NOTES

The Historical Development of J. J. Sylvester's Four Point Problem

RICHARD E. PFIEFER
San Jose State University
San Jose, CA 95192

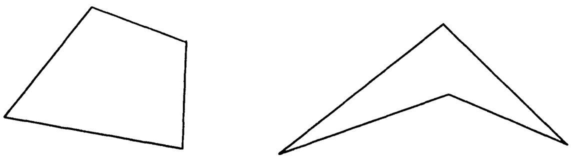
Historically, it would seem that the first question given on local probability, since Buffon, was the remarkable four-point problem of Prof. Sylvester.

M. W. Crofton (1885) [6]

1. The Early History In the *Educational Times* of 1864 [15], question 1491, J. J. Sylvester proposed what became known as his four point problem:

> Show that the chance of four points forming the apices of a reentrant quadrilateral is $1/4$ if they be taken at random in an indefinite plane, but $1/4 + e^2 + x^2$, where e is a finite constant and x a variable quantity, if they be limited by an area of any magnitude and of any form.

The limiting area mentioned above is understood to be a convex region [18].



convex quadrilateral **reentrant quadrilateral**

FIGURE 1.

Since there are two questions stated, we first report the results on the probability, P, that four points taken at random in the plane form a reentrant quadrilateral. The readers of the *Educational Times* set to work on this problem and here is a list of some of the solvers and their published answers [9].

SOLVER	PROBABILITY, P
Cayley and Sylvester	$1/4$
G. C. DeMorgan	$1/2$
J. M. Wilson	$1/3$
C. M. Ingleby	$P < 1/2$
(Name Unknown)	$3/8$
W. S. B. Woolhouse	$35/12\pi^2$

At the time, no one could detect whether any of the probabilities computed above was *the* solution, and J. J. Sylvester concluded "This problem does not admit of a determinate solution" [16].

Cayley and Sylvester "solved" the problem by assuming that A, B, and C are the three points which form the largest triangle, then the fourth point D gives a reentrant quadrilateral only if it falls within the original triangle, $\triangle ABC$, out of the four equal-area triangles as shown in FIGURE 2. Cayley and Sylvester knew that this argument was insufficient in that it was possible, by an equally good argument, to obtain an inconsistent result [18].

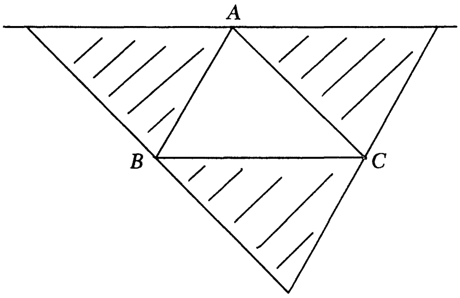


FIGURE 2.

In his solution to Question 1491, W. S. B. Woolhouse decided that he could assume that the four points were contained inside a circle, K, of radius $r > 0$, then compute the probability for four points inside K, and finally take the limit as $r \to \infty$. In other words, he wanted to treat the plane as a circle of infinite radius. (See FIGURE 3.) He noted that:

$$\begin{aligned} P(K) &= \text{Prob (reentrant quadrilateral)} \\ &= 4 \text{ Prob (One point is inside the triangle formed by the other 3 points.)} \\ &= 4 \text{ Mean (Expected) Triangle Area of 3 points/Area of } K \\ &= 4\, M(K)/A(K), \text{ where} \end{aligned}$$

$$M(K) = 1/(A(K))^3 \iint_{P\in K} \iint_{Q\in K} \iint_{R\in K} \frac{1}{2} \left\| \begin{matrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{matrix} \right\| dy_3\, dy_2\, dy_1\, dx_3\, dx_2\, dx_1. \quad (1)$$

This formula defines $M(K)$ for any closed, bounded convex plane region K. When K is a circle of radius r, $A(K) = \pi r^2$.

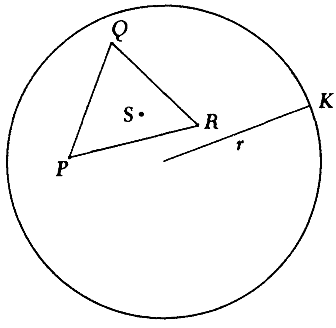


FIGURE 3.

The computation of $M(K)$, for K a circle of radius r, will be left to the interested reader with the hints: use Crofton's Formula and polar coordinates. (See Solomon [**14**] for Crofton's Formula and some related computations.) After a few pages, you will obtain:

$$M(K)=\left(1/(\pi r^2)^3\right)(35/48\pi^2)(\pi r^2)^4, \quad \text{and } P(K)=35/12\pi^2.$$

Noting that $P(K)$ is independent of the radius r, Woolhouse concluded that the solution to Question 1491 was $35/12\pi^2$.

The culprit responsible for these inconsistent results is, of course, the phrase "at random in the plane." From this sample of two of the solutions, we see that each of the solvers above had used his own intuitive interpretation of the phrase and arrived at different answers. In subsequent issues of the *Educational Times*, there was a great deal of spirited discussion of what "at random in the plane" should mean. Crofton [**5**], in 1868, wrote about these differences of opinion and the discordant results in the four-point problem:

> this arises, not from any inherent ambiguity in the subject matter, but from the weakness of the instrument employed; our undisciplined conceptions of a novel subject requiring to be repeatedly and patiently reviewed, tested, and corrected by the light of experience and comparison, before they are purged from all latent error.

The discussion, of course, would not be completely resolved until probability theory in terms of appropriate measures was developed in the following century. For a discussion of the necessity of such a measure in geometric probability see the monograph by Kendall and Moran [**11**, pp. 9–13].

2. The Variational Four-Point Problem Now we concentrate on the computation of the probability, $P(K)$, of the four points forming a re-entrant quadrilateral when they are taken at random inside a closed, bounded, convex plane region K. As previously shown, $P(K)=4M(K)/A(K)$ where $M(K)$, defined by formula (1), is the mean (or expected) area of the triangle formed by three points taken at random in K and $A(K)$ is the area of K.

As noted above, Woolhouse obtained

$$M(K)=(35/48\pi^2)(\pi r^2), \quad \text{and} \quad P(K)=35/12\pi^2$$

when $K=D$ is a circular disk.

Question No. 1229 from the *Educational Times* of 1865 [**17**] proposed by S. Watson and solved by J. J. Sylvester, was to show that $M(K)=(1/12)A(K)$ when K is a triangle. It follows that $P(K)=1/3$ when $K=\Delta$ is a triangle. By 1867, Woolhouse [**19**] had computed $M(K)$ and $P(K)$ for K a square, and K a regular hexagon. We summarize these values in TABLE 4.

It should be noted that, if T is a non-singular affine transformation of the plane, then $P(T(K))=P(K)$. This explains why squares and parallelograms, or circles and

TABLE 4

K	Triangle	Square or Parallelogram	Regular Hexagon	Circle or Ellipse
$M(K)$	$A(K)/12$	$11A(K)/144$	$289A(K)/3888$	$35A(K)/48\pi^2$
$P(K)$	$1/3$	$11/36$	$289/972$	$35/12\pi^2$

ellipses have the same probability. Making the convention that $A(K)=1$ will simplify TABLE 4, so we will adopt that convention and only allow area-preserving affine transformations T.

Notice (from TABLE 4) that $1/3 > 11/36 > 289/972 > 35/12\pi^2$. J. J. Sylvester asked for the shape of the regions K that gave the maximum and minimum probabilities $P(K)$ (see [6]), and the conjecture was:

(i) $P(D) \leqslant P(K)$, when D is bounded by a circle or an ellipse;

(ii) $P(\triangle) \geqslant P(K)$, when $\triangle$ is bounded by a triangle.

The computational problem of finding $P(K)$ and this new variational problem (or conjecture) are now collectively known as Sylvester's Four Point Problem [2, 3, 11, 12, 13, 14].

The first (not quite rigorous) proof of (i) was given by M. W. Crofton [6] in 1885. But it was not until 1917 that a complete proof of both (i) and (ii) was given by W. Blaschke [2]. Blaschke gave another proof in 1923 [3]. The two proofs given by Blaschke use the same geometric ideas so we will outline Blaschke's solution emphasizing these ideas.

3. The Geometry of Blaschke's Solution of Sylvester's Four Point Problem W. Blaschke actually proved the equivalent conjecture for the mean value $M(K)$, namely:

(i′) $M(D) \leqslant M(K)$, where D is bounded by a circle or an ellipse;

(ii′) $M(\triangle) \geqslant M(K)$, where $\triangle$ is bounded by a triangle. And he showed that equality holds if and only if K is an ellipse in (i′) or K is a triangle in (ii′).

To simplify the expression for $M(K)$, we continue the convention that $A(K)=1$.

Blaschke's solution depends on the geometry of the multiple integral

$$M(K)=\int_a^b\int_a^b\int_a^b\int_{\alpha_1}^{\beta_1}\int_{\alpha_2}^{\beta_2}\int_{\alpha_3}^{\beta_3}\frac{1}{2}\left\|\begin{matrix}1 & x_1 & y_1\\ 1 & x_2 & y_2\\ 1 & x_3 & y_3\end{matrix}\right\|dy_3\,dy_2\,dy_1\,dx_3\,dx_2\,dx_1,$$

and, more specifically, the "inside" triple integral,

$$I(x_1,x_2,x_3)=\int_{\alpha_1}^{\beta_1}\int_{\alpha_2}^{\beta_2}\int_{\alpha_3}^{\beta_3}|Ay_1+By_2+Cy_3|\,dy_3\,dy_2\,dy_1.$$

(See FIGURE 5.)

Here $A=(x_3-x_2)$, $B=(x_1-x_3)$, and $C=(x_2-x_1)$. We may assume x_1, x_2, x_3 are distinct.

Since we cannot sketch the surface given by the integrand $f(y_1,y_2,y_3)=|Ay_1+By_2+Cy_3|$, let's look at the integral

$$I=\int_{\alpha_1}^{\beta_1}\int_{\alpha_2}^{\beta_2}|Ay_1+By_2|\,dy_2\,dy_1,$$

as shown in FIGURE 6.

Since the integral, I, represents the volume under the surface $z=|Ay_1+By_2|$ and above the rectangle $R=[\alpha_1,\beta_1]\times[\alpha_2,\beta_2]$, we can see from FIGURE 6 that the value of I is a strictly increasing function of the distance, d, of the center of R, $(m_1,m_2)=(\alpha_1+\beta_1,\alpha_2+\beta_2)/2$, from the line $Ay_1+By_2=0$ in the (y_1,y_2) plane. The corresponding fact holds for our integral $I(x_1,x_2,x_3)$.

In addition, the convexity of K forces the determinants

$$l = \begin{vmatrix} 1 & x_1 & \alpha_1 \\ 1 & x_2 & \alpha_2 \\ 1 & x_3 & \alpha_3 \end{vmatrix} \quad \text{and} \quad u = \begin{vmatrix} 1 & x_1 & \beta_1 \\ 1 & x_2 & \beta_2 \\ 1 & x_3 & \beta_3 \end{vmatrix}$$

to have opposite signs. (Remember, x_1, x_2, x_3 may be in any order.) This is illustrated in FIGURE 5.

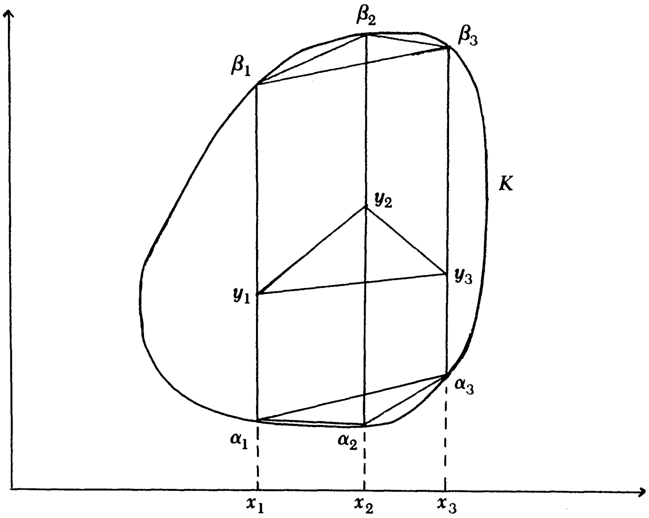


FIGURE 5.

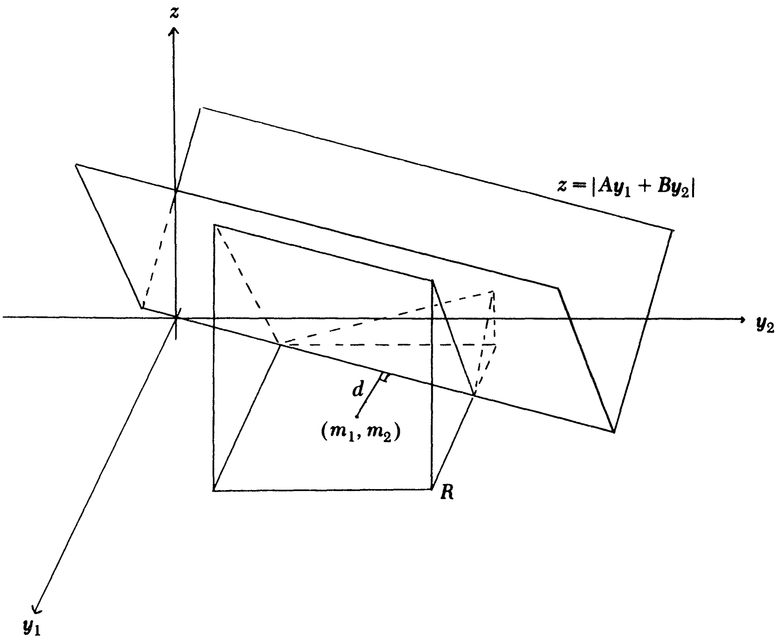


FIGURE 6.

Using the formula for distance, d, from the center $(m_1, m_2, m_3) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \alpha_3 + \beta_3)/2$ to the plane $Ay_1 + By_2 + Cy_3 = 0$, we have

$$d = \frac{|Am_1 + Bm_2 + Cm_3|}{\sqrt{A^2+B^2+C^2}} = \frac{\left\|\begin{matrix} 1 & x_1 & \alpha_1+\beta_1 \\ 1 & x_2 & \alpha_2+\beta_2 \\ 1 & x_3 & \alpha_3+\beta_3 \end{matrix}\right\|}{2\sqrt{A^2+B^2+C^2}} = \frac{\left\|\begin{matrix} 1 & x_1 & \alpha_1 \\ 1 & x_2 & \alpha_2 \\ 1 & x_3 & \alpha_3 \end{matrix}\right\| + \left\|\begin{matrix} 1 & x_1 & \beta_1 \\ 1 & x_2 & \beta_2 \\ 1 & x_3 & \beta_3 \end{matrix}\right\|}{2\sqrt{A^2+B^2+C^2}}$$

$$= |l + u| / 2\sqrt{A^2+B^2+C^2}.$$

Note that d is a constant multiple of the area of the "midpoint triangle" with vertices $(x_1, m_1), (x_2, m_2), (x_3, m_3)$.

Now, if we allow the three intervals in FIGURE 5 to vary vertically, subject to the constraint that l and u have opposite sign (or one or both are zero), we see that d attains its minimum value of zero when $l = -u$ and d attains its maximum when $l = 0$ or $u = 0$. In other words, $I(x_1, x_2, x_3)$ attains the smallest value, call it $I^*(x_1, x_2, x_3)$, if the three intervals in FIGURE 5 have their midpoints $(x_1, m_1), (x_2, m_2), (x_3, m_3)$ on a line N. $I(x_1, x_2, x_3)$ attains its maximum value, call it $\bar{I}(x_1, x_2, x_3)$, when the lower endpoints $(x_1, \alpha_1), (x_2, \alpha_2), (x_3, \alpha_3)$ lie on a line $\bar{N}$ (i.e. $l = 0$). See FIGURE 7, where we have used the x-axis for both lines N and $\bar{N}$. We may always use the x-axis for the line N (or $\bar{N}$) because of the invariance of $I(x_1, x_2, x_3)$ under transformations of the form $T(x, y) = (x, y - (mx + b))$.

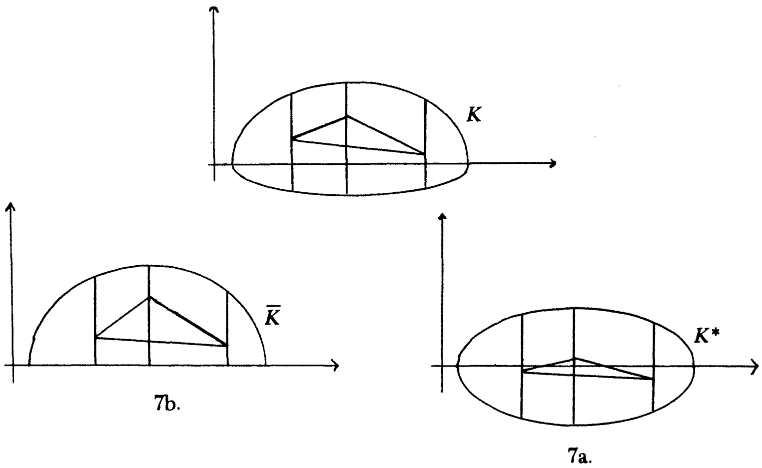


FIGURE 7.

Since x_1, x_2, x_3 was an arbitrary triple, centering all of the vertical line segments of K in a line will form a new set K^*, (FIGURE 7a) with the property $M(K^*) \leqslant M(K)$. Setting all of the vertical line segments on top of the x-axis will form a new set $\bar{K}$, (FIGURE 7b) with the property $M(\bar{K}) \geqslant M(K)$. These two operations, of forming K^* and $\bar{K}$ from a given set K, are called the Steiner Symmetrization of K and the Schüttelung of K, respectively, in the line N, and are well-known in the geometry of convex sets [1, 4, 8]. It is easy to see that K, K^*, and $\bar{K}$ all have the same area and it is not difficult to show that K^* and $\bar{K}$ are convex whenever K is convex. It is also well known that there exists a sequence of Steiner Symmetrizations of K (respectively Schüttelung operations of K) in a sequence of lines $N_1, N_2, \ldots,$ (resp. $\bar{N}_1, \bar{N}_2, \ldots,$) which converges to a circle [4, 8] (respectively, triangle [1]). See FIGURE 8.

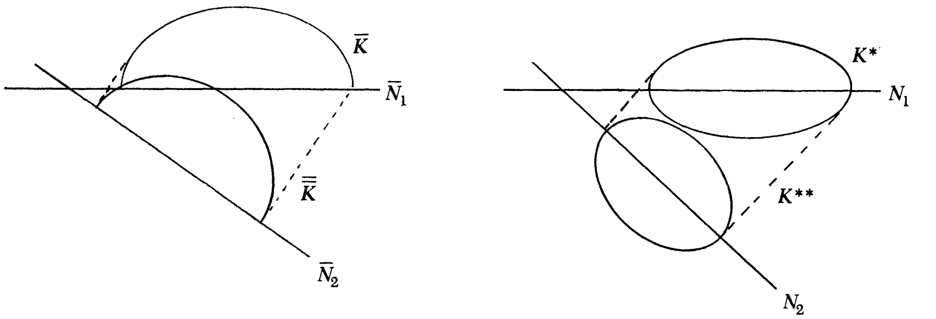


FIGURE 8.

If we denote these sequences by $K_0 = K$, $K_1 = K^*$, $K_2 = (K^*)^*$, etc., and $K^0 = K$, $K^1 = \bar{K}$, $K^2 = (\bar{\bar{K}})$, etc., and observe that $M(K)$ is a continuous function of K, we see that $\{M(K_i)\}$ is a decreasing sequence with limit $M(D)$, where D is bounded by a circle, and $\{M(K^i)\}$ is an increasing sequence with limit $M(\triangle)$, where $\triangle$ is bounded by a triangle. This completes the outline of the proof of (i′) and (ii′). The equality conditions are established by observing that the first Symmetrization (resp., Schüttelung) of K can be made to strictly decrease (resp., increase) $M(K)$ unless K is an ellipse (resp., triangle).

4. The Generalization of Sylvester's Four Point Problem to Three Dimensions; an Unsolved Problem If we let K be a three dimensional, compact, convex set of volume 1, and define $M(K)$ to be the mean (or expected) value of the volume of the tetrahedron formed by 4 points taken at random (uniform distribution) from K, the natural generalization of Sylvester's Four Point Problem would be the conjecture:

(a) $M(K) \geqslant M(D)$, where D is bounded by a sphere or an ellipsoid; and equality holds if and only if K is a solid ellipsoid;

(b) $M(K) \leqslant M(\triangle)$ where $\triangle$ is a solid tetrahedron, and equality holds if and only if K is a tetrahedron.

Blaschke stated in 1917 [2] that both (a) and (b) were true, and that the proofs, as outlined in section 3, would carry over to higher dimensions. For (a), he was correct. This was verified by Groemer [7] in 1973. However, conjecture (b) is still unsolved! Why doesn't the proof of (ii′) carry over to the three-dimensional problem? In two dimensions, the convexity of K forced the "lower" and "upper" determinants l and u to have opposite sign. For Blaschke's proof to work in three dimensions, the convexity of K must force the corresponding determinants

$$l = \begin{vmatrix} 1 & x_1 & y_1 & \alpha_1 \\ 1 & x_2 & y_2 & \alpha_2 \\ 1 & x_3 & y_3 & \alpha_3 \\ 1 & x_4 & y_4 & \alpha_4 \end{vmatrix}, \quad \text{and} \quad u = \begin{vmatrix} 1 & x_1 & y_1 & \beta_1 \\ 1 & x_2 & y_2 & \beta_2 \\ 1 & x_3 & y_3 & \beta_3 \\ 1 & x_4 & y_4 & \beta_4 \end{vmatrix}$$

to have opposite sign. But this is not the case! Referring to FIGURE 9, K is a tetrahedron and

$$l = \begin{vmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 1 \end{vmatrix} = 4, \quad \text{and} \quad u = \begin{vmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 2 \\ 1 & -1 & 0 & 1 \\ 1 & 0 & -1 & 2 \end{vmatrix} = 4.$$

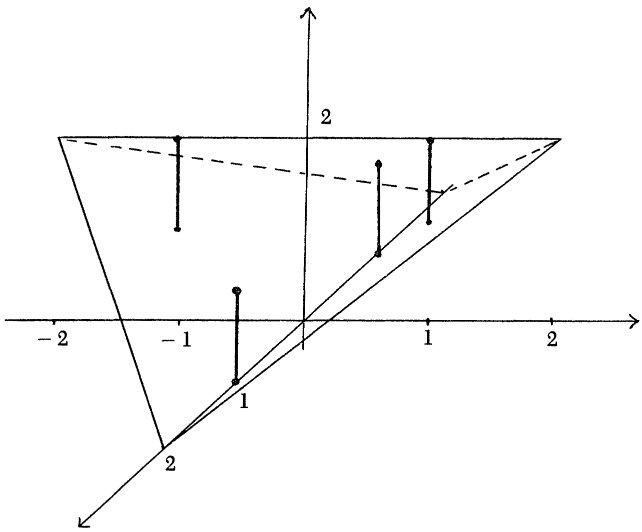


FIGURE 9.

If we shifted the four intervals shown in FIGURE 9, using the Schüttelung process, the corresponding function d would *decrease* to 0. (The verification of this is left to the reader.) For Blaschke's solution of (ii′) to carry over to (b), d should increase! This indicates that the Schüttelung procedure may not be the correct procedure. (Perhaps (b) is not the correct generalization of (ii′).) If you would care to work on this problem and/or a restatement of (b), please do.

By the way, the thought may have occurred to the reader that perhaps a maximum does not exist. That is, maybe there is no set Δ of volume one such that $M(K) \leqslant M(\Delta)$ for all K. That would not be correct. Using the result of John [**10**], that every three-dimensional, compact, convex set of K of volume 1 is contained inside an ellipsoid of volume $3^3 = 27$, it is (relatively) easy to show that such a maximum set exists. This, too, will be left to the reader.

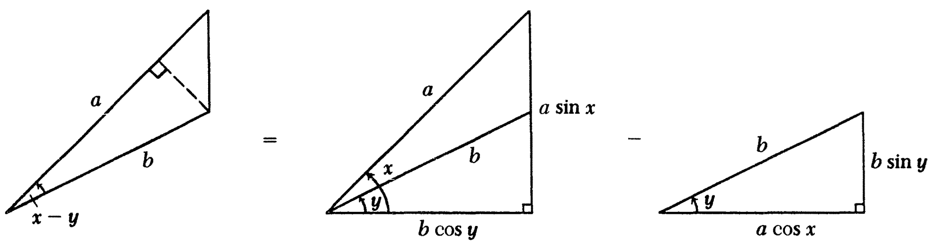
5. Concluding Remarks We have followed a single linear sequence of events in the development of Sylvester's Four Point Problem in order to arrive at the current status of only one of its offspring. There have been many other relatives along the way. There is a survey of results in the book by Santaló [**13**]. Solomon's book [**14**] gives related results and concentrates on some of the computations using Crofton's Theorem. For more historical remarks, the article by Klee [**12**] is recommended.

REFERENCES

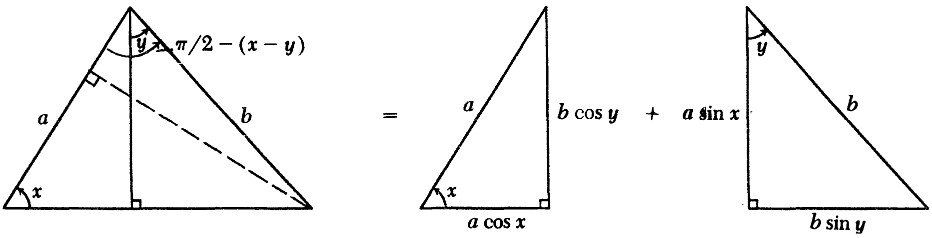
1. T. Biehl, Über Affine Geometrie XXXVIII, Über die Schüttlung von Eikörpern, *Abh. math. Semin. Hamburg Univ.* Bd 2 (1923), 69–70.
2. W. Blaschke, Über affine Geometrie XI: Lösung des "Vierpunktproblems" von Sylvester aus der Theorie der geometrischen Wahrscheinlichkeiten, *Leipziger Berichte* 69 (1917), 436–453.
3. ______, *Vorlesungen über Differentialgeometrie II: Affine Differentialgeometrie*, Springer, Berlin, 1923.
4. T. Bonnesen and W. Fenchel, *Theorie Der Konvexen Körper*, Chelsea, New York, 1948.
5. M. W. Crofton, On the Theory of Local Probability, *Royal Society of London*, *Philosophical Transactions* 158 (1868), 181–200.
6. ______, Probability, *Encyclopedia Brittanica*, 9th ed., 19 (1885), 768–788.

7. H. Groemer, On some mean values associated with a randomly selected simplex in a convex set, *Pacific J. Math.* 45 (1973), 525–533.
8. H. Hadwiger, *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie*, Springer-Verlag, Berlin-Göttingen-Heidelberg, 1957.
9. C. M. Ingleby, *Mathematical Questions and their Solutions from the Educational Times*, W. J. C. Miller, Editor, 5 (1865), 108.
10. F. John, *Extremum Problems with Inequalities as Subsidiary Conditions, Studies and Essays Presented to R. Courant*. Interscience Publishers, Inc., New York, 1948.
11. M. G. Kendall and P. A. P. Moran, Geometrical Probability, *Griffin's Statistical Monographs and Courses*, No. 10, M. G. Kendall, Editor, Hafner Publishing Company, New York, 1963.
12. V. Klee, What is the expected volume of a simplex whose vertices are chosen at random from a given convex body?, *Amer. Math. Monthly* 76 (1969), 286–288.
13. L. A. Santaló, *Integral Geometry and Geometric Probability*, Addison-Wesley Publ. Co., Reading, MA, 1976.
14. H. Solomon, *Geometric Probability*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1978.
15. J. J. Sylvester, *The Educational Times*, London, April (1864).
16. ______, *Report of the British Association* 35 (1865), 8–9.
17. S. Watson, Question No. 1229, *Mathematical Questions and their Solutions from the Educational Times*, W. J. C. Miller, editor, 4 (1865), 101.
18. W. S. B. Woolhouse, Some Additional Observations on the Four-Point Problem, *Mathematical Questions and their Solutions from the Educational Times*, W. J. C. Miller, editor, 7 (1867), 81.
19. ______, Question no. 2471, *Mathematical Questions and their Solutions from the Educational Times*, W. J. C. Miller, editor, 8 (1867), 100–105.

Proof without Words: Area and Difference Formulas



$$\sin(x-y) = \sin x \cos y - \cos x \sin y$$



$$\cos(x-y) = \cos x \cos y + \sin x \sin y$$

—SIDNEY H. KUNG
JACKSONVILLE UNIVERSITY