

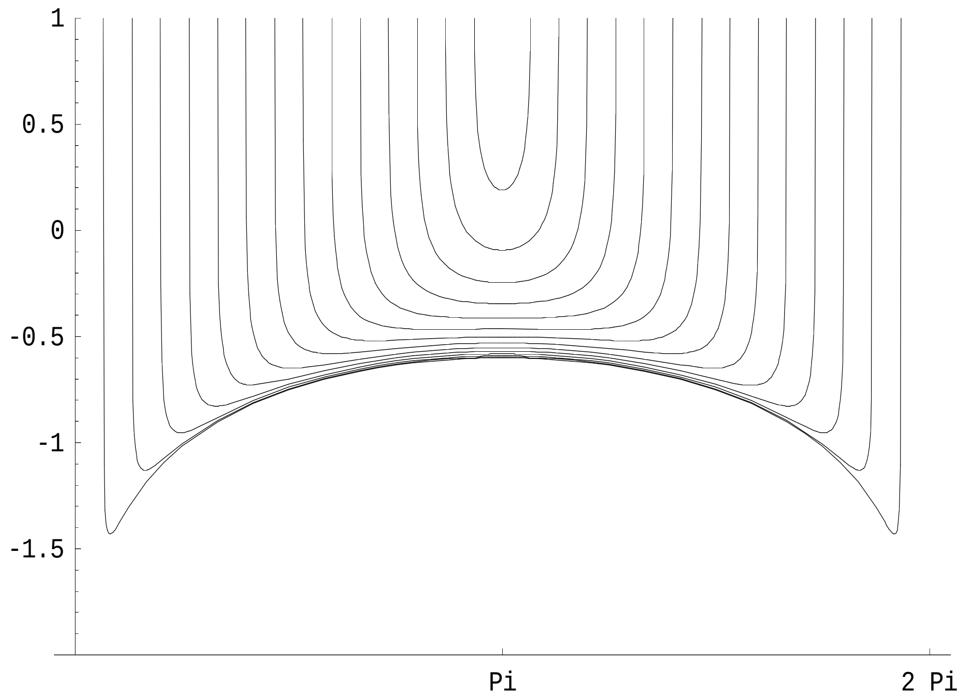

Figure 5 The function $\log_{10}[g(\lambda \mid \gamma)]$ for $\gamma = k\pi/15$ (k = 1,..., 14).

Proof. For a given λ and γ the locus of R is the spherical ellipse. Hence the probability that R is placed so the distance RP + RQ is less than some value $\lambda > \gamma$, is just the area function $G(\lambda, \gamma)$ given by Theorem 5. Then $g(\lambda \mid \gamma)$ must be the partial derivative of G with respect to λ keeping γ constant.

In (30) the terms E_m and F_m are functions of γ only and so are constants with respect to the differentiation. Thus we obtain

$$g(\lambda|\gamma) = \frac{\partial G}{\partial \lambda} = -\frac{1}{\pi}\left((E_m - F_m)\frac{\partial F}{\partial \lambda} + F_m\frac{\partial E}{\partial \lambda}\right) \tag{38}$$

Considering first the function $E(\nu/2, m')$ we have

$$E(\tfrac{\nu}{2},\ m') = \int_0^{\frac{\nu}{2}} \frac{d\phi}{\sqrt{1 - m' \sin^2\phi}} \tag{39}$$

where by (23) and $\nu = \pi - \mu$ we have $\sin\nu/2 = (\cos\lambda/2)/\cos\gamma/2$. Using the substitution

$$\sin\phi = \frac{\cos\frac{\theta}{2}}{\cos\frac{\gamma}{2}} \tag{40}$$

the integral transforms to

$$E(\tfrac{\nu}{2},\ m') = -\frac{1}{2}\int_\pi^\lambda \frac{\sin^2\frac{\theta}{2}\ d\theta}{\sqrt{\cos^2\frac{\gamma}{2} - \cos^2\frac{\theta}{2}}} \tag{41}$$

The partial derivative of (41) with respect to λ is now just value of the integrand at $\theta = \lambda$, i.e.

$$\frac{\partial E}{\partial \lambda} = -\frac{1}{2}\frac{\sin^2\frac{\lambda}{2}}{\sqrt{\cos^2\frac{\gamma}{2} - \cos^2\frac{\lambda}{2}}} \tag{42}$$

Similarly

$$\frac{\partial F}{\partial \lambda} = -\frac{1}{2}\frac{1}{\sqrt{\cos^2\frac{\gamma}{2} - \cos^2\frac{\lambda}{2}}} \tag{43}$$

Now substituting from (42) and (43) into (38) we arrive at the statement of the theorem. That the integral of $g(\lambda \mid \gamma)$ with respect to λ, over the range $\gamma < \lambda < 2\pi - \gamma$, is 1 therefore turns out to be equivalent to Legendre's identity (20).

A pseudo-elliptic integral.

We digress briefly to mention a curious integral which can be derived from Theorem 6.

Integrals which look as if they should result in an elliptic function but which actually evaluate to trigonometric functions are often called pseudo-elliptic integrals. Theorem 6 leads to an integral *of elliptic functions* which evaluates to trigonometric functions (a slightly different kind of pseudo-elliptic integral) that we were not able to trace in the literature but which is probably an example of a class of such integrals. The integral in question arises in the following way.

Since the probability density function for γ is $(\sin\gamma)/2$ the unconditional probability density for λ is given by

$$g(\lambda) = \frac{1}{2} \int_{\gamma=0}^{\min\{\lambda,\ 2\pi-\lambda\}} g(\lambda|\gamma)\ \sin\gamma\ d\gamma \tag{44}$$

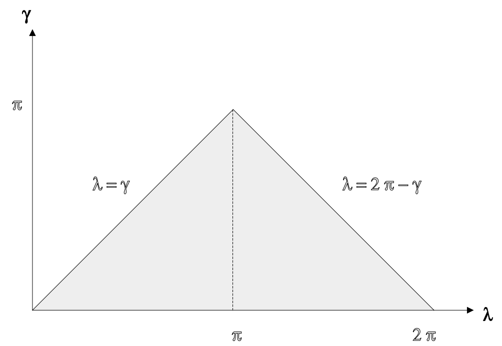

Figure 6 Region over which $g(\lambda \mid \gamma)$ is positive.

In Figure 6 the region in which $g(\lambda \mid \gamma)$ is non-zero is shaded and $g(\lambda \mid \gamma)$ has an axis of λ-symmetry about $\lambda = \pi$. Hence the integral $g(\lambda)$ is symmetric about $\lambda = \pi$ and for the moment we shall assume that $0 \le \lambda \le \pi$.

Slightly manipulating the result in Theorem 6 we can write (44) as

$$g(\lambda) = \frac{1}{2\pi}\int_{\gamma=0}^{\lambda} F_m\ \Delta\ \sin\gamma\ d\gamma + \frac{1}{2\pi}\int_{\gamma=0}^{\lambda} (E_m - F_m\cos^2\frac{\gamma}{2})\ \frac{\sin\gamma}{\Delta}\ d\gamma \tag{45}$$

where $m = \sin^2(\gamma/2)$ and

$$\Delta = \sqrt{\cos^2\frac{\gamma}{2} - \cos^2\frac{\lambda}{2}} \tag{46}$$

The following indefinite integral for the complete elliptic function F_m with respect to m is readily verified

$$\frac{1}{2}\int F_m dm = E_m - m' F_m \tag{47}$$

Using $m = \sin^2(\gamma/2)$ this enables us to integrate the first term in (45) by parts and show that

$$\int_{\gamma=0}^{\lambda} F_m \, \Delta \, \sin\gamma \; d\gamma \; = \; \int_{\gamma=0}^{\lambda} (E_m \; - \; F_m \cos^2\frac{\gamma}{2}) \; \frac{\sin\gamma}{\Delta} \; d\gamma \qquad (48)$$

i.e. the two terms in (45) are equal. Hence

Theorem 7. For $0 \leq \lambda \leq \pi$

$$g(\lambda) \; = \; \frac{1}{2}\int_{\gamma=0}^{\lambda} g(\lambda|\gamma)\sin\gamma \; d\gamma \; = \; \frac{1}{\pi}\int_{\gamma=0}^{\lambda} F(\frac{\pi}{2}, \; \sin^2\frac{\gamma}{2})\sqrt{\cos^2\frac{\gamma}{2} \; - \; \cos^2\frac{\lambda}{2}} \;\; \sin\gamma \; d\gamma \qquad (49)$$

This expresses g(λ) as a particular integral involving a complete elliptic function. However, g(λ) is in fact the density function for $\lambda = \alpha + \beta$ independent of γ and so is just the convolution of two independent arc lengths given in Theorem 3. Using the fact that g(λ) is symmetric about π this yields

Theorem 8.

$$\frac{1}{2\pi}\int_{\gamma=0}^{\lambda} F(\frac{\pi}{2}, \; \sin^2\frac{\gamma}{2})\sqrt{\cos^2\frac{\gamma}{2} \; - \; \cos^2\frac{\lambda}{2}} \;\; \sin\gamma \; d\gamma \; =$$

$$\begin{cases} \frac{\pi}{8}\left(\sin\lambda \; - \; \lambda\cos\lambda\right) & (0 \leq \lambda \leq \pi) \\ \frac{\pi}{8}\left(\sin(2\pi-\lambda) \; - \; (2\pi-\lambda)\cos(2\pi-\lambda)\right) & (\pi \leq \lambda \leq 2\pi) \end{cases} \qquad (50)$$

We were also able to verify (50) directly by expanding the complete elliptic function as a power series in m, but not by any subtle substitution.

The probability density function of r = $\alpha + \beta + \gamma$.

Since

$$T_3(r) \; = \; \frac{1}{2}\int_{\gamma \, = \, 0}^{r/2} g(r-\gamma|\gamma) \; \sin\gamma \; d\gamma \qquad (51)$$

we have immediately from Theorem 6

Theorem 9. For $0 \le r < 2\pi$

$$T_3(r) = \frac{1}{4\pi} \int_{\gamma=0}^{r/2} \left(\frac{E(\pi/2,\ \sin^2\gamma/2) - F(\pi/2,\ \sin^2\gamma/2)\ \cos^2(r-\gamma)/2}{\sqrt{\cos^2\gamma/2 - \cos^2(r-\gamma)/2}} \right) \sin\gamma\ d\gamma \qquad (52)$$

The integral is improper at $\gamma = r/2$ ($r < 2\pi$) but converges. At $r = 2\pi$ the integral diverges at the right endpoint and $T_3(r)$ has a singularity due to the singularity of $F(\pi/2, m)$ at $m = 1$.

Thus we have the interesting result that the most probable length of a triangular tour on a sphere is equal to the circumference, and this is the maximum possible tour length on 3 points.

The righthand side of (52) seems remarkably difficult to integrate explicitly. A surface plot of the integrand is given in Figure 7.

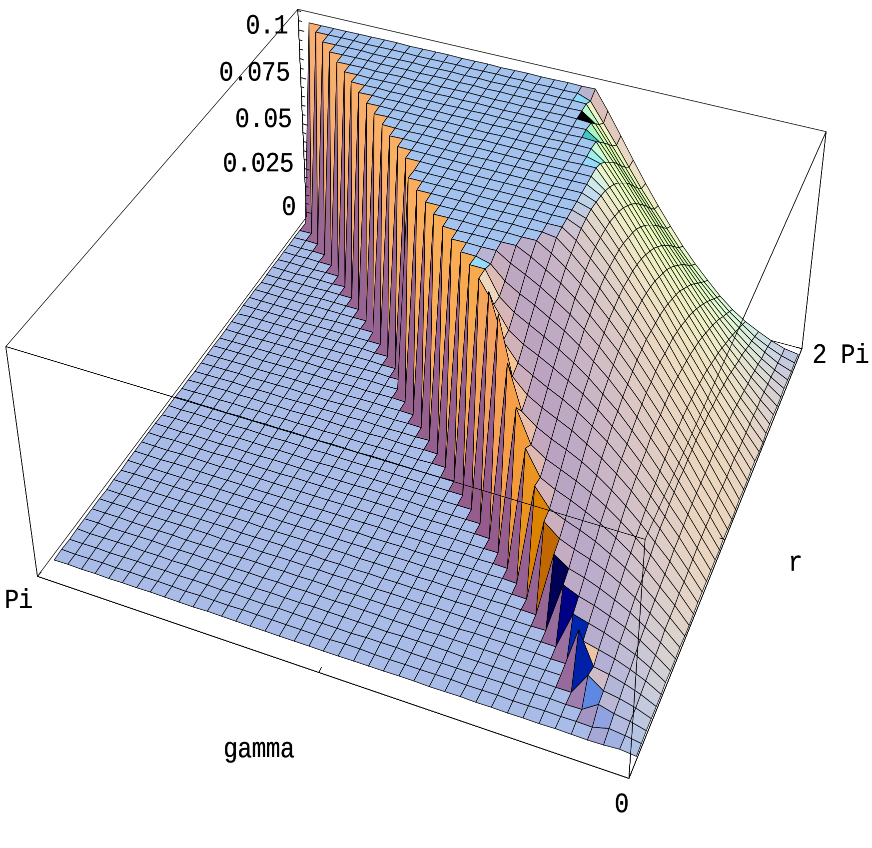

Figure 7 The integrand as a function of r and γ.

We can use (52) to compute the graph of $T_3(r)$ by numerical integration. This yields the top graph of Figure 8.

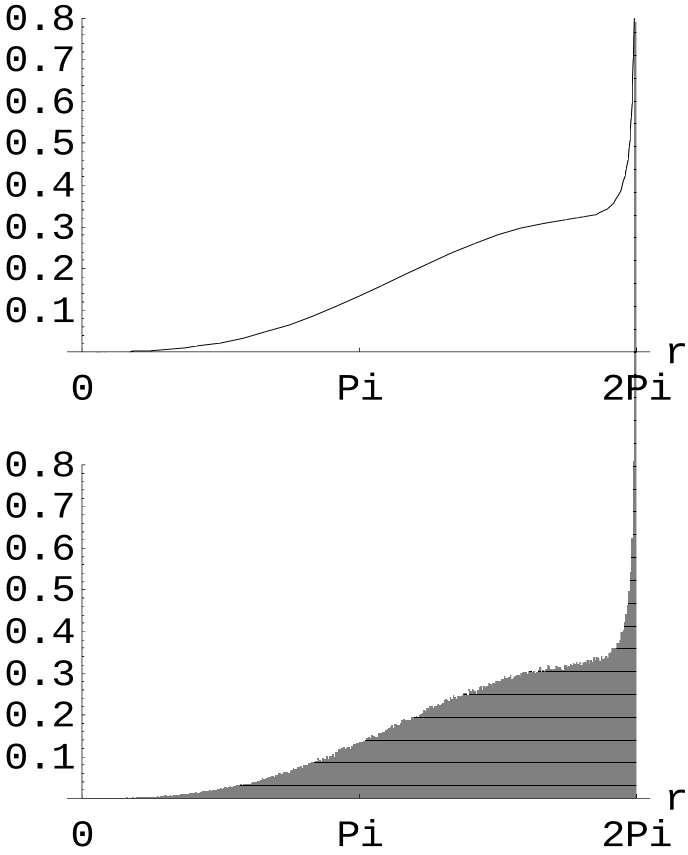

Figure 8 Top: $T_3(r)$ derived from (52) by numerical integration. Bottom: Histogram for 1 million trials - the vertical axis is proportional frequency nomalised by a factor $360/2\pi$ (1 degree bins).

Experimental results.

It is straightforward to carry out a simulation of randomly scattering points on the surface of a sphere. If we adopt a (*longitude*, *latitude*) system of coordinates, then the longitude α is uniformly distributed over $(0, 2\pi)$, while the latitude β is distributed according to the density function $\frac{1}{2}\cos\beta$. Hence if u is a uniformly distributed random variable on (0,1), samples for α and β are given by

$$\alpha = 2\pi u$$
$$\beta = \sin^{-1}(2u-1) \tag{53}$$

and the distance r between any two points (α_1, β_1), (α_2, β_2) is given by

$$\cos r = \sin\beta_1 \sin\beta_2 + \cos\beta_1 \cos\beta_2 \cos(\alpha_1 - \alpha_2) \tag{54}$$

Samples for u were generated using the triple pseudo-random number generator of Wichmann and Hill [Wichmann 1982] which has a period of ~ 3.10^{13} and proven good random characteristics. The experimental procedure was to place three points at random and measure the sum of the sides of the triangle, this being the length of a closed tour on three points. The result of carrying out 10^6 trials gives the frequency histogram shown in the bottom graph of Figure 8.

Remark. In fact any odd number of points will have the property that the maximum tour length is $(n\text{-}1)\pi$ rather than $n\pi$. However, this requires that $(n\text{-}1)/2$ points are concentrated at one pole and $(n\text{-}1)/2$ points at the other. The probability that this will occur, rapidly tends to zero as n increases, and for $n > 5$ the phenomenon of a sharp cutoff at $(n\text{-}1)\pi$ is not observed.

Conclusion.

We have determined an integral formula for $T_3(r)$ and shown that experimental data closely conforms to the resulting predictions. Knowing the distribution of triangular tour lengths on the surface of a sphere also tells us the distribution of areas and of the sum of the angles. As a byproduct of the investigation we also demonstrated in Theorem 5 that the area of a 'spherical ellipse' is simply expressed in terms of Heuman's lambda function. Knowledge of $T_3(r)$ also gives (in a rather trivial sense - because there is essentially only one tour when $n = 3$) knowledge of the distribution of *minimal* tour lengths for the case $n = 3$; this is of interest since the detailed nature of this distribution has remained an intractable problem, certainly since the Beardwood, Halton and Hammersley paper.

There remains the goal of deriving as much information as possible regarding $T_n(r)$ for general n. In particular an accurate estimate for $T_n(r)$ in the vicinity of $\sqrt{n}$ when n is large would be extremely useful. One way to approach this is first to derive the distribution of the length of an *open* tour obtained by placing n points at random and joining them up in the order generated. This is $\sigma_{n-1}(r)$ and is given in Theorem 2 (although we have postponed the proof until the second paper in this series). Second we might then convolve this distribution with the distribution of the distance between the last and first points. Although it might appear, since the last point is placed randomly, that this distance should also be distributed as d(r), the case $n = 3$ amply illustrates that the closing edge length distribution actually depends upon the sum of the first n-1 edges. This is because that sum constrains the position of the n*th* point so that it cannot be uniformly distributed over the surface of the sphere.

In general if we let the sum of the first n-1 edges be λ_{n-1}, and the length of the closing edge be γ, then the probability density function of γ is $h_n(\gamma|\lambda_{n-1})$, where in the case $n = 3$,

$h_3 = h(\gamma|\lambda)$ is the 'conditional inverse' of the function $g(\lambda|\gamma)$ given by Theorem 6. We then have

$$T_n = \sigma_{n-1} \otimes h_n \qquad (55)$$

where $\otimes$ denotes convolution. Our belief is that, as *n* becomes large, the constraint imposed by λ_{n-1} on the placing of the last point becomes progressively weaker so that $h_n(\gamma|\lambda_{n-1}) \to d(\gamma)$ as $n \to \infty$, and $T_n \sim \sigma_n$ as $n \to \infty$. As remarked in the introduction, simulation experiments suggest that $T_n - \sigma_n \to 0$ quite rapidly, but it remains to construct a formal proof.

References

[Beardwood 1959] J. Beardwood, J. H. Halton and J. M. Hammersley. *The shortest path through many points.* Proc. Cambridge Philos. Soc. **55**, 299-327, 1959.

[Borel 1925] E. Borel. *Principes et formules classiques du Calcul des Probabilités.* Traité du Calcul des probabilités et de ses Applications. Paris, Gauthier-Villars, 1925.

[Byrd 1971] Paul F. Byrd and Morris D. Friedman. *Handbook of Elliptic Integrals for Engineers and Scientists* (Second Edition, Revised), Springer-Verlag, 1971.

[Gečiauskas 1966] E. Gečiauskas. *Distribution function of a distance between two random points in an oval*. Lietucos Matematikos Rinkinys, **6**, 245-248 (In Russian).

[Grötschel 1991] M. Grötschel and O. Holland. *Solutions of large-scale symmetric traveling salesman problems*. Mathematical Programming **51**:141-202. North Holland, 1991.

[Held 1970] M. Held and R. M. Karp. *The travelling salesman problem and minimum spanning trees*. Oper. Res. **18**:1138-1162, 1970.

[Held 1971] M. Held and R. M. Karp. *The travelling salesman problem and minimum spanning trees: part II*. Math. Programming **1**:6-25, 1971.

[Heuman 1941] C. Heuman. *Tables of Complete Elliptic Integrals*. Journal of Mathematics and Physics **20**:127-144, 1941.

[Karp 1977] R. M. Karp. *Probabilistic analysis of partitioning algorithm for the Travelling-Salesman Problem in the plane*. Mathematics of Operations Research, **2**(3):209-224, August 1977.

[Kendall 1963] M. G. Kendall and P. A. P. Moran. *Geometrical Probability*. Charles Griffin & Co., London, 1963.

[Stein 1977] D. Stein. *Scheduling Dial-a-Ride transportation systems: an asymptotic approach.* Ph.D. thesis, Harvard University, Cambridge, MA. 1977.

[Wichmann 1982] B. A. Wichmann and I. D. Hill. *A pseudo-random number generator*. National Physical Laboratory, U.K. Report DITC 6/82, 1982.