

Sylvester's Four Point Problem

Introduction to Crofton's Formula in Geometric Probability

Sajid Bin Mahamud

Introduction

James Joseph Sylvester posed the following problem in the Educational times of 1864:

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

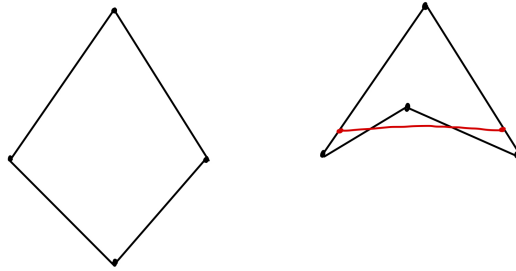


Figure 1: Reentrant Quadrilateral

In simple terms, part (1) of problem asks us to show that the probability of the event that four points sampled uniformly randomly on the infinite plane forms a concave quadrilateral is $\frac{1}{4}$. Part (2) asks for the probability in any region of bounded area. The problem, posed in the late nineteenth century, predates the invention of measure theoretic formulation of probability. As a result, the fact that ‘points taken at random in an indefinite plane’ is undefined evaded even the great J.J. Sylvester.

Definition 0.0.1. A set of point $K \subset \mathbb{R}^d$ is convex if it contains the line segments connecting any pair of points within K . The convex hull of set K , denoted by $\text{Conv}(K)$ is the unique minimal convex set that contains K .

On a later issue of the Education times, Sylvester noted, “this problem does not admit of a determinate solution” and modified the question. Which, in modern terminology, translates to:

Problem Find the probability of the event where the convex hull of four randomly distributed points on any region K of the plane is a triangle.

In this paper, I investigate the problem when K is a disk of radius r , D_r . At first, I will attempt the problem with techniques from class and, hopefully, demonstrate that we need tools outside traditional probability theory to solve the problem reasonably.

Calculation: Brute Force

Let D_r be a disk of radius $r > 0$. Sample four random points from the disk D_r :

$$M, N, P, Q \sim \text{Unif}(D_r)$$

We have to find the probability of the event where one of the four points is contained in the triangle formed by the remaining three. The event is given by the union of four similar events:

$$E = [M \in \Delta(N, Q, P)] \vee [N \in \Delta(M, P, Q)] \vee [Q \in \Delta(M, N, P)] \vee [P \in \Delta(M, N, Q)]$$

These events are mutually exclusive. The proof follows from the fact that only one point can be contained inside the triangle formed by the rest at any given time. It follows from symmetry that all the events have equal probability. As a result, for simplicity, we can write:

$$Pr[E] = 4Pr[Q \in \Delta(M, N, P)]$$

Now note that computing $Pr[Q \in \Delta(M, N, P)]$ is the same as computing the ratio of $\text{Area} \Delta(M, N, P)$ to $\text{Area}(D_r)$. This follows from the following observation: Q is either contained in $\Delta(M, N, P)$ or contained in its complement $D_r - \Delta(M, N, P)$. As a result, we get:

$$Pr[E] = \frac{4\mathbb{E}[\text{Area}(M, N, P)]}{\text{Area}(D_r)}$$

We know $\text{Area}(D_r) = \pi r^2$. Then the problem boils down to finding the expected area of $\text{Area} \Delta(M, N, P) = \text{Area}(M, N, P)$. The integral we have to evaluate:

$$\mathbb{E}[\text{Area}(M, N, P)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{Area}(m, n, p) f_{M, N, P}(m, n, p) dm dn dp$$

This is not as bad as it looks when you assume that $r = 1$. Because we've assumed

$$M, N, P \sim \text{Unif}(D_1)$$

We can define the joint probability density function to be:

$$f_{M, N, P}(m, n, p) = \begin{cases} \frac{1}{\pi^3} & \text{when } (m, n, p) \in D_1 \\ 0 & \text{when } (m, n, p) \notin D_1 \end{cases}$$

Now let

$$m = (x_1, y_1), n = (x_2, y_2), p = (x_3, y_3)$$

We can then use the shoelace formula to compute:

$$\text{Area}(m, n, p) = \frac{1}{2} \left| \det \begin{bmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix} \right|$$

Now, assume that D_1 is a disk of radius 1 centered at the origin, so that we can use polar coordinates:

$$m = (x_1, y_1) = (\sqrt{r_1} \cos \theta_1, \sqrt{r_1} \sin \theta_1)$$

$$n = (x_2, y_2) = (\sqrt{r_2} \cos \theta_2, \sqrt{r_2} \sin \theta_2)$$

$$p = (x_3, y_3) = (\sqrt{r_3} \cos \theta_3, \sqrt{r_3} \sin \theta_3)$$

Then:

$$\text{Area}(m, n, p) = \frac{1}{2} \left| \det \begin{bmatrix} 1 & 1 & 1 \\ \sqrt{r_1} \cos \theta_1 & \sqrt{r_2} \cos \theta_2 & \sqrt{r_3} \cos \theta_3 \\ \sqrt{r_1} \sin \theta_1 & \sqrt{r_2} \sin \theta_2 & \sqrt{r_3} \sin \theta_3 \end{bmatrix} \right|$$

Recall the trig identity:

$$\sin(\alpha - \beta) = \sin \alpha \cos \beta - \cos \alpha \sin \beta$$

We can further simplify:

$$\text{Area}(m, n, p) = \frac{1}{2} |\sqrt{r_1 r_3} \sin(\theta_1 - \theta_3) + \sqrt{r_1 r_2} \sin(\theta_2 - \theta_1) + \sqrt{r_2 r_3} \sin(\theta_3 - \theta_2)|$$

We can further specify what the joint pdf looks like:

$$\begin{aligned} f_{M,N,P}(r_1, r_2, r_3, \theta_1, \theta_2, \theta_3) \\ = \begin{cases} \frac{1}{\pi^3} & \text{when } (\sqrt{r_1}, \sqrt{r_2}, \sqrt{r_3}) \in [0, 1]^3 \text{ and } (\theta_1, \theta_2, \theta_3) \in [0, 2\pi]^3 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Now, let's write the full integral:

$$\begin{aligned} \mathbb{E}[\text{Area}(M, N, P)] \\ &= \frac{1}{\pi^3} \int_{[0, 2\pi]} \int_{[0, 2\pi]} \int_{[0, 2\pi]} \int_{[0, 1]} \int_{[0, 1]} \int_{[0, 1]} \text{Area}(m, n, p) \, dr_1 \, dr_2 \, dr_3 \, d\theta_1 \, d\theta_2 \, d\theta_3 \\ &= \frac{1}{\pi^3} \int_{[0, 2\pi]^3} \int_{[0, 1]^3} \text{Area}(r_1, r_2, r_3, \theta_1, \theta_2, \theta_3) \, dr_1 \, dr_2 \, dr_3 \, d\theta_1 \, d\theta_2 \, d\theta_3 \\ &= \frac{1}{\pi^3} \int_{[0, 2\pi]^3 \times [0, 1]^3} \frac{1}{2} |\sqrt{r_1 r_3} \sin(\theta_1 - \theta_3) + \sqrt{r_1 r_2} \sin(\theta_2 - \theta_1) + \sqrt{r_2 r_3} \sin(\theta_3 - \theta_2)| \, dr_1 \cdots \end{aligned}$$

Evaluating this integral simply by adjusting the bounds of integration is going to be a nightmare. As a result, we need an important tool called **Crofton's Formula**, named after Morgan Crofton. I consider it to be a sin to throw formulas around without motivating them, so let's first see where the formula comes from.

Crofton's Formula

Let's take things down by one notch. Instead of computing the area of a convex hull of three points in the unit disk, let's compute the area of the convex hull of two points. That is, we want to find the expected length of a line segment drawn from a disk of radius 1. Let's simplify the problem by considering a disk of radius r . Let $\vec{Q}_1 = (x_1, y_1)$ and $\vec{Q}_2 = (x_2, y_2)$

be points sampled from a uniform distribution on the domain $D_r = B_r(0,0)$, or a disk of radius r :

$$\vec{Q}_1, \vec{Q}_2 \sim \text{Unif}(D_r)$$

Take the convex hull of \vec{Q}_1 and \vec{Q}_2 : the line connecting points (x_1, y_1) and (x_2, y_2) :

$$\text{Conv}(\vec{Q}_1, \vec{Q}_2) = \vec{Q}_1(1-t) + \vec{Q}_2 t \quad \text{for } t \in [0, 1]$$

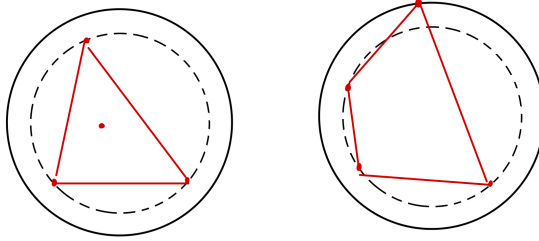
Our aim is to find the expected area of $\text{Conv}(\vec{Q}_1, \vec{Q}_2)$ or the length in this case:

$$\mathbb{E}_r[\text{Area}(\text{Conv}(\vec{Q}_1, \vec{Q}_2))] = \mathbb{E}_r[|\vec{Q}_1 - \vec{Q}_2|]$$

Consider a new domain D_{r+h} , a disk of radius $r+h$. Like before, sample points from the domain:

$$\vec{Q}_1, \vec{Q}_2 \sim \text{Unif}(D_{r+h})$$

One can ask what is the relationship between the expected areas of points sampled from the domains: D_r and D_{r+h} .



We first compute the joint probability of the points in the bigger domain, D_{r+h} :

$$P_{r+h}(\vec{Q}_1, \vec{Q}_2) = \begin{cases} \left(\frac{\pi r^2}{\pi(r+h)^2} \right)^2 & \text{when } \vec{Q}_1, \vec{Q}_2 \in D_r \\ \left(\frac{\pi r^2}{\pi(r+h)^2} \right) \left(1 - \frac{\pi r^2}{\pi(r+h)^2} \right) & \text{when } \vec{Q}_1 \in D_{r+h} - D_r, \vec{Q}_2 \in D_r \\ \left(\frac{\pi r^2}{\pi(r+h)^2} \right) \left(1 - \frac{\pi r^2}{\pi(r+h)^2} \right) & \text{when } \vec{Q}_2 \in D_{r+h} - D_r, \vec{Q}_1 \in D_r \\ \left(1 - \frac{\pi r^2}{\pi(r+h)^2} \right)^2 & \text{when } \vec{Q}_1, \vec{Q}_2 \in D_{r+h} - D_r \end{cases}$$

Now consider the case when $\vec{Q}_1, \vec{Q}_2 \in D_r$:

$$\begin{aligned}
P_{r+h}(\vec{Q}_1, \vec{Q}_2) &= \left(\frac{\pi r^2}{\pi(r+h)^2} \right)^2 \\
&= \left(\frac{r^4}{r^4 + 4r^3h + 6r^2h^2 + 4rh^3 + h^4} \right) \\
&= \left(\frac{1}{1 + \frac{4h}{r} + \frac{6h^2}{r^2} + \frac{4h^3}{r^3} + \frac{h^4}{r^4}} \right) \\
&= \left(\frac{1}{1 + 4\frac{h}{r} + 6\frac{h^2}{r^2} + 4\frac{h^3}{r^3} + \frac{h^4}{r^4}} \right) \\
&= \frac{1}{\left(1 + 4\frac{h}{r} + 6\frac{h^2}{r^2} + 4\frac{h^3}{r^3} + \frac{h^4}{r^4} \right)} \\
&= \frac{1}{\left(1 + \frac{4h}{r} + \frac{6h^2}{r^2} + \frac{4h^3}{r^3} + \frac{h^4}{r^4} \right)}
\end{aligned}$$

From now onwards, assume that $|h| \ll r$. Then we can use the power series expansion to get:

$$P_{r+h}(\vec{Q}_1, \vec{Q}_2) = 1 - \frac{4h}{r} + o(h)$$

We can similarly compute the probability one of the points is contained inside D_r and the other in $D_{r+h} - D_r$:

$$\begin{aligned}
P_{r+h}(\vec{Q}_i \in D_r, \vec{Q}_j \in D_{r+h} - D_r) &= 2 \left(\frac{\pi r^2}{\pi(r+h)^2} \right) \left(1 - \frac{\pi r^2}{\pi(r+h)^2} \right) \\
&= 2 \left(\frac{r^2}{(r+h)^2} \right) \left(1 - \frac{r^2}{(r+h)^2} \right) \\
&= 2 \left(\frac{r^2}{r^2(1 + \frac{h}{r})^2} \right) \left(1 - \frac{r^2}{r^2(1 + \frac{h}{r})^2} \right) \\
&= 2 \left(\frac{1}{(1 + \frac{h}{r})^2} \right) \left(1 - \frac{1}{(1 + \frac{h}{r})^2} \right)
\end{aligned}$$

Using binomial approximation with $|\frac{2h}{r}| \ll r$ and power series expansion, we get:

$$\begin{aligned}
P_{r+h}(\vec{Q}_i \in D_r, \vec{Q}_j \in D_{r+h} - D_r) &\approx \left(\frac{2}{1 + \frac{2h}{r}} \right) \left(1 - \frac{1}{1 + \frac{2h}{r}} \right) \\
&\approx 2 \left(1 - \frac{2h}{r} \right) \cdot \frac{2h}{r} \\
&= \frac{4h}{r} + o(h)
\end{aligned}$$

Lastly, compute the probability that both the points are in $D_{r+h} - D_r$:

$$\begin{aligned}
P_{r+h}(\vec{Q}_1, \vec{Q}_2 \in D_{r+h} - D_r) &= \left(1 - \frac{\pi r^2}{\pi(r+h)^2}\right)^2 \\
&\approx \left(1 - \frac{1}{1 + \frac{2h}{r}}\right)^2 \\
&\approx \left(\frac{2h}{r}\right)^2 \\
&= o(h)
\end{aligned}$$

Now, recall that our original ambition was to compute

$$\mathbb{E}_r[\text{Area}(\text{Conv}(\vec{Q}_1, \vec{Q}_2))] = \mathbb{E}_r[|\vec{Q}_1 - \vec{Q}_2|]$$

Denote $P_{r+h} = P$. Then, using conditional expectation, one can show that

$$\begin{aligned}
&\mathbb{E}[|\vec{Q}_1 - \vec{Q}_2| \text{ given } \vec{Q}_1, \vec{Q}_2 \in D_{r+h}] \\
&= P(\vec{Q}_1, \vec{Q}_2 \in D_r) \cdot \mathbb{E}[|\vec{Q}_1 - \vec{Q}_2| \text{ given } \vec{Q}_1, \vec{Q}_2 \in D_r] \\
&+ P(\vec{Q}_1 \in D_r, \vec{Q}_2 \in D_{r+h} - D_r) \cdot 2\mathbb{E}[|\vec{Q}_1 - \vec{Q}_2| \text{ given } \vec{Q}_1 \in D_r, \vec{Q}_2 \in D_{r+h} - D_r] \\
&+ P(\vec{Q}_1, \vec{Q}_2 \in D_{r+h} - D_r) \mathbb{E}[|\vec{Q}_1 - \vec{Q}_2| \text{ given } \vec{Q}_1, \vec{Q}_2 \in D_{r+h} - D_r]
\end{aligned}$$

This is becoming cumbersome, so let's take \vec{Q}_1, \vec{Q}_2 to be implicit:

$$\begin{aligned}
&= \mathbb{E}[\text{length} \mid D_{r+h}] \\
&= P(D_r) \cdot \mathbb{E}[\text{length} \mid D_r] \\
&+ 2P(D_r, D_{r+h} - D_r) \cdot \mathbb{E}[\text{length} \mid D_r, D_{r+h} - D_r] \\
&+ P(D_{r+h} - D_r) \cdot \mathbb{E}[\text{length} \mid D_{r+h} - D_r]
\end{aligned}$$

Take length to be implicit and plug in values from before:

$$\begin{aligned}
\mathbb{E}[D_{r+h}] &= \left(1 - \frac{4h}{r}\right) \mathbb{E}[D_r] + \frac{4h}{r} \mathbb{E}[D_r, (D_{r+h} - D_r)] \\
\mathbb{E}[D_{r+h}] - \mathbb{E}[D_r] &= \frac{4h}{r} (\mathbb{E}[D_r, (D_{r+h} - D_r)] - \mathbb{E}[D_r]) \\
\frac{\mathbb{E}[D_{r+h}] - \mathbb{E}[D_r]}{h} &= \frac{4}{r} \mathbb{E}[D_r, (D_{r+h} - D_r)] - \frac{4}{r} \mathbb{E}[D_r]
\end{aligned}$$

Taking the limit as $h \rightarrow 0$, we get:

$$\mathbb{E}'[D_r] = \frac{4}{r} \mathbb{E}[D_r, (D_r - D_r)] - \frac{4}{r} \mathbb{E}[D_r]$$

We need to be careful about taking the limits on the right hand side. Consider the first term of the summand: $\mathbb{E}[\text{length} \mid \vec{Q}_i \in D_r, \vec{Q}_j \in D_{r+h} - D_r]$. As $h \rightarrow 0$, the domain $D_{r+h} - D_r$ collapses to C_r , the boundary of the disk D_r . As a result:

$$\lim_{h \rightarrow 0} \mathbb{E}[\text{length} \mid \vec{Q}_i \in D_r, \vec{Q}_j \in D_{r+h} - D_r] = \mathbb{E}[\text{length} \mid \vec{Q}_i \in D_r, \vec{Q}_j \in C_r]$$

Meaning, limit of the expected length of points exclusively sampled from D_r and $D_{r+h} - D_r$ approaches the expected length for when one of the points is sampled from D_r and the other from its boundary C_r , as $h \rightarrow 0$.

Rewrite to get Crofton's formula for a convex domain with two randomly sampled points:

$$\mathbb{E}'[\text{length} | D_r] = \frac{4}{r} \mathbb{E}[\text{length} | D_r, C_r] - \frac{4}{r} \mathbb{E}[\text{length} | D_r]$$

Let's try to unpack here. The equation says that the rate of change of expected length when both points are sampled from D_r , can be expressed as a sum of two expected lengths: (1) expected length when one point is sampled from the circumference and the other from the interior, and (2) both points are sampled from the interior, weighted by constants dependent on r . We make an important remark here.

Remark 1. *The expected length of the line segment $\ell(\vec{Q}_1, \vec{Q}_2)$, is a function of the radius r :*

$$\mathbb{E}[|\vec{Q}_1 - \vec{Q}_2|] = \mathbb{E}[\text{length} | D_r] = kr$$

for some constant k .

This follows from simple observation about Euclidean distance. Say two points \vec{Q}_1, \vec{Q}_2 sampled from a unit disk D_1 are k -unit away from each other:

$$|\vec{Q}_1 - \vec{Q}_2| = k$$

If we now enlarge the disk to have radius r , then, for the linearly transformed points \vec{Q}'_1, \vec{Q}'_2 now sampled from the bigger disk, D_r we have:

$$|\vec{Q}'_1 - \vec{Q}'_2| = r|\vec{Q}_1 - \vec{Q}_2| = kr$$

Since expectation is linear, we know

$$\mathbb{E}[|\vec{Q}'_1 - \vec{Q}'_2|] = \mathbb{E}[k|\vec{Q}_1 - \vec{Q}_2|] = k \mathbb{E}[|\vec{Q}_1 - \vec{Q}_2|]$$

I mention this because this tell us that rate of change of expected length should be a constant:

$$\mathbb{E}'[\text{length} | D_r] = k$$

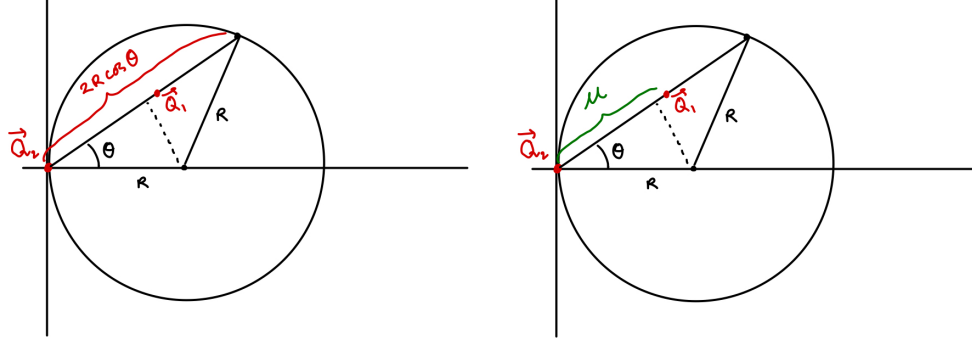
This observation simplifies Crofton's formula into:

$$kr = \frac{4}{r} \mathbb{E}[\text{length} | D_r, C_r] - \frac{4}{r} kr$$

We have one remaining expected to compute which is relatively simpler:

$$\mathbb{E}[|\vec{Q}_1 - \vec{Q}_2| \text{ given } \vec{Q}_1 \in D_r, \vec{Q}_2 \in C_r]$$

Note that due to the symmetry of the circumference, we can fix \vec{Q}_2 at the origin: $\vec{Q}_2 = (0, 0)$. Then:



$$\mathbb{E}[|\vec{Q}_1 - \vec{Q}_2| \text{ given } \vec{Q}_1 \in D_r, \vec{Q}_2 = (0, 0)] = \int_{D_r} \sqrt{x_2^2 + y_2^2} f_{\vec{Q}_1}$$

Here $f_{\vec{Q}_1}$ is simply the probability distribution function of \vec{Q}_1 which we know equals the inverse of the area of D_r .

$$\mathbb{E}[|\vec{Q}_1 - \vec{Q}_2| \text{ given } \vec{Q}_1 \in D_r, \vec{Q}_2 = (0, 0)] = \frac{1}{\text{Area}(D_r)} \int_{D_r} \sqrt{x_2^2 + y_2^2}$$

We can parametrize \vec{Q}_2 as a function of μ and θ where $\mu = \sqrt{x_2^2 + y_2^2}$ where $\theta = [-\frac{\pi}{2}, \frac{\pi}{2}]$. From the picture above, we see that as $\mu \in [0, 2R \cos(\theta)]$. As a result, we get:

$$\begin{aligned} \mathbb{E}[|\vec{Q}_1 - \vec{Q}_2| \text{ given } \vec{Q}_1 \in D_r, \vec{Q}_2 = (0, 0)] &= \frac{1}{\pi r^2} \int_{-\pi/2}^{\pi/2} \int_0^{2R \cos(\theta)} \mu \, d\mu \, d\theta \\ &= \frac{32r}{9\pi} \end{aligned}$$

Plugging into Crofton's formula, we get:

$$\begin{aligned} kr &= \frac{4}{r} \cdot \frac{32r}{9\pi} - \frac{4}{r} kr \\ k &= \frac{128}{45\pi} \end{aligned}$$

Recall the remark we made on the last page:

$$\mathbb{E}[|\vec{Q}_1 - \vec{Q}_2|] = \mathbb{E}[\text{length } |D_r|] = kr$$

Setting $r = 1$, we get:

$$\mathbb{E}[|\vec{Q}_1 - \vec{Q}_2|] = \mathbb{E}[\text{length } |D_1|] = k$$

So we have successfully computed the expected length of a line drawn from the unit disk at random.

Conclusion

The nice thing about Crofton's formula is that it generalizes for any convex set K , and n uniformly, randomly chosen points from K . We can then use a similar approach to find the expected area of a triangle whose vertices are sampled uniformly, randomly from the unit disk:

$$\mathbb{E}[\text{Area}(\Delta) \text{ given that } \Delta \in D_1] = \frac{35}{48\pi}$$

The computation becomes slightly involved so I will leave it as an exercise for the reader. As a result, we know that the probability of the event where the convex hull of four randomly distributed points on a unit disk is a triangle, denoted by E is:

$$P(E) = \frac{4 \mathbb{E}[\text{Area}(\Delta)]}{\pi} = \frac{35}{12\pi^2}$$

I like the result because its considered to be one of the earliest works in the world of geometric probability. Nice extensions to the problem can be found in the references below.

References

- [1] Eisenberg, B., & Sullivan, R. (2000). Crofton's Differential Equation. The American Mathematical Monthly, 107(2), 129–139. <https://doi.org/10.2307/2589434>
- [2] Pfeifer, R. E. (1989). The Historical Development of J. J. Sylvester's Four Point Problem. Mathematics Magazine, 62(5), 309–317. <https://doi.org/10.2307/2689482>
- [3] Solomon H. (1978). Geometric Probability / Herbert Solomon. Society for Industrial and Applied Mathematics; 1978.