NOTES

Edited by **Ed Scheinerman**

Four Shots for a Convex Quadrilateral

Christian Blatter

1. INTRODUCTION. Problem A6 of the 2006 Putnam competition read as follows [7]: "Four points are chosen uniformly and independently at random in the interior of a given circle. Find the probability that they are the vertices of a convex quadrilateral." It seems that only one of the contestants obtained the correct answer $1 - \frac{35}{12\pi^2} \doteq 0.70448$.

A6 is an instance of Sylvester's famous *four points problem* [6]: What is the probability of the event K that four points taken at random in the plane form the vertices of a convex quadrilateral? (To be precise, Sylvester asked for the complementary probability of a "re-entrant" quadrilateral.) This problem has an interesting history [4] which is intimately tied to the development of the notion of probability measure. In fact, Sylvester's problem, together with Buffon's needle problem, is one of the prime paradigms in geometric probability theory. In 1917 Blaschke [1], [2, §§24–25] proved that the probability of K is maximal (namely $0.70448\ldots$) when the shape from which the points are chosen is an ellipse and minimal (namely $\frac{2}{3}$) when it is a triangle. Blaschke's results refer to convex shapes only; for nonconvex shapes the situation becomes combinatorially more challenging, see [5].

In geometric probability theory the random objects (points, lines, etc.) are usually assumed to be uniformly distributed with respect to a natural geometric measure (angle, area, etc.), so the relevant probabilities are directly proportional to geometric quantities like area.

Let us now imagine a marksman firing four shots at the origin of the (x, y)-plane. Under these circumstances the four bullet holes will not be uniformly distributed on a circle or a square. It is much more natural to model their distribution by a bivariate Gaussian whose parameters depend on the markman's skill, the quality of his rifle and maybe other factors. So we pose Sylvester's question under these new assumptions: What is the probability that the four bullet holes form the vertices of a convex quadrilateral? In this note we shall prove the following result:

Theorem. *Let four points be chosen independently and distributed according to the same normal distribution in the plane. Then the probability of the event K that they form the vertices of a convex quadrilateral is given by*

$$P[K] = \frac{6}{\pi} \arcsin\left(\frac{1}{3}\right) \doteq 0.64904.$$

Note that in this setup the probability of K is even smaller than in the case of four points uniformly distributed in a triangle. One of the referees has pointed out that our theorem is not new: Hiroshi Maehara arrived at the same result already in 1978 [3]. His paper, apart from being in Japanese, treats the geometrical combinatorics in a manner different from ours and is more demanding in terms of abstract probability. In the light of these circumstances it seems profitable to take another look at the problem.

2. OUTLINE OF THE PROOF. Let $A_i = (x_i, y_i)$ $(1 \le i \le 4)$ be the four random points. The first two points A_1, A_2 determine a line g_{12}, and the remaining two points A_3, A_4 determine a line g_{34}. With probability 1 the two lines are well defined and intersect in a point S.

As can be seen from Figure 1, the four points are the corners of a convex quadrilateral if and only if one of the following is true:

(a) The point S separates the pair $\{A_1, A_2\}$ as well as the pair $\{A_3, A_4\}$,

(b) The point S separates neither the pair $\{A_1, A_2\}$ nor the pair $\{A_3, A_4\}$.

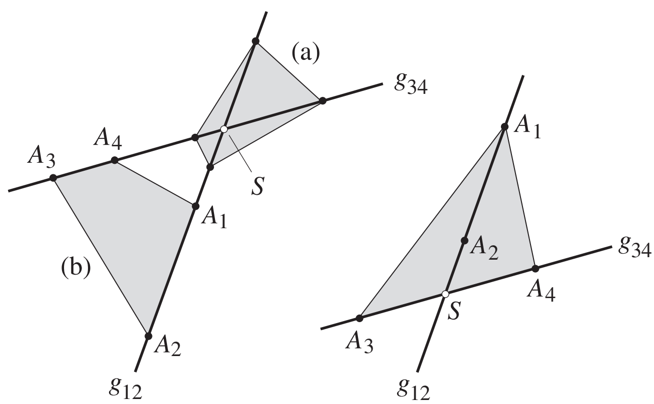

Figure 1.

In terms of probabilities, we shall proceed as follows:

In Section 3 we consider only the two points A_1, A_2, and the line g_{12} joining them. We compute the distribution of g_{12} among all lines in $\mathbb{R}^2$, as well as the distribution of A_1 and A_2 on g_{12}. Using the latter we can determine the probability $p(v_*)$ that a point S at given position v_* on g_{12} separates A_1 and A_2. The resulting expressions will serve for the second pair $\{A_3, A_4\}$ as well.

In Section 4 we describe the position of S on g_{12} and g_{34} by the relevant parameters, among them the angle θ between the two lines. This will allow us to compute the probability that one of the above cases (a) or (b) occurs, given g_{12} and g_{34}.

The proof ends in Section 5 with a "grand" integration, using the distributions of g_{12} and g_{34} determined in Section 3. Fortunately the computations can be arranged in such a way that no proper integrals $\int_a^b e^{-x^2/2}\,dx$ appear in the process.

3. TWO POINTS. Given any measure μ we denote the corresponding differential by $\mathrm{d}\mu$. If $x_1, \dots, x_r$ are real variables then $\mathrm{d}(x_1, \dots, x_r)$ denotes this differential with respect to Lebesgue measure in the product space of the x_i. Instead of $\mathrm{d}(x)$ we write $\mathrm{d}x$.

If two points A_1 and A_2 are chosen independently and normally distributed in the (x, y)-plane, then their common probability distribution on the probability space $\Omega = \mathbb{R}^2 \times \mathbb{R}^2$ is

$$\mathrm{d}\mu = \frac{1}{4\pi^2} e^{-(x_1^2+y_1^2+x_2^2+y_2^2)/2}\,\mathrm{d}(x_1, y_1, x_2, y_2).$$

With no loss of generality we have assumed the distribution to be rotationally symmetric with standard deviation equal to $\sqrt{2}$.

The two points determine the angle $\phi := \arg(x_2 - x_1, y_2 - y_1) - \frac{\pi}{2}$. With the help of ϕ we now introduce in the (x, y)-plane new coordinates (u, v) as follows (Figure 2):

$$x = u\cos\phi - v\sin\phi,$$
$$y = u\sin\phi + v\cos\phi.$$

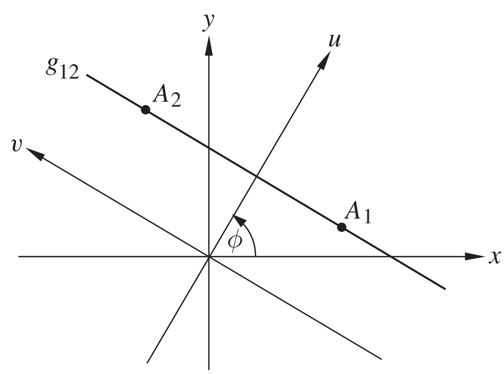

Figure 2.

In this way the line g_{12} becomes a line $u = \text{constant}$, $-\infty < v < \infty$, going vertically upwards in the (u, v)-plane. The two points A_1, A_2 now have coordinates (u, v_1), (u, v_2), where $v_2 > v_1$. One should view u and ϕ as "coordinates" of g_{12}, whereas v_1 and v_2 describe the positions of the points A_1 and A_2 on g_{12}.

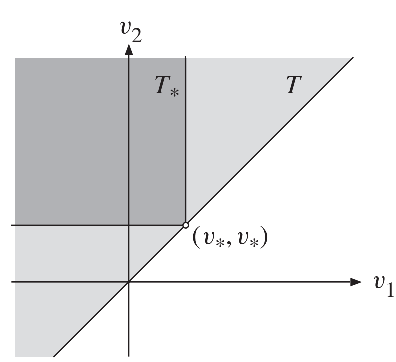

Figure 3.

Let $T := \{(v_1, v_2) \mid -\infty < v_1 < v_2 < \infty\}$ be the region above the line $v_2 = v_1$ in an abstract (v_1, v_2)-plane (Figure 3). The formulae

$$\left.\begin{aligned} x_i &= u\cos\phi - v_i\sin\phi \\ y_i &= u\sin\phi + v_i\cos\phi \end{aligned}\right\} \qquad (i = 1, 2)$$

realize a parametric representation of Ω by means of the new variables u, ϕ, v_1, v_2. It is easy to see that this representation maps the set

$$\tilde{\Omega} := \{(u, \phi, v_1, v_2) \mid -\infty < u < \infty,\ \phi \in \mathbb{R}/(2\pi),\ (v_1, v_2) \in T\}$$

bijectively onto $\Omega \setminus \{(A_1, A_2) \mid A_1 = A_2\}$. One has

$$x_i^2 + y_i^2 = u^2 + v_i^2 \qquad (i = 1, 2);$$

furthermore, the Jacobian computes to

$$\left|\frac{\partial(x_1, y_1, x_2, y_2)}{\partial(u, \phi, v_1, v_2)}\right| = \begin{vmatrix} \cos\phi & -u\sin\phi - v_1\cos\phi & -\sin\phi & 0 \\ \sin\phi & u\cos\phi - v_1\sin\phi & \cos\phi & 0 \\ \cos\phi & -u\sin\phi - v_2\cos\phi & 0 & -\sin\phi \\ \sin\phi & u\cos\phi - v_2\sin\phi & 0 & \cos\phi \end{vmatrix} = v_2 - v_1.$$

As a consequence we get the following expression for $\mathrm{d}\mu$ on $\tilde{\Omega}$:

$$\mathrm{d}\tilde{\mu} = \frac{e^{-u^2}}{\sqrt{\pi}}\,\mathrm{d}u \cdot \frac{\mathrm{d}\phi}{2\pi} \cdot \frac{1}{\sqrt{4\pi}}(v_2 - v_1)e^{-(v_1^2+v_2^2)/2}\,\mathrm{d}(v_1, v_2).$$

This shows in particular that the new random variables U, Φ, and (V_1, V_2) are independent. The two factors

$$\frac{e^{-u^2}}{\sqrt{\pi}}\,\mathrm{d}u \cdot \frac{\mathrm{d}\phi}{2\pi} \tag{1}$$

describe the distribution of the random line g_{12} among all lines in $\mathbb{R}^2$, whereas the function

$$f(v_1, v_2) := \frac{1}{\sqrt{4\pi}}(v_2 - v_1)e^{-(v_1^2+v_2^2)/2}$$

is nothing else but the probability density of (V_1, V_2) on T. Indeed, one has

$$\int_T f(v_1, v_2)\,\mathrm{d}(v_1, v_2) = 1,$$

for otherwise $\int_{\tilde{\Omega}} \mathrm{d}\tilde{\mu}$ would not compute to 1.

The point S has coordinates (u, v_*), where the actual value v_* depends on the line g_{34} as well. Assume for the moment that v_* is given. As indicated in Section 2 we have to compute the probability $p(v_*)$ that S separates A_1 and A_2, i.e., that $v_1 \le v_* \le v_2$ holds:

$$p(v_*) = \int_{T_*} f(v_1, v_2)\,\mathrm{d}(v_1, v_2) = \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{v_*} \int_{v_*}^{\infty} (v_2 - v_1)e^{-(v_1^2+v_2^2)/2}\,dv_2\,dv_1,$$

where the domain T_* is shown in Figure 3. One has

$$\int_{v_*}^{\infty} v_2 e^{-v_2^2/2}\,dv_2 = e^{-v_*^2/2}, \qquad \int_{-\infty}^{v_*} (-v_1)e^{-v_1^2/2}\,dv_1 = e^{-v_*^2/2}$$

and consequently

$$\begin{aligned} p(v_*) &= \frac{1}{\sqrt{4\pi}}\left(\int_{-\infty}^{v_*} e^{-v_1^2/2}\,dv_1 \cdot e^{-v_*^2/2} + e^{-v_*^2/2} \cdot \int_{v_*}^{\infty} e^{-v_2^2/2}\,dv_2\right) \\ &= \frac{1}{\sqrt{4\pi}} e^{-v_*^2/2} \int_{-\infty}^{\infty} e^{-v^2/2}\,dv = \frac{1}{\sqrt{2}}\,e^{-v_*^2/2}. \end{aligned} \tag{2}$$

4. FOUR POINTS. We argue in the same way in the case of the two points $A_3 = (x_3, y_3)$ and $A_4 = (x_4, y_4)$. Thus the line g_{34} connecting these two points becomes a line $\bar{u} =$ constant in a $(\bar{u}, \bar{v})$ coordinate system, where the $\bar{u}$-axis encloses the (random) angle $\theta = \bar{\phi} - \phi$ with the u-axis constructed in Section 3 (Figure 4). In particular one has

$$u = \bar{u}\cos\theta - \bar{v}\sin\theta, \qquad \bar{u} = u\cos\theta + v\sin\theta. \tag{3}$$

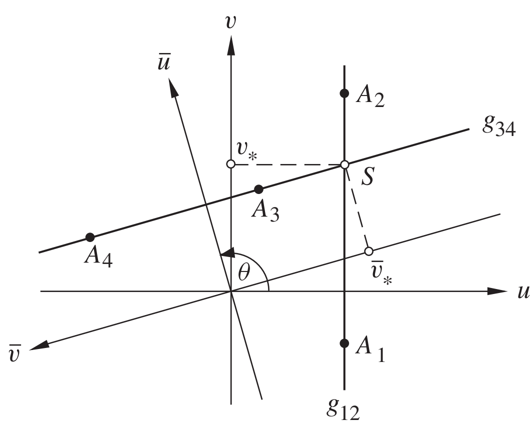

Figure 4.

The coordinates of the point S in the (u, v) and $(\bar{u}, \bar{v})$ coordinate systems are (u, v_*) and $(\bar{u}, \bar{v}_*)$ respectively, where v_* and $\bar{v}_*$ depend on u, $\bar{u}$, and θ only: from (3) we get

$$v_* = \frac{\bar{u} - u\cos\theta}{\sin\theta}, \qquad \bar{v}_* = \frac{\bar{u}\cos\theta - u}{\sin\theta}. \tag{4}$$

From what we have said in Section 2 it follows that the four points A_i $(1 \le i \le 4)$ are the corners of a convex quadrilateral if and only if one of the following is true:

(a) $v_1 < v_* < v_2$, and simultaneously $\bar{v}_3 < \bar{v}_* < \bar{v}_4$,

(b) Neither $v_1 < v_* < v_2$ nor $\bar{v}_3 < \bar{v}_* < \bar{v}_4$ holds.

We conclude that the conditional probability (given u, $\bar{u}$, and θ) for the event K computes to

$$P[K \mid U, \bar{U}, \Theta] = p(v_*)p(\bar{v}_*) + \big(1 - p(v_*)\big)\big(1 - p(\bar{v}_*)\big),$$

and using (2) we obtain

$$P[K \mid U, \bar{U}, \Theta] = 1 + e^{-(v_*^2 + \bar{v}_*^2)/2} - \frac{1}{\sqrt{2}}\big(e^{-v_*^2/2} + e^{-\bar{v}_*^2/2}\big),$$

where v_* and $\bar{v}_*$ are given by (4).

5. INTEGRATION. The distributions of U and $\bar{U}$ are apparent in (1), and the variable $\Theta = \bar{\Phi} - \Phi$ is uniformly distributed mod 2π. Therefore $\mathrm{d}P[K \mid u, \bar{u}, \theta]$ has the following form:

$$dP[K \mid u, \bar{u}, \theta] = \frac{e^{-u^2-\bar{u}^2}}{\pi}\left(1 + e^{-(v_*^2+\bar{v}_*^2)/2} - \frac{1}{\sqrt{2}}\left(e^{-v_*^2/2} + e^{-\bar{v}_*^2/2}\right)\right) d(u, \bar{u})\, \frac{d\theta}{2\pi}. \tag{5}$$

It remains to integrate (5) with respect to u, $\bar{u}$, and θ.

In the exponents of (5) certain positive definite quadratic forms in u and $\bar{u}$ appear whose coefficients depend on θ. Let $Q(u, \bar{u}) := au^2 + 2bu\bar{u} + c\bar{u}^2$ be such a form, and let $\Delta := ac - b^2$ be the discriminant of Q. Then we have the following simple formula:

$$\int_{\mathbb{R}^2} e^{-Q(u,\bar{u})}\, d(u, \bar{u}) = \frac{\pi}{\sqrt{\Delta}}. \tag{6}$$

In order to prove (6), write Q as

$$Q(u, \bar{u}) = a\left(u + \frac{b}{a}\bar{u}\right)^2 + \frac{\Delta}{a}\,\bar{u}^2$$

and do two successive integrations.

We return to (5). The first form we have to consider is

$$\begin{aligned} Q_1(u, \bar{u}) &:= u^2 + \bar{u}^2 + \frac{v_*^2 + \bar{v}_*^2}{2} = u^2 + \bar{u}^2 + \frac{(u^2 + \bar{u}^2)(1 + \cos^2\theta) - 4u\bar{u}\cos\theta}{2\sin^2\theta} \\ &= \frac{(u^2 + \bar{u}^2)(2 + \sin^2\theta) - 4u\bar{u}\cos\theta}{2\sin^2\theta}; \end{aligned}$$

its discriminant is given by

$$\Delta_1 = \frac{(2 + \sin^2\theta)^2 - 4\cos^2\theta}{(2\sin^2\theta)^2} = \frac{8\sin^2\theta + \sin^4\theta}{4\sin^4\theta},$$

and hence

$$\frac{1}{\sqrt{\Delta_1}} = \frac{2\,|\sin\theta|}{\sqrt{9 - \cos^2\theta}}. \tag{7}$$

In a similar way the discriminant of the form

$$\begin{aligned} Q_2(u, \bar{u}) &:= u^2 + \bar{u}^2 + \frac{v_*^2}{2} = u^2 + \bar{u}^2 + \frac{\bar{u}^2 - 2u\bar{u}\cos\theta + u^2\cos^2\theta}{2\sin^2\theta} \\ &= \frac{u^2(1 + \sin^2\theta) - 2u\bar{u}\cos\theta + \bar{u}^2(1 + 2\sin^2\theta)}{2\sin^2\theta} \end{aligned}$$

appearing in (5) computes to

$$\Delta_2 = \frac{(1 + \sin^2\theta)(1 + 2\sin^2\theta) - \cos^2\theta}{(2\sin^2\theta)^2} = \frac{4\sin^2\theta + 2\sin^4\theta}{4\sin^4\theta};$$

hence

$$\frac{1}{\sqrt{\Delta_2}} = \frac{\sqrt{2}\,|\sin\theta|}{\sqrt{3-\cos^2\theta}}. \tag{8}$$

We now can integrate (5) with respect to u and $\bar{u}$. Using (6) we obtain

$$\mathrm{d}P[K \mid \theta] = \left(1 + \frac{1}{\sqrt{\Delta_1}} - 2\frac{1}{\sqrt{2}}\frac{1}{\sqrt{\Delta_2}}\right)\frac{\mathrm{d}\theta}{2\pi}. \tag{9}$$

Since from (7) and (8) we get

$$\frac{1}{2\pi}\int_0^{2\pi}\frac{d\theta}{\sqrt{\Delta_1}} = \frac{4}{\pi}\int_0^{\pi/2}\frac{\sin\theta}{\sqrt{9-\cos^2\theta}}\,d\theta = \frac{4}{\pi}\int_0^{1/3}\frac{dt}{\sqrt{1-t^2}} = \frac{4}{\pi}\arcsin\frac{1}{3}$$

and

$$\frac{1}{2\pi}\int_0^{2\pi}\frac{d\theta}{\sqrt{\Delta_2}} = \frac{2\sqrt{2}}{\pi}\int_0^{\pi/2}\frac{\sin\theta}{\sqrt{3-\cos^2\theta}}\,d\theta = \frac{2\sqrt{2}}{\pi}\int_0^{1/\sqrt{3}}\frac{dt}{\sqrt{1-t^2}}$$
$$= \frac{2\sqrt{2}}{\pi}\arcsin\frac{1}{\sqrt{3}},$$

it follows that by integrating (9) with respect to θ we altogether obtain

$$P[K] = 1 + \frac{4}{\pi}\arcsin\frac{1}{3} - \frac{4}{\pi}\arcsin\frac{1}{\sqrt{3}}.$$

Because of

$$1 - \frac{4}{\pi}\arcsin\frac{1}{\sqrt{3}} = \frac{2}{\pi}\left(\frac{\pi}{2} - 2\arcsin\frac{1}{\sqrt{3}}\right) = \frac{2}{\pi}\arcsin\frac{1}{3}$$

this can be simplified to the stated result.

ACKNOWLEDGMENTS. The author wishes to thank the referees for various suggestions to make this paper more readable. Particular thanks go to the referee who pointed to the reference [**3**].

REFERENCES

1. W. Blaschke, Lösung des "Vierpunktproblems" von Sylvester aus der Theorie der geometrischen Wahrscheinlichkeiten, *Ber. Verh. Sächs. Akad. Wiss. Leipzig Math.-Phys. Kl.* **69** (1917) 436–453.
2. ———, *Vorlesungen über Differentialgeometrie, II. Affine Differentialgeometrie,* Springer-Verlag, Berlin, 1923.
3. H. Maehara, On Sylvester's problem, *Proc. Inst. Stat. Math., Tokyo* **25** (1978) 81–86.
4. R. E. Pfiefer, The historical development of J. J. Sylvester's four point problem, *Math. Mag.* **62** (1989) 309–317.
5. E. R. Scheinerman and H. S. Wilf, The rectilinear crossing number of a complete graph and Sylvester's "Four Point Problem" of geometric probability, this MONTHLY **101** (1994) 939–943.
6. J. J. Sylvester, Question 1491, *The Educational Times*, London, April, 1864.
7. Year 2006 Problems, *William Lowell Putnam Mathematics Competition Directory*, available at `http://www.unl.edu/amc/a-activities/a7-problems/putnamindex.shtml`.

Department of Mathematics, Eidg. Technische Hochschule, CH-8092 Zurich, Switzerland
christian.blatter@math.ethz.ch