Biographical Note for Dan Flath

Dan Flath, whose first education was in electrical engineering, teaches mathematics at Macalester College in St. Paul, Minnesota, and writes on number theory, algebra, and calculus. The rest of his time is devoted to his family, including five children.

Biographical Note for Stan Wagon

Stan Wagon is Professor of Mathematics and Computer Science at Macalester College in St. Paul, Minnesota. His recent work has centered on applications of *Mathematica* to both abstract mathematics and concrete modeling, and he has written several books to illustrate the power of modern symbolic and numerical computing. He was on one of the first-place teams in the 2002 SIAM 100-Digit Challenge, and that led to a book about the ten challenging problems of that contest. His other interests include the many different types of cross-country skiing: recent highlights include the completion of a 100-mile ski race and a 10-day ski traverse in the Monashee Range of Canada during which his group sighted a wolverine, which is very rare. Math and snow can be combined, and he has captained several prize-winning snow sculpture teams that built giant snow sculptures with mathematical themes (see http://stanwagon.com for images).

Finding a Hidden Coin

DAN FLATH
STAN WAGON
Macalester College
St. Paul, Minnesota 55105
{flath,wagon}@macalester.edu

1. The St. Basil Cake

In the Greek Orthodox tradition New Year's Day is also the Feast Day for St. Basil (329–379), who was known for his generosity and support of philanthropic projects ([4]; see [5] for a recipe). There are two stories that associate St. Basil with coins in cakes: (1) he had cakes baked for the poor and included coins in the cake as a way of helping out those in need; (2) when Emperor Valens demanded taxes from his region, St. Basil gathered jewelry and coins from those who could afford it. But Basil's piety so impressed the emperor's emissary that the items were returned. No record was kept of what belonged to whom, so St. Basil made a cake with the materials baked into it. By a miracle, each recipient got the slice with his own items in it. Well, the probability of that is so small that it would indeed be a miracle.

In this article we will discuss a mathematical problem arising from the modern St. Basil tradition, which calls for the inclusion of a coin in the batter of the St. Basil's Cake, also called *Vasilopitta*; the person who gets the coin in his or her slice will have good luck in the coming year.

There are apparently two traditions: the classic one is to include a coin in the batter, but another is to slide the coin under the cake as it is placed on a serving platter. These versions leads to two distinct mathematical problems, in the style of the Buffon needle problem. Indeed, there is even a third tradition calling for additional trinkets to be baked in the cake, but we will content ourselves with a single coin.

Figure 1 A St. Basil's cake of 216-mm diameter with a 24-mm diameter U.S. quarter baked into it. Four cuts yielded one that struck the coin. Do you think the probability of such a strike is more or less than 50%? (Photo by Ruth Dover.)

The coin-on-the-bottom protocol leads to the following problem in geometric probability, which was discussed in [**1**]:

> Suppose a round cake has radius R^*, a round coin has radius r, and the coin is placed randomly in the cake (parallel to the base). Suppose n slices through the center are made, cutting the cake into $2\,n$ congruent pieces. What is the probability that the knife strikes the coin?

This version has the coin in a horizontal position. If, more generally, the coin is in an arbitrary orientation then its horizontal projection is an ellipse, and the corresponding geometric problem is quite a bit harder. In this paper we will show how to solve either version.

One must clarify the meaning of *random*. The natural way to do it is to assume that the coin is placed so that it does not protrude outside the cake. That is, we assume that the center of the coin is placed randomly within the disk of radius $R = R^* - r$.

The problem was investigated by Savvidou in [**1**] under the assumption that the coin is horizontal. This is the same as mixing a sphere, such as a marble, into the batter, since its projection in any orientation is a disk, so we will often refer to this as the sphere problem. But the results of [**1**] are based on the assumption that the distance of the coin's center from the center of the cake is uniformly distributed in [0, R]; this does not conform to the reality of kneading and baking, since it implies that the coin is as likely as not to be closer to the center than the border. For typical dimensions the coin is much more likely to lie closer to the boundary, because the area of the inner disk of radius $R/2$ has area that is a quarter that of the full disk of radius R.

In this paper we will first give a complete solution to the horizontal coin problem, where elementary geometry yields a simple formula for the probability. Then we will attack the general orientation problem, which first requires solving the horizontal problem for an ellipse. We will show how symbolic algebra (equation-solving and integration) can be combined with numerical algorithms (finding roots of a polynomial and double integration) to compute the probability that a knife strikes the tilted coin. Our model yields probabilities that are about half of those presented in [**1**].

2. Finding the Marble

We will start by getting a complete solution to the marble problem. Since the projection is a disk, this is identical to the problem of the horizontal coin. We assume that the distribution of the disk is uniform with respect to area in the plane; that is, the probability of finding the center of the disk in any region is proportional to the region's area.

Now, to solve the problem we note that the probability that the knife strikes the disk can be computed by adding up the areas of the $2n$ congruent regions where the disk's center could lie so that the disk avoids the knife (see Fig. 1). We call these regions the *safe regions*, and their complement *unsafe*. This is a simple geometry problem, since all the relevant dimensions and areas can be determined by standard formulas. The safe regions shown in Figure 1 could be moved inward so that they meet tightly at a common center and form a circle-like region, but it would not be a perfect circle (because the distance from M to the border (see Fig. 3) is not constant).

We let r denote the radius of the disk that represents the coin and, in this section only, assume that the cake's radius is $1 + r$. Since the disk's center cannot be within r of the cake's boundary, the circle of possibilities (the *viable region*) for the center has radius 1. If the cake has radius R and the disk has radius r_1, then one can translate to our assumption by setting $r = r_1/(R - r_1)$. Figure 2 shows how the widened cuts and shrunken rim (both by r) define the safe regions, which are shaded. Figure 3 is a closeup that shows the various dimensions needed to determine the area of one of the safe pieces.

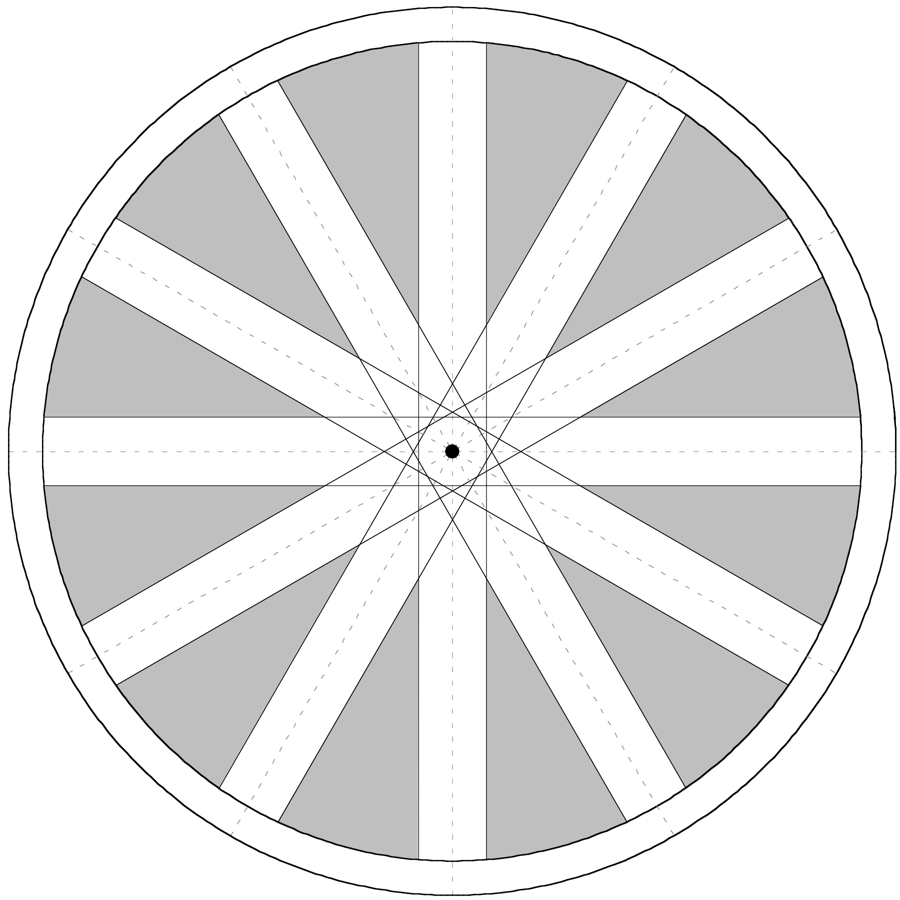


Figure 2 The shaded sectors are the locations of the center of a marble that will not be struck by the knife.

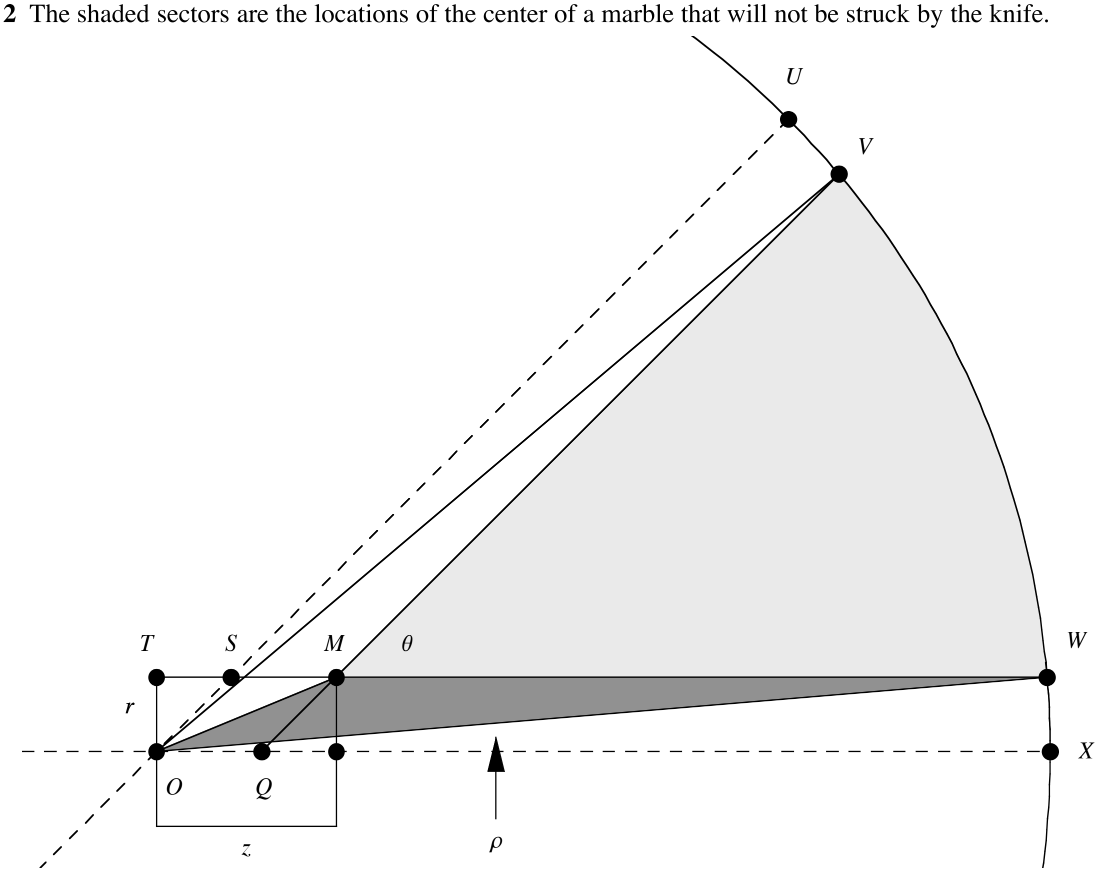


Figure 3 A close-up view of the sector determined by two cuts (dashed lines).

Refer to Figure 3, in which the viable circle has radius 1; we use θ for the angle of size π/n where the cuts (dashed) meet. We assume here that $r < \sin(\theta/2)$, for otherwise point M is not inside the circle and the safe region is empty. In other words, if $n \geq \pi/(2 \arcsin r)$ then the probability of a hit is 1. We need z, the horizontal extent of rhombus $OQMS$; it is $r \cot\theta + r \csc\theta$ (to derive the second summand, draw the perpendicular from M to the angled cut; if $n = 1$ then M is taken to be $(0, r)$ and $z = 0$).

The safe region consists of $2n$ copies of region MVW. This area can be computed by taking the area of the circular sector OVW and subtracting twice the area of $\triangle MOW$. We need ρ, the angle shown in the diagram; it is $\arcsin r$. The sector then has area $\frac{\pi}{2n} - \rho$ and the two triangles, having height r and base $|TW| - z$, have total area $r\left(\sqrt{1-r^2} - r\cot\theta - r\csc\theta\right)$. It follows that the entire safe area in the upper half of the circle is

$$n\left(\frac{\pi}{2n} - \rho - r\left(\sqrt{1-r^2} - r\cot\theta - r\csc\theta\right)\right).$$

Subtracting from $\pi/2$, dividing by $\pi/2$, and using a trig identity gives the probability of a hit, which is

$$P(r, n) = \begin{cases} \frac{2n}{\pi}\left(\arcsin r + r\left(\sqrt{1-r^2} - r\cot\frac{\pi}{2n}\right)\right), & \text{if } n < \pi/(2\arcsin r) \\ 1, & \text{if } n \geq \pi/(2\arcsin r) \end{cases}$$

If we have a cake of radius R and sphere of radius r, then $P(R, r, n)$, the probability of a hit, is $P(r/(R-r), n)$.

Thus we learn that if the cake has radius 16 cm and the marble has radius 1 cm then, if $n = 2$, the probability of a hit is 0.16. When $n = 4$ it is 0.31. In [**1**] these two probabilities were presented as 0.30 and 0.47, respectively.

The simple formula allows for easy exploration. For example, suppose one wishes to know how large the cake should be so that the probability of finding a sphere of half-inch radius using 5 cuts is exactly $1/2$. Simple root-finding shows that this happens if the cake's diameter is 11.92 inches.

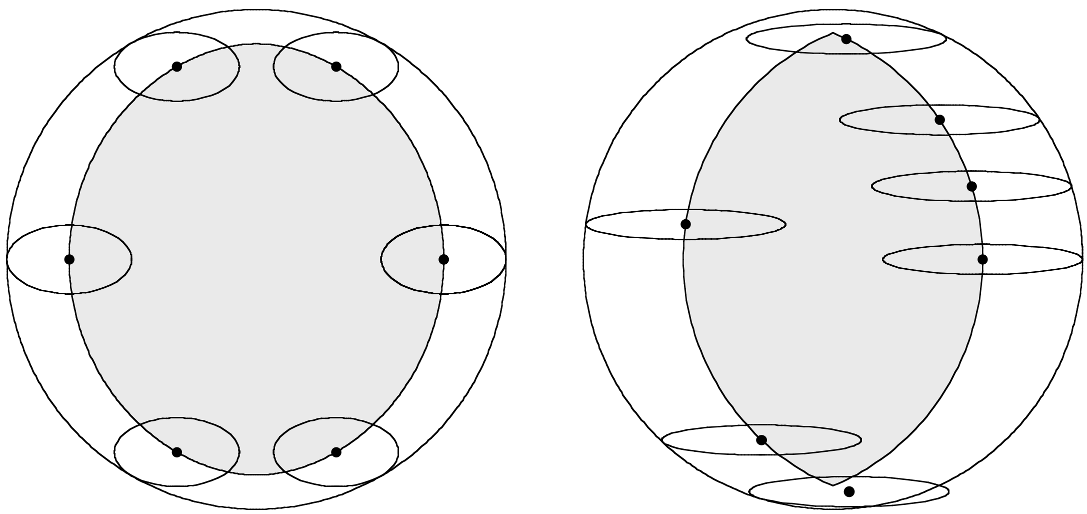
3. Striking an Ellipse

Before considering arbitrary orientations of a coin in space, we need to extend the work of §2 to the case that the planar shape being hidden in the thin cake is an ellipse lying in the x-y plane. That is, we want the probability of the knife striking an ellipse with semimajor axis r and semiminor axis m. We allow the ellipse to be skewed so that its major axis is at angle η from the x-axis. Thus we use $e_\eta(t)$ for the skewed ellipse $R_\eta \cdot (r\cos t, m\sin t)$, where R_η is the matrix that rotates η radians in the positive direction. We use just $e(t)$ for the unskewed ellipse $e_0(t)$; moreover, we generally think of this ellipse as being mobile: having an arbitrary center. We always take the first of the n slices to coincide with the horizontal diameter of the cake. The cake is assumed to have radius 1. We use ϵ to denote m/r.

Our initial approach, which did work, was entirely numerical, using algorithms for optimization, root-finding, and integration to get the needed areas. But an algebraic method, when it exists, will usually speed things up tremendously; further, the two approaches can be used as a check on each other. There are times when numerics is simpler, since one can avoid issues of root-choice and other special cases; a numerical approach is also more general, as it will work in situations where the algebra is hopeless. But algebraic formulas have an elegance that can't be beat, and modern software makes it so painless to solve quartics or perform symbolic integrations.

The Viable Region

Given an unskewed ellipse $e(t)$, the first step is to determine the *viable region* V. By this, we mean the region where the center of $e(t)$ can lie so that the ellipse lies entirely inside the unit disk (Fig. 4); the boundary of this region, ∂V, is called the *viable curve*. Thus we seek a parametrization $\boldsymbol{p}(b)$ of ∂V; it is easy to see that V must be convex. The following theorem derives the parametrization and, as an unexpected but useful bonus, shows how the region and its boundary can be described using pure polynomial algebra.



(a) (b)

Figure 4 The viable regions are shaded. In (b), $m < r^2$ and there are vertices at the top and bottom of the curve; they arise because of the discontinuity in the point on the ellipse that touches the cake boundary.

Theorem 1(a). The viable region is the component of the origin in the complement of the curve defined by

$$\boldsymbol{p}(b) = (x(b),\, y(b)) = \left(\cos b - \frac{r^2 \cos b}{\sqrt{r^2 \cos^2 b + m^2 \sin^2 b}},\ \sin b - \frac{m^2 \sin b}{\sqrt{r^2 \cos^2 b + m^2 \sin^2 b}}\right)$$

(b). The viable boundary has a vertex if and only if $m < r^2$.

(c). Let $\alpha = 1 - x^2 - y^2 + m^2 + r^2$ and $\tau = m^2(1 - x^2 + r^2) - r^2(y^2 - 1)$. Then the parametrized curve of (a) is the set of points (x, y) inside the unit circle for which $\delta = 0$ where

$$\delta = -4\, m^2\, r^2\, \alpha^3 + \alpha^2\, \tau^2 + 18\, m^2\, r^2\, \alpha\, \tau - 4\, \tau^3 - 27\, m^4\, r^4.$$

(d). The highest point of V has y-coordinate $1 - m$ if $m \geq r^2$ and $\sqrt{(r^2 - m^2)\,(1 - r^2)}\,/\,r$ otherwise.

Proof. (a) For a polar angle b in the first quadrant, we ask: Which point $e(\tau)$ on the right half of the ellipse is such that the tangent to the ellipse at that point is perpendicular to the b-radius of the unit circle? The answer (see Fig. 4) is given by the dot product $e'(\tau) \cdot (\cos b, \sin b) = 0$, which reduces to $\tau = \arctan(\epsilon \tan b)$. Then the ellipse will just touch the unit circle if the center is such that $e(\tau)$ is on the unit circle and has polar angle b. This occurs if the center is at $(\cos b, \sin b) - e(\tau)$. Therefore the locus of the centers is given by

$$(\cos b - r \cos(\arctan(\epsilon \tan b)), \ \sin b - m \sin(\arctan(\epsilon \tan b)))$$

and standard trig identities reduce this to the asserted parametrization (taking care that the formula is correct in all four quadrants).

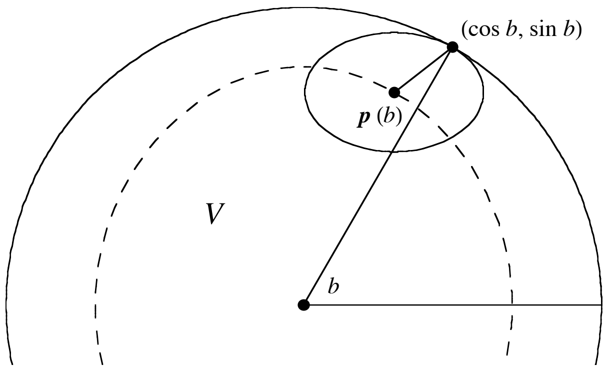


Figure 5 An ellipse tangent to the unit circle on the inside.

(b). If the ellipse is very eccentric then it can be tangent to the circle on the inside but also protrude outside (e.g., the lowest ellipse in Fig. 4(b); see also Fig. 7). For this not to happen the radius of curvature of the ellipse must be no greater than 1 at each point. But we can resolve this purely algebraically, since a protrusion arises iff there is a b between 0 and $\pi/2$ so that $\boldsymbol{p}(b)$ is left of the y-axis. Using (a) this condition reduces to $(1-\epsilon^2)\cos^2 b < r^2 - \epsilon^2$. This condition can occur iff $\epsilon < r$, or $m < r^2$, in which case it happens for those b such that $b > \arccos\sqrt{(r^4 - m^2)/(r^2 - m^2)}$. So if we let $b_{\max}$ denote the largest b in the first quadrant so that $\boldsymbol{p}(b)$ is in V, then $b_{\max} = \pi/2$ unless $m < r^2$, in which case it is $\arccos\sqrt{(r^4 - m^2)/(r^2 - m^2)}$. It turns out that both cases are covered by a one-liner in $\mathbb{C}$: $b_{\max} = \mathrm{Re}\left(\arccos\sqrt{(r^4 - m^2)/(r^2 - m^2)}\right)$ (with a special case if $m = r$).

(c). This expression can be obtained from the parameterization in (a) by using *Mathematica*'s `Eliminate` function. But the result can also be derived from classic conic section algebra since a point is on ∂V iff a certain ellipse intersects the unit circle in precisely one point; for details see [**2**, chap. 18]. An illustrative plot of $\delta = 0$ appears in Figure 6. For the interior of the viable region the expression in (c) is positive.

(d). Details are omitted (easy algebra).

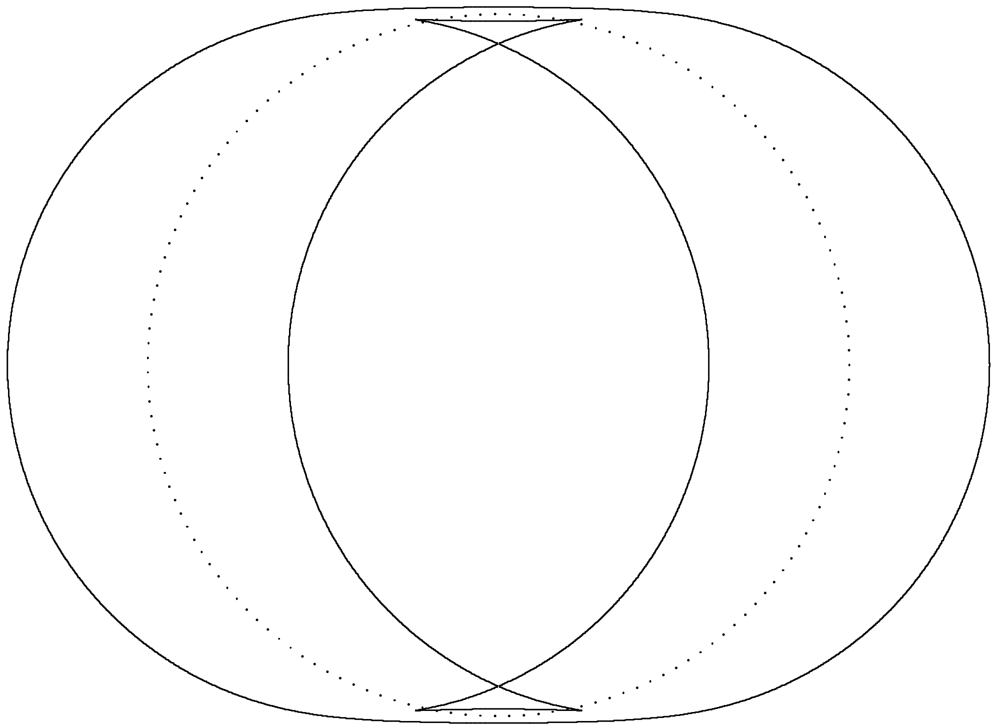


Figure 6 The viable region can be described by a polynomial condition, which is satisfied by the two solid curves; but that includes points outside the (dashed) unit circle.

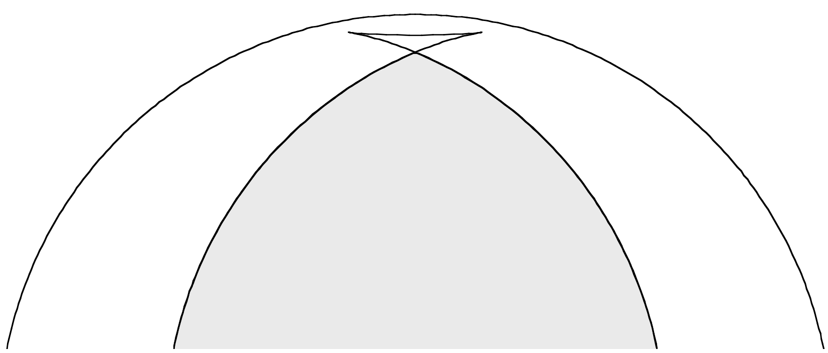


Figure 7 When $m < r^2$ the viable region has a vertex and the curve defined by $\boldsymbol{p}(b)$ extends outside the region.

Corollary 1. If $0 \le b_1 \le b_2 \le b_{\max}$, then the area of that part of V that lies between the rays obtained by connecting the origin to the points $\boldsymbol{p}(b_i)$ is

$$\int_{b_1}^{b_2} \frac{1}{2}\,(x(b),\, y(b)) \cdot (y'(b),\, -x'(b))\ db$$

Proof. This is a consequence of Green's Theorem, one case of which asserts that for

a simply connected region R with nice boundary curve ∂R in the plane

$$\text{Area of } R = \int\int_R dx\, dy = \int_{\partial R} \frac{-y\,dx + x\,dy}{2}$$

In our application where R is a wedge of V, the integrand

$$-y\,dx + x\,dy = (-r \sin\theta)\, d(r\cos\theta)\ +\ r\cos\theta\, d(r\sin\theta) = 0\, dr$$

vanishes along the boundary rays from the origin with θ constant.

Now it came as a bit of a surprise that the integral of Corollary 1 can be expressed in closed form (the integrand, which we do not show here explicitly, is quite complicated). We must be liberal and allow the use of elliptic functions, but that is perfectly reasonable as there are very fast algorithms for computing them. Thus we let $E(\phi \mid M)$ denote the elliptic integral of the second kind, which is $\int_0^{\phi} 1 \Big/ \sqrt{(1 - M \sin^2 \theta)}\, d\theta$ (see [**3**]).

Corollary 2. If $0 \le b_1 \le b_2 \le b_{\max}$, then the area of the viable region enclosed by the rays from the origin to $\boldsymbol{p}(b_1)$ and $\boldsymbol{p}(b_2)$ is $A(b_2) - A(b_1)$, where

$$A(b) =$$
$$\tfrac{1}{4}\left(2\,b + 2\,m\,r \arctan(\epsilon \tan b)\; - 4\,r\,E(b \mid 1 - \epsilon^2) + \sqrt{2}\;(r^2 - m^2)\,(\sin 2\,b) \Big/ \sqrt{m^2 + r^2 + (r^2 - m^2)\cos}\right.$$

and the arctan term is set to $\pi/2$ if $\tan b$ is infinite. Note that $A(0) = 0$ and if $m = 0$ then $A(b)$ is simply $(b - r \sin b)/2$.

Proof. Simply obtain the explicit expression for the integrand in Corollary 1 and integrate using computer algebra. As a check, take the derivative of the result. Indeed, simply taking the derivative of $A(b)$ proves the corollary. As a further check one can make sure the symbolic result agrees with numerical integration over random intervals.

To get the area over a b-interval not contained in the first quadrant, just break up the interval as necessary so that the symmetry of V can be used.

The formula for $\boldsymbol{p}(b)$ makes no reference to the cuts in the cake, so they all work for an ellipse oriented at an acute angle η to the horizontal. For example, the viable boundary for the ellipse $e_\eta(t)$ is given by $R_\eta \cdot \boldsymbol{p}(b)$, the rotation of $\boldsymbol{p}(b)$ by η.

The Thickness of an Ellipse

Next we bring in the cuts and find the unsafe strips: the regions parallel to the cuts where the ellipse center cannot lie (Fig. 8). More notation: Let the cuts be $C_1, \ldots, C_n$ where C_1 is the horizontal diameter of the circle and they are in counterclockwise order. Let C_i^+ and C_i^- denote the lines parallel to C_i that define the unsafe strip determined by C_i. The one marked by a negative sign is the one that would be below the equator when the cut and the parallels are rotated clockwise into the horizontal diameter. We sometimes call these lines *tropics*. Because of symmetry we can work exclusively in the upper half of the circle. Let T_i be the sector of the unit circle determined by C_i and C_{i+1}. We can break the viable region in the upper semicircle into n subregions by defining $V_i = V \cap T_i$; this is called the *ith viable region*. Let S_i be the subregion of T_i lying between the tropics C_i^+ and C_{i+1}^- (where C_{n+1}^- always means C_1^+). The vertex of the region S_i (the intersection point of the tropics) is denoted M_i. When $n = 1$ we define M_1 to be the intersection point of the tropic C_1^+ with the infinite line connecting the origin to the highest point of V.

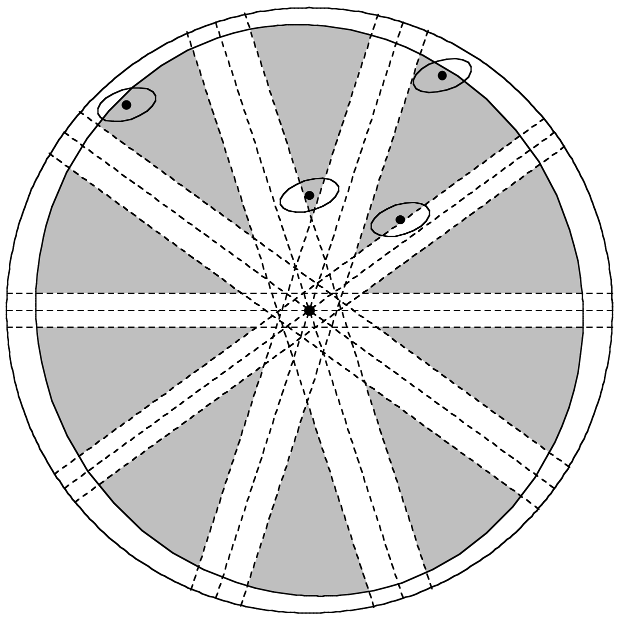


Figure 8 A cake with 5 cuts (dashed). The shaded pieces are the safe regions for the ellipse $e_{0.3}(0.1, 0.05)$. The widths of the unsafe strips vary, as does the radial distance of the viable boundary.

The amount the horizontal cut must be fattened on the positive side — the distance from C_1 to C_1^+ — is simply the length of the vertical projection of the ellipse. Once this is found, rotation will yield the relevant formula for any cut. Let r_i denote the half-width of the ith unsafe strip.

Proposition 1. For an ellipse $e_\eta(t)$ and $1 \le i \le n$,

$$r_i = \sqrt{m^2 \cos^2(\eta - (i-1)\,\pi\,/\,n) + r^2 \sin^2(\eta - (i-1)\,\pi\,/\,n)}\,,$$

Proof. This is a simple calculus exercise: by rotational symmetry we may assume $i = 1$; then we use the y-derivative to find the lowest point on the ellipse and determine the ellipse's projection onto the y-axis. That turns out to be twice the radical in the statement The simplicity of the formula suggests that a proof without calculus exists; it does, but we omit the details since all we need is the formula.

Note that the ith viable region includes parts of the unsafe strips. We will use the term *ith safe region* to refer to the safe part of V_i (the shaded areas in Fig. 8). It seems obvious that the ith safe region is the triangle-like region bordered by the two tropics emanating outward from M_i and the viable curve (which we called S_i). But this needs proof. The following lemma shows that the shape of the safe regions is as one expects from the diagrams.

Lemma 1. Suppose an ellipse $e_\eta(t)$ is given. We work only in the upper half of the circle.

(a). The region S_i is disjoint from all unsafe strips.

(b). If the ith safe region has positive area, then M_i is a safe point, and so is in the interior of the viable region V.

(c). If M_i is in the interior of V, then the ray from M_i along C_i^+ in the direction away from the origin (i.e., the ray that does not cross C_i) intersects ∂V in exactly one point. The same is true for the ray along C_{i+1}^-. The ith safe region is the sector of V determined by M_i and these two intersection points.

Proof. (a). The ellipse $e_\eta(t)$ centered at M_i is tangent to C_i and C_{i+1}. Now take any cut C_k with k different from i and $i+1$.The cut cannot touch the ellipse since the ellipse lies entirely within T_i. It follows that M_i is outside the unsafe strip defined by C_k. Hence this strip does not intersect S_i.

(b). The hypothesis implies that there is a safe point P inside S_i. Consider the line connecting P to M_i. We will show that the entire segment from P to M_i is safe, which shows V contains M_i. Consider the ellipse $e_\eta(t)$ centered at P and slide it down the segment so its center moves to M_i. The only way the center could leave the safe region is by hitting a cut or the unit circle. But the center stays inside S_i so the ellipse cannot hit C_i or C_{i+1}; and by (a) it cannot hit any of the other cuts. Moreover, the ellipse can't hit the circle because its center is moving closer to the center of the circle and so each point on the ellipse is closer to the center (exercise using Law of Cosines). Because the safe region is an open set, M_i must be in the interior of V.

(c). The convexity of V implies that any ray from a point in its interior crosses ∂V exactly once. The last assertion is just a restatement of the definition that the ith safe region consists of points in $V \cap S_i$.

The Ellipse Probability

With all the tools in place, we can now mimic the work of §2 to get the probability that the skewed ellipse $e_\eta(t)$ is struck by one of n cuts. First we get the total viable area in the upper half of the unit disk, which is just $2\,A(b_{\max})$, with A from Corollary 2. Then, for each $i = 1, 2, \ldots, n$, we need to find the area of the ith safe region. That can be done by just mimicking the approach of §2. First, more notation: the point z of Figure 3 becomes z_i, which can be computed as $r_i \cot(\pi/n) + r_{i+1} \csc(\pi/n)$.

Instead of the angle(s) ρ shown in Fig. 3, we need the values of b_i^+ and b_i^- which determine the points on ∂V where the tropics intersect that curve. That is, given r, m, η, n, and i, we seek the point on C_i^+ that is on ∂V. Assume first that $i = 1$, so we seek the point on ∂V having height r_1. Since $\boldsymbol{p}$ gives V for an unskewed ellipse, the point b we seek is the solution of $R_\eta \cdot \boldsymbol{p}(b) = (x, r_1)$; this equation, if we use c and s to abbreviate $\cos b$, $\sin b$, respectively, and q for $1 \big/ \sqrt{r^2\,c^2 + m^2\,s^2}$, becomes

$$(\cos\eta)\, s\,(1 - m^2\,q) + (\sin\eta)\, c\,(1 - r^2\,q) \;=\; r_1\ .$$

There is no reason to expect this equation to have a simple solution for b, which occurs in both c and s, but in some sense it does. One could feed the equation directly to *Mathematica*'s `Solve` command but it is much faster to use `Eliminate`. First we rewrite it as the following denominator-free system.

$$s^2 + c^2 = 1$$
$$q^2(r^2\,c^2 + m^2\,s^2) = 1$$
$$(\cos\eta)\, s\,(1 - q\,m^2) + (\sin\eta)\, c\,(1 - q\,r^2) = r_1$$

Then using `Eliminate` to eliminate s and q returns the following complicated, but by modern standards not monstrous, equation. At the beginning of our investigation we used numerical techniques to solve the equation $R_\eta \cdot \boldsymbol{p}(b) = r_1$ for b and it came as a pleasant surprise that it can be done by pure algebra. Here we use $c_{j\,\eta}$ to abbreviate $\cos(j\,\eta)$, and similarly for $\sin(j\,\eta)$.

$$
\begin{aligned}
&(m^2 - r^2)^2 c^8 - 4(m^2 - r^2)^2 r_1 s_\eta c^7 + \frac{1}{2}(m^2 - r^2) \\
&\quad \left(m^4 + 2r^2 m^2 + 8r_1^2 m^2 - 6m^2 + r^4 + 2r^2 - 8r^2 r_1^2 + (m^2 - r^2)^2 c_{4\eta} + 2(m^2 - r^2) c_{2\eta} (m^2 + r^2 - 2r_1^2 - 1)\right) c^6 + \\
&\quad 2(m^2 - r^2) r_1 (m^4 + (-2r^2 - 2r_1^2 + 5) m^2 + (m^4 - 2r^2 m^2 + m^2 + r^4 - r^2) c_{2\eta} - r^2 (r^2 - 2r_1^2 + 1)) s_\eta c^5 + \\
&\quad (s_\eta^4 r^8 - r_1^2 r^6 + r_1^4 r^4 - 2m^2 s_\eta^4 r^4 + m^2 r_1^2 r^4 - r_1^2 r^4 - 2m^2 r_1^4 r^2 + m^4 r_1^2 r^2 + 10m^2 r_1^2 r^2 + \\
&\qquad (m^8 - 6m^6 + (4r^2 + 6) m^4 - 6r^2 m^2 + r^4) c_\eta^4 + m^4 r_1^4 + m^4 s_\eta^4 - m^6 r_1^2 - 9m^4 r_1^2 - \\
&\qquad (m^6 - (r^2 + 3) m^4 + r^2 (r^2 + 2) m^2 - r^6 + r^4) c_{2\eta} r_1^2 + 2(2m^6 + (r^4 - 9r^2 + 3) m^4 + (6r^4 - 2r^2) m^2 - r^6) c_\eta^2 s_\eta^2) c^4 - \\
&\quad 2m^2 r_1 (2m^4 + (-5r^2 - 4r_1^2 + 4) m^2 + (2m^4 + (2 - 5r^2) m^2 + r^2 (3r^2 - 2)) c_{2\eta} + r^2 (r^2 + 4r_1^2 - 2)) s_\eta c^3 - \\
&\quad m^2 (2(m^2 - 1)(m^4 - 2m^2 + r^2) c_\eta^4 + 2(m^4 + (r^4 - 4r^2 + 1) m^2 + r^4) s_\eta^2 c_\eta^2 + \\
&\qquad r_1^2 (-2m^4 + (r^2 + 2r_1^2 - 6) m^2 + (-2m^4 + r^2 m^2 - r^4 + 2r^2) c_{2\eta} + r^2 (r^2 - 2r_1^2 + 2))) c^2 + \\
&\quad (r_1 s_\eta m^6 + r_1 s_{3\eta} m^6 - 4r_1^3 s_\eta m^4 - 2r^2 r_1 s_\eta m^4 + r_1 s_\eta m^4 - 2r^2 r_1 s_{3\eta} m^4 + r_1 s_{3\eta} m^4) c + \\
&\quad m^8 c_\eta^4 - 2m^6 c_\eta^4 + m^4 c_\eta^4 + m^4 r_1^4 - m^6 r_1^2 - m^4 r_1^2 - m^6 c_{2\eta} r_1^2 - m^4 c_{2\eta} r_1^2 = 0
\end{aligned}
$$

All symbols in the preceding equation are parameters except for c, which is $\cos b$. The expression is a degree-8 polynomial equation in c so we cannot solve it by algebra. But neither do we need general root-finding, since there are well-studied algorithms that find all the roots of a polynomial. In *Mathematica* the `NSolve` function implements such an algorithm. So our algorithm will just use `NSolve`, delete any imaginary roots, and select the real root that is numerically closest to giving the desired y-coordinate of r_1. There is one slight problem in that when $m = 0$ the formula works out to something quite different. We will not go into the details because this case, where the ellipse is just a line segment, is given a totally algebraic solution in §5.

This formula assumed $i = 1$, but the case of other i-values reduces to this case after a clockwise rotation of $i\pi/n$. Figure 9 shows how it all works, with the intersection points added and the safe regions shown in gray.

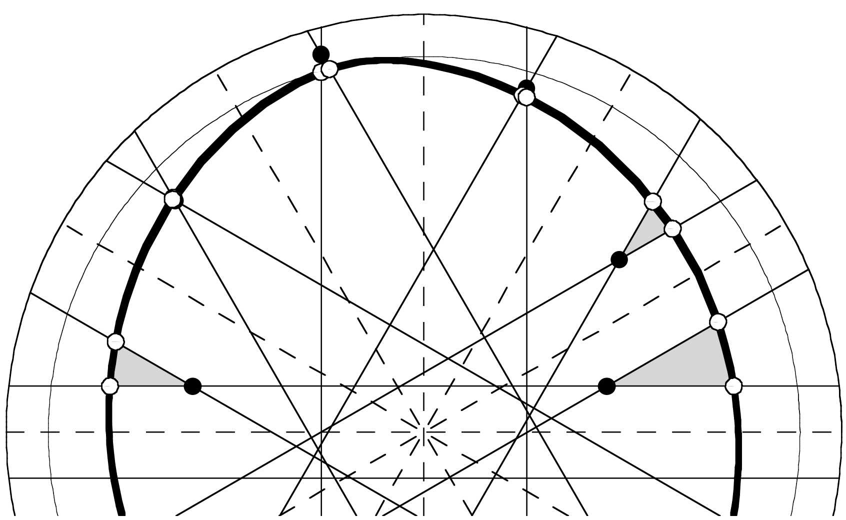


Figure 9 The situation for the ellipse $e_{0.2}(0.25,\ 0.1)$ and 6 cuts. The safe regions are shown in gray, the viable boundary is the thick curve, and the cuts are the dashed lines. The black dots are the points M_i (two of them are outside V) and the white ones are the intersection of the rays from M_i with ∂V. There are three visible safe regions, as the third and fourth are empty and the fifth is too small to see.

So we can now formulate an algorithm for computing the probability that a cut strikes an ellipse. In §5 we will describe a purely algebraic method for handling the case that $m = 0$, which corresponds to an ellipse that degenerates to a line segment. Thus we assume that $m > 0$ in the following algorithm. The algorithm that follows also assumes that there are at least two cuts. When there is one cut the geometry is a little different and there is no problem in formulating a similar algorithm, but we omit the details.

Algorithm for computing $P_{\text{ellipse}}(r, m, \eta, n)$, where $n \geq 2$ and $m > 0$

1. Determine $b_{\max}$ so that ∂V is given by $\boldsymbol{p}(b)$ over $[-b_{\max}, b_{\max}]$ and $[\pi - b_{\max}, \pi + b_{\max}]$. If $m \geq r^2$, $b_{\max} = \pi/2$, else $b_{\max} = \arccos\sqrt{(r^4 - m^2)/(r^2 - m^2)}$ (see proof of Theorem 3(b)).

2. Use Corollary 2 to compute the total viable area in the upper half-circle: it is $2\,A(b_{\max})$.

3. Compute the half-widths r_i of the unsafe strips using Proposition 1, the lengths z_i as defined earlier, and the coordinates of the points M_i using the polar form: length is $\sqrt{r_i^2 + z_i^2}$; angle is $\arctan(r_i / z_i)$

4. For each $i = 1, \ldots, n$, determine whether the safe region inside V_i has area 0 as follows. If $|M_i| \geq 1 - m$ then the safe area is 0. Otherwise, let $M_i' = R_{-\eta}(M_i)$; if the y-coordinate of M_i' is greater than or equal to $\sqrt{(1 - r^2)(r^2 - m^2)} \big/ r$, which is the greatest height of V in the case $m < r^2$, then the safe area is 0. Finally, if $\delta \leq 0$, δ as in Theorem 1(c) with $(x, y) = M_i$, then the safe area is 0. These three conditions capture all possibilities for the safe area to be 0 since the height test eliminates the possibility that δ can be negative but the point lie outside V (see Fig. 6).

5. For those i for which the safe area in V_i is nonzero, compute the area by first finding the b-values by solving the degree-8 polynomial for $\cos b$. Take the inverse cosine of the roots and retain only the real values. Add in the negatives to account for negative b, which can occur. Delete any values that are not in $[-b_{\max}, b_{\max}]$ or $[\pi - b_{\max}, \pi + b_{\max}]$. From these, choose the correct one by selecting the value so that $\boldsymbol{p}(b)$ is in the first quadrant w.r.t. M_i'. There can be extraneous roots so one last step is needed: choose the b-value so that $\boldsymbol{p}(b)$ is the best match to the desired equation $\boldsymbol{p}(b) = y$. Then use Corollary 2 to determine the sector area and subtract the area of two triangles whose heights are r_i and r_{i+1} and whose bases are obtainable from M_i and the ∂V points on the sector.

6. The probability is obtained by dividing the total safe area by the total viable area.

This algorithm is quite fast, taking only a fraction of a second to compute a value of $P_{\text{ellipse}}(r, m, \eta, n)$. Example: $P_{\text{ellipse}}(0.25, 0.1, 0.2, 6) = 0.96$ takes a tenth of a second

This algorithm is not pure algebra, but it is close, as numerical methods occur only in the calls to the elliptic integral and the determination of the roots of a polynomial. It is possible to do formulate a purely numerical algorithm based on root-finding and numerical integration, and such work yields a valuable check on what is, regardless of the method, a somewhat complicated algorithm.

4. Random Coins

In order to understand coins that are randomly placed in a cake, we must clarify the concept of a random orientation. It is natural to assume that the orientation is purely random among all possible orientations, and that is what we shall do even though, in reality, this is not exactly so. First, if the cake has finite height, then a coin near the top or bottom is more likely to be horizontal than vertical, since a vertical coin would not be entirely within the cake. Similarly, a coin near the outer edge of the cake is more likely to have a vertical orientation than a horizontal one. The first case can be taken care of mathematically by assuming the cake to be a cylinder of infinite height in both directions. The second admits no easy solution. So we simply live with the fact that our assumption of purely random orientations might not conform exactly to physical reality, but is of little importance when the coin is small relative to the cake.

We describe an orientation of the coin by means of the upward pointing vector of length r through the center of the coin and perpendicular to it. The tip Q of the vector lies in the northern hemisphere of the sphere concentric with the coin. Consider the orienting point Q in spherical coordinates (ϕ, θ) where ϕ, the latitude, is the angle made with the vertical axis of the sphere $(0 \le \phi \le \pi/2)$ and θ is the longitude $(0 \le \theta < 2\pi)$. Moreover, m, the semiminor axis of the ellipse that is a horizontal projection of the coin, equals $r\cos\phi$. See Figure 10.

Intuitively, we say that an orientation is random if the orienting point Q is equally likely to be anywhere in the northern hemisphere. Such an approach is inadequate since there are an infinite number of points in the northern hemisphere. The way out of this difficulty is to define randomness to mean that the probability that Q lies in a nice subset of the hemisphere is proportional to the area of the subset. Since the area of the northern hemisphere is $2\pi r^2$ and has total probability is 1, for a nice region E we have

$$\text{Probability of } Q \text{ in } E = \frac{\text{Area of } E}{2\pi r^2} = \frac{1}{2\pi}\int\!\!\int_E \sin\phi \, d\phi \, d\theta.$$

(We here used the spherical coordinate formula $d\,\text{Area} = r^2 \sin\phi \, d\phi \, d\theta$ for area on a sphere of radius r.) Thus the joint probability density function (pdf) of the distribution of the angles ϕ and θ is $\sin\phi/(2\pi)$ for $0 \le \phi \le \frac{\pi}{2}$, $0 \le \theta < 2\pi$. Since the pdf is a product of density functions for the two angles, $\sin\phi = f(\theta)\, g(\phi)$ with $f(\theta) = 1/(2\pi)$ and $g(\phi) = \sin\phi$, we can conclude that the two angles are distributed independently.

Notice particularly that for a random point on the sphere the distribution of the latitude angle ϕ is not uniform. This is because there is much more surface area near the equator than at the north pole. The fact that more humans live hear the equator than near the arctic circle is due to geometry as much as geography.

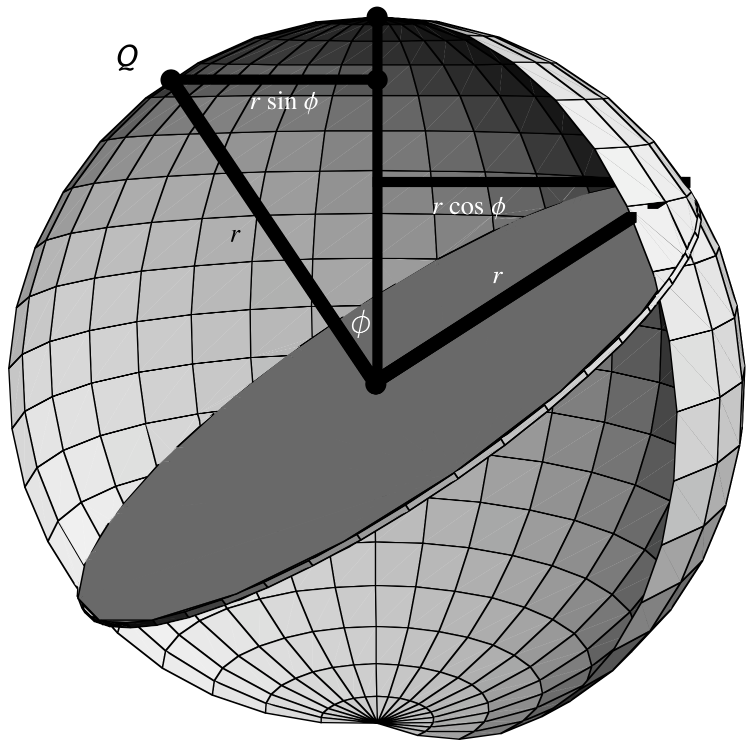


Figure 10 A disk in a sphere. The randomness of the orientation is governed by the location of the point on the sphere that the diameter normal to the disk strikes.

The cumulative distribution function (cdf) of ϕ is given by

$$P(\phi < \phi_0) = \frac{\text{Area of spherical cap } \phi < \phi_0}{\text{Area of hemisphere}} = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\phi_0} \sin\phi \, d\phi \, d\theta = 1 - \cos\phi_0 \, .$$

and thus $P(\phi > \phi_0) = \cos\phi_0$. This is essentially the classic result of Archimedes that the surface area of a spherical cap is proportional to its height.

The cdf of the semiminor axis $m = r\cos\phi$ is now easily computed. For $m_0 \le r$ we have

$$P(m < m_0) = P(r\cos\phi < m_0) = P(\phi > \arccos\tfrac{m_0}{r}) = \cos(\arccos\tfrac{m_0}{r}) = \tfrac{m_0}{r}.$$

Taking the derivative with respect to m_0 gives the pdf of m, namely the constant $1/r$, $0 \le m \le r$.

This gives the pleasing and somewhat surprising result that the semiminor axis is *uniformly distributed* from 0 to r and therefore the final probability that the knife strikes the tilted coin is simply the following double integral of the ellipse probability.

$$\frac{2}{\pi r} \int_0^r \int_0^{\pi/2} P_{\text{ellipse}}(r, m, \eta, n) \, d\eta \, dm.$$

Mathematica's `NIntegrate` can handle this two-dimensional integral. Figure 11 shows the results of 24 such double integrations where r is 1/16. In particular, the final probability when $n = 4$ is

0.247, well under the flat-coin probability of 0.312. For the case of the cake and quarter shown in Figure 1, a hit is just a little less likely than a miss because $r = \frac{12}{108-12} = \frac{1}{8}$ and $n = 4$, and the strike probability works out to 0.48,

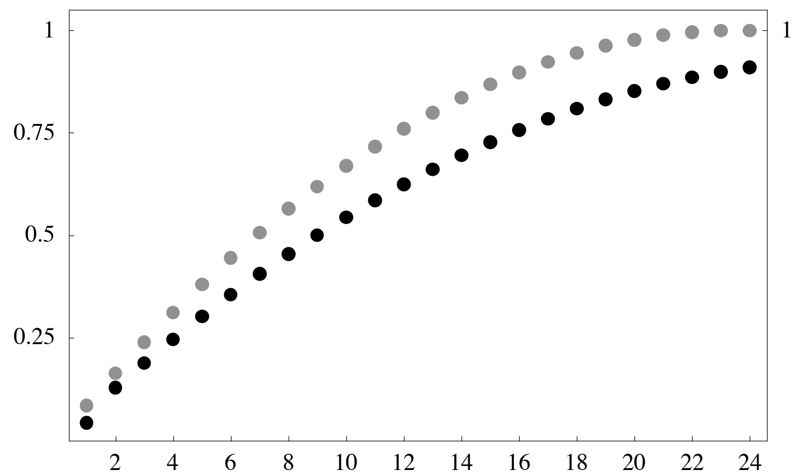


Figure 11 The black dots are the probability of the knife striking a coin of radius 1 / 16 that of the cake, as n goes from 1 to 24. The gray dots show the same probability when the coin is assumed to be horizontal.

5. Coda à la Buffon

A nice special case of the St. Basil problem is to compute the probability of finding a needle hidden in a cake, since a needle is an ellipse with vanishing minor axis. There are two versions of this problem, depending on whether one works in two dimensions or three.

Consider first the planar case: a needle (segment) of half-length L is placed randomly inside a disk of unit radius. What is the probability $P_{\text{needle}}(L, n)$ that one of n equally spaced diameter cuts will strike the needle? While numerical work is necessary to average over all orientations of the needle with respect to the cuts, the problem for a fixed orientation can be completely solved algebraically because the viable region is quite simple. We present here a summary of the method.

Let us begin with the case of fixed orientation, so we seek $P_{\text{needle}}(L, \eta, n)$. As usual let η be the angle the needle makes with the horizontal cut. The boundary of the viable region consists of two arcs that are translations of arcs of the unit circle by distance L in the directions parallel to the orientation. Thus the viable region is the intersection of two circular discs of radius 1 with centers separated by distance $2\,L$ (Fig. 12). The area of the viable region is $A(L) = 2\left(\arccos L - L\sqrt{1-L^2}\right)$, independent of the needle's orientation.

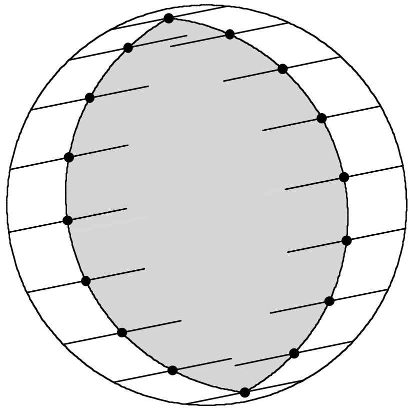


Figure 12 The shaded region is where the center of a tilted segment of length 0.6 can lie so that the segment is inside the unit disk; it is bounded by two circular arcs.

The safe region is a union of wedges each bounded by a circular arc of the viable boundary and two segments sharing an endpoint at one of the points M_i at the intersection of two tropics C_i^+ and C_{i+1}^-. A calculus exercise then produces the appropriate area formula.

Lemma 2. Let Area(p, q, α, β), for $p^2 + q^2 \le 1$ and $0 \le \beta - \alpha \le 2\pi$ be the area bounded by the circular arc from $A = (\cos\alpha, \sin\alpha)$ to $B = (\cos\beta, \sin\beta)$ and the segments from (p, q) to A and B. Then

$$\text{Area}(p, q, \alpha, \beta) = \frac{1}{2}\left(\beta - \alpha - p(\sin\beta - \sin\alpha) + q(\cos\beta - \cos\alpha)\right)$$

In the application of the preceding area formula, the point (p, q) is the intersection of two lines (tropics) and the angles α and β depend on the intersection of a line and a circle, so all parameters are easily determined algebraically.

If $n = 1$ the geometry is very simple and one gets the following formula for $P_{\text{needle}}(L, \eta, 1)$, the probability that a needle of half-length L and tilt angle η strikes the equatorial diameter:

$$\begin{cases} \dfrac{\arcsin(2L\sin\eta) + 2L\sin\eta\left(\sqrt{1-4L^2\sin^2\eta} - 2L\cos\eta\right)}{2\left(\arccos L - L\sqrt{1-L^2}\right)}, & \text{if } L < \cos\eta \\ 1 - \dfrac{\arccos(2L\sin\eta) - 2L\sin\eta\sqrt{1-(2L\sin\eta)^2}}{\arccos L - L\sqrt{1-L^2}}, & \text{if } \cos\eta \le L \le \min\left(\csc\eta, \sqrt{2}\right)/2 \\ 1, & \text{otherwise} \end{cases}$$

When $n > 1$ a case analysis is needed to get the precise algorithm, since different configurations can arise. One of the safe sectors (in upper half-plane) might contain a vertex of the viable curve, and its area computation is then done in two parts. One of the two tropics that bound a given safe sector might pass through the center of a bounding arc, in which case the area formulas are different than when the center of the bounding arc is on neither tropic. We omit the precise details, but the bottom line is that one can formulate a completely algebraic algorithm that is very fast when using approximate real numbers and that, when the input is symbolic, computes a symbolic answer. As an example, we have the following formula.

$$P_{\text{needle}}\left(\frac{1}{10}, 0, 4\right) = \frac{100\left(\operatorname{arccsc} 5 - \operatorname{arccsc} 10 + \operatorname{arccsc}\sqrt{50}\right) + 8\sqrt{6} - 3\sqrt{11} + 8}{100\operatorname{arcsec} 10 - 3\sqrt{11}}$$

Now consider the planar random orientation problem. First note that the distribution of η, the angle the needle makes with the horizontal cut, is uniform; this is because even though there are boundary effects that restrict the orientations when the center of the needle is too close to the edge, they are perfectly symmetric under rotation. Moreover, given L, the area of the viable region is independent of η. Thus the probability that a cut strikes the needle can be obtained by averaging the probabilities for each η, yielding:

$$P_{\text{needle}}(L, n) = \frac{2}{\pi}\int_0^{\pi/2} P_{\text{needle}}(L, \eta, n)\, d\eta = \frac{2n}{\pi}\int_0^{\pi/(2n)} P_{\text{needle}}(L, \eta, n)\, d\eta$$

Though the integrand is quite explicit, the integral must in general be done numerically. We have implemented this combined algebraic/numerical approach and the results agree with our purely numerical algorithm, which is comforting.

The 3-dimensional needle problem requires dealing with the fact that the half-length of the projected needle will vary between 0 and L. Moreover, as in §4, we need to make the assumptions that the cake is unbounded in height and that all orientations of the needle in 3-space are equally likely. But then the probability can be worked out. Referring to Figure 10, we see that the projected half-length of the needle — for which we use λ — is $L \sin\phi$. Here ϕ is the latitude of the point on the sphere pierced by the needle. Then using the latitude distribution from §4:

$$P(L\sin\phi > \lambda_0) = P\left(\phi > \arcsin\frac{\lambda_0}{L}\right) = \cos\left(\arcsin\frac{\lambda_0}{L}\right) = \sqrt{1 - \left(\frac{\lambda_0}{L}\right)^2} = \frac{1}{L}\sqrt{L^2 - \lambda_0^2}\,.$$

Thus the cumulative distribution of λ is given by $P(\lambda < \lambda_0) = 1 - \sqrt{L^2 - \lambda_0^2}\,\big/ L$. Because λ's density function is the derivative of this, the probability of hitting the needle is the following, where the substitution $u = \sqrt{L^2 - \lambda^2}$ (then replacing u by λ) is used to simplify the integral and also symmetry is used to reduce the domain of η-integration to $\pi/(2n)$.

$$\frac{2n}{\pi}\int_0^L \int_0^{\frac{\pi}{2n}} P_{\text{needle}}(\lambda, \eta, n)\, \frac{\lambda}{L\sqrt{L^2 - \lambda^2}}\, d\eta\, d\lambda = \frac{2n}{\pi L}\int_0^L \int_0^{\frac{\pi}{2n}} P_{\text{needle}}\left(\sqrt{L^2 - \lambda^2}, \eta, n\right) d\eta\, d\lambda$$

For example, the probability of hitting a needle of full length $1/4$ using 7 cuts is 0.498.

Figure 13 contains a top view of a cake with 200 needles of length $1/4$ having random space orientations embedded into it. The hits, which were computed by actually checking for intersections of the projections with the seven cuts, are marked by thicker lines. It takes only 17 seconds to compute the double integral above and get 0.493. The hit ratio in this example is 0.47, consistent with the prediction. An experiment with one million needles saw 487706 hits.

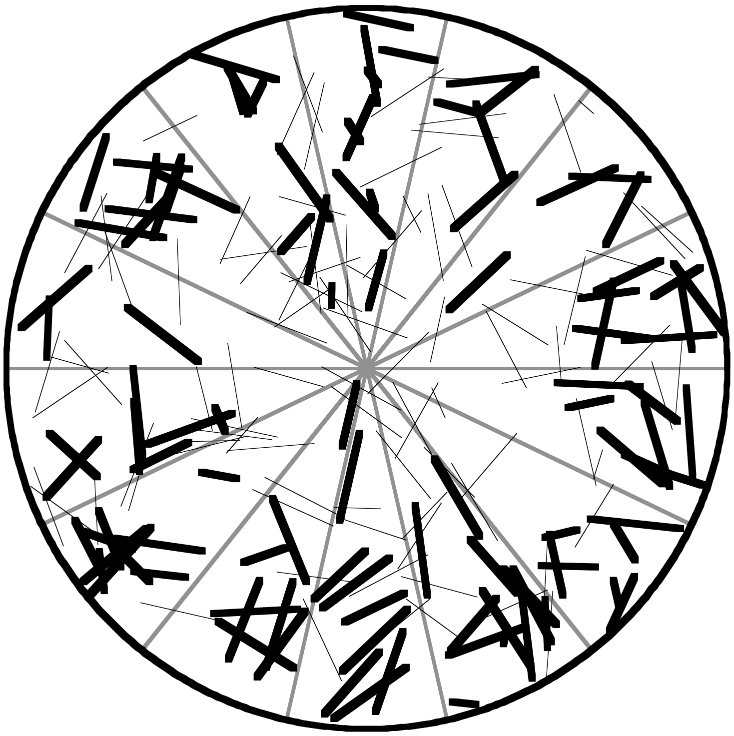


Figure 13 Two hundred random needles of length $1/4$ in a cake of unit radius. Think of this as a top view of a cake containing many copies of the needle in random orientation. In this example, the proportion of needles struck by one of the seven cuts (the thicker segments) was 0.47. The true probability is 0.493.

6. Conclusion

Without baking a lot of St. Basil's cakes, one has very little intuition about the probabilities in question, but the modeling done here, though imperfect in some fine detail, gives us good approximations to the actual probabilities. This project is a tremendous example of the power of numerical and symbolic algorithms. We first obtained the probabilities using purely numerical algorithms, calling various numerical integrations and root-finding algorithms thousands of times. But we later found many algebraic simplifications, which is what we presented here, and though some of the expressions are complicated, their use yielded algorithms many times faster than the purely numerical approach.

Here are some open questions related to this problem.

(1) Assume the cake has height h. Work out the probabilities where h is taken into account when computing the distribution of orientations, always under the assumption that the coin is entirely within the cake. This means that if the center of the coin is near the upper or lower border of the cake, then the set of possible orientations is restricted, since the latitude of the point on the sphere in Figure 10 is restricted. Closely related to this is the problem of resolving the slight difficulty at the border of the cake, where not every orientation can occur.

(2) Suppose a circle of radius r lies inside a unit disk. If one can make n straight cuts in the disk, how should those cuts be made to maximize the chance of striking the circle? If, say, $r = 1/16$ and $n = 8$, and the cuts are equispaced cake cuts, then, by the formula of §2, the probability of a hit is 0.56. But one

can do much better by just using parallel horizontal cuts at y-values $\pm 7\, r$, $\pm 5\, r$, $\pm 3\, r$, $\pm r$, for which the probability of a hit is 0.645. But we do not see how to prove that this is the best choice. If n is odd, the best choice appears to be to use cuts at heights 0, $\pm 2\, r$, $\pm 4\, r$, and so on.

As pointed out to us by Bill Briggs, this last question is related to the real-world problem of diagnosing prostate cancer. Both the gland and a cancerous tumor can be modeled by ellipsoids. A simplified version of the problem is: Suppose an ellipsoid G is given and it contains a smaller ellipsoid E whose center and dimensions are randomly distributed on some given intervals so that E is inside G. Suppose we can make m one-dimensional piercings of G. How should those lines be chosen to maximize the probability of striking E? In reality, there would be several disjoint several cancerous ellipsoids within the gland. One might start by assuming that the tumors are spherical.

Mathematica code for computing the probabilities discussed here is available from the authors.

Acknowledgement. We are grateful to Adam Strzebonski of Wolfram Research, Inc., for helpful advice regarding symbolic algebra.

References

1. C. Savvidou, The St. Basil's cake problem, *Mathematics Magazine* **78** (2005) 48–51.

2. G. Salmon, *A Treatise on Conic Sections*, 6th ed., Chelsea, New York, 1960.

3. http://mathworld.wolfram.com/EllipticIntegraloftheSecondKind.html

4. http://www.stbasil.goarch.org/about/vasilopita.asp

5. http://www.catholicculture.org/lit/activities/view.cfm?id=993