

Extended Essay

Mathematics: Analysis and Approaches HL

May 2022

“Average distance between two randomly selected points on the circumference of a unit circle and the perimeter of a unit square.”

Word Count: 3954

Contents

1. Introduction.....	2
2. Background – Bertrand Paradox.....	2
3. Average Chord Length.....	7
4. Applications of Findings About the Average Chord Length of a Unit Circle.....	11
4.1 Estimating the Value of Pi.....	11
4.2 Matter in Circular Orbit.....	15
5. Average Distance Between Two Randomly Selected Points on the Perimeter of a Unit Square.....	17
5.1 Approach.....	19
5.2 Adjacent Sides.....	19
5.3 Same Side.....	29
5.4 Opposite Sides.....	32
5.5 Averaging all the possible outcomes.....	35
5.6 Double-checking the result using computation.....	36
6. Conclusion.....	37
7. Bibliography.....	38

Introduction

Randomness is an interesting concept, both in general and specifically in mathematics. Not only is it challenging to express randomness in mathematical terminology, it is also challenging to find common ground on how problems involving randomness should be approached. These difficulties become even more prominent when infinity gets involved in a question with randomness as well, two concepts that are amongst the most infamous and controversial ones in mathematics. My involvement with the intersection of these two topics began thanks to me stumbling upon the Bertrand Paradox.

Background - Bertrand Paradox

Last year, I stumbled upon the Bertrand Paradox¹. Paradoxes in mathematics were always interesting to me so I started researching about this one as well. Basically, the problem is as follows: *Consider an equilateral triangle inscribed in a circle. A chord of the circle is chosen at random. What is the probability that the chord is longer than a side of the triangle?* The question seems easy enough. But the tricky part is: how does someone define a *randomly selected chord*? The three main approaches to the question are as follows:

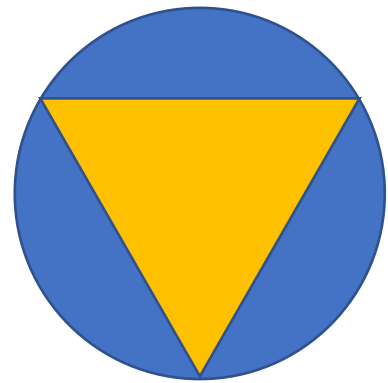


Figure 1. Equilateral triangle inscribed in a circle

¹ "Bertrand Paradox (Probability)." Wikipedia, Wikimedia Foundation, 27 Dec. 2021, [https://en.m.wikipedia.org/wiki/Bertrand_paradox_\(probability\)](https://en.m.wikipedia.org/wiki/Bertrand_paradox_(probability)).

1. The *random endpoints* method, draw two random points on the circumference of the circle and draw a chord joining these two points. In order to calculate the probability in question, assume that the triangle is rotated so one of its vertices is coinciding with one of the randomly selected points. Notice that in order for the chord to be longer than a side of the equilateral triangle the other point should lie on the opposite arc that lies between the other two

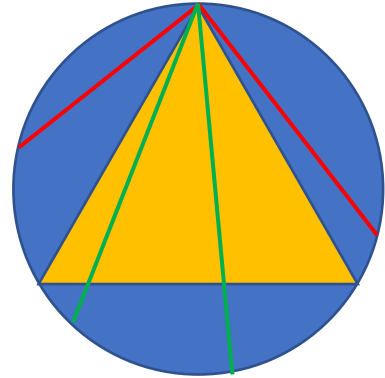


Figure 2. Random endpoints method of Bertrand Paradox

vertices of the triangle. The arc is one third of the circumference of the circle, since the inscribed triangle is equilateral. Therefore, the probability that a random chord is longer than a side of the triangle is $\frac{1}{3}$.

2. The *random radial point* method. Draw a radius line that passes through the middle of one of the sides of the triangle. Choose a random point in the circle. In order to calculate the probability in question, assume that the triangle and the radius line is rotated so that the randomly selected point is on the radius line. Now, draw a chord, that is perpendicular to the radius line, which has the randomly chosen point as its midpoint. Because of

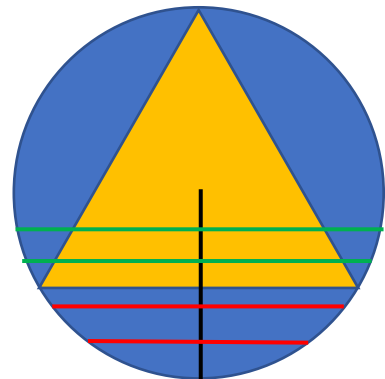


Figure 3. Random radial point method of Bertrand Paradox

the centroid theorem² the intersection of the radius line with the side of the triangle is exactly half a radius. Therefore, the randomly selected point has to be on the one half of the radius line that is

² "Centroid of a Triangle." Centroid, https://www.varsitytutors.com/hotmath/hotmath_help/topics/centroid#:~:text=The%20centroid%20theorem%20states%20that,midpoint%20of%20the%20opposite%20side.

inside the triangle, in order for the chord to be longer than one of the sides of the triangle. In conclusion, the probability that a random chord is longer than a side of the triangle is $\frac{1}{2}$.

3. The *random midpoint* method. Draw a second smaller circle that is inscribed inside the triangle. Now choose a random point anywhere in the bigger circle that the triangle is inscribed in. Draw a chord with the chosen random point as the midpoint. Notice that the chord is longer than a side of the triangle only if its midpoint falls inside the smaller circle. Because of the centroid theorem, the radius of the smaller circle is half of the

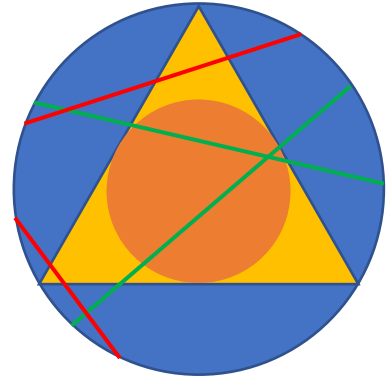


Figure 4. Random midpoint method of Bertrand Paradox

original circle. Therefore, because of the circle area formula ($\pi \cdot r^2$), the area of the smaller circle is $\frac{1}{4}$ of the bigger one. In conclusion, the probability that a random chord is longer than a side of the triangle is $\frac{1}{4}$.

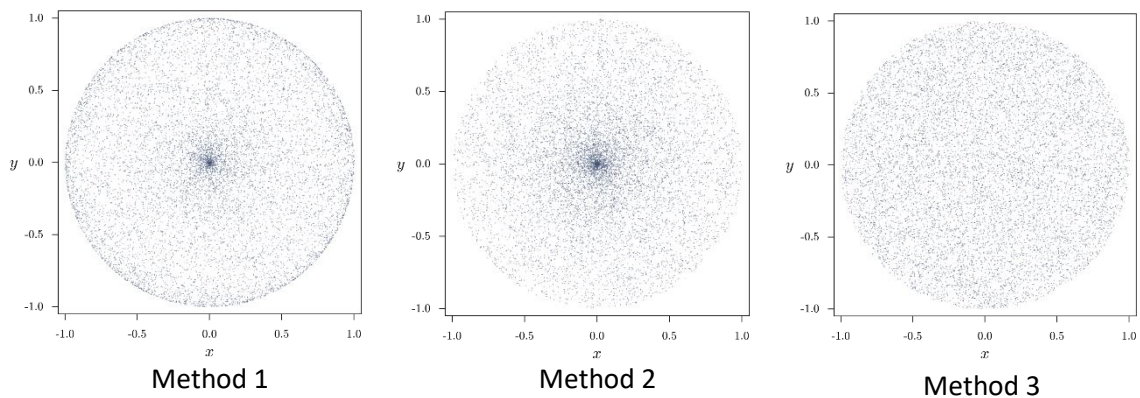


Figure 5. Midpoint distributions of the chords generated using different methods of the Bertrand Paradox

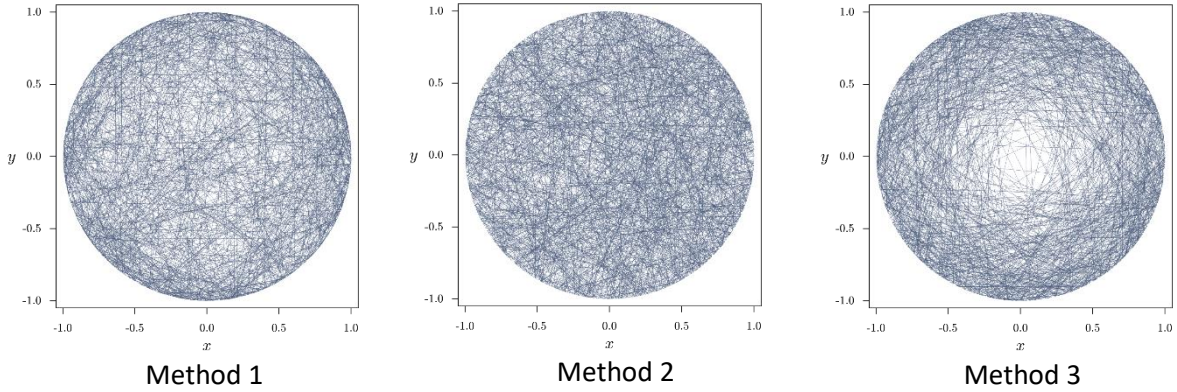


Figure 6. Distributions of the chords generated using different methods of the Bertrand Paradox

This seems very odd, three different answers to the same simple question. The problem lies in the ambiguity of the question. The statement “*randomly selected chord*” is open to a lot of interpretation on how exactly the chords are randomized. To prove this, I came up with a fourth method myself:

4. Imagine infinite lines that are moving from one side of the circle to the other, forming chords on the way. Since it is assumed that the orientation of the chords is evenly distributed, we can just consider that all the lines are moving at the same angle. Now rotate the triangle in order to allow one of its sides to be parallel to the lines. Now, draw another equal triangle inscribed inside the circle that is 180° rotated. Join the vertices of the sides of the triangles that are parallel to the line.

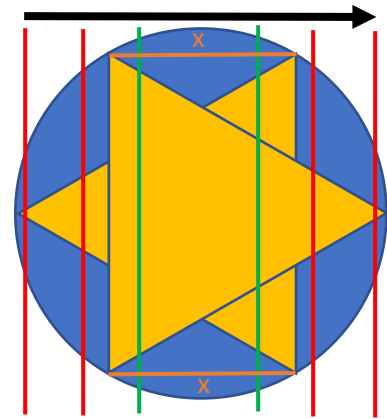


Figure 7. A 4th method for the Bertrand Paradox

Call this newly drawn lines x . Notice that the chords should intersect with x in order to be shorter than a side of the triangles. Therefore, we need to calculate the length of x . Since the shape is a hexagon star, the six small triangles formed on the outside are equilateral as well. Because of the centroid theorem the height of one of the big triangles is 1.5 radius. Therefore, if a vertical line

was to be drawn from one of their vertices to the mid-point of the opposite side to form two right triangles, one side of the big triangles would equal $\frac{1.5}{\sin(60)} = \sqrt{3}r$. This would mean that the length of one of the sides of the small equilateral triangles equals to $\frac{\sqrt{3}}{3}r$. Notice that two sides of the small equilateral triangles and x form an isosceles triangle with angles 30, 30, and 120. Use the sine law to find the length of x :

$$\frac{x}{\sin(120)} = \frac{\left(\frac{\sqrt{3}}{3}\right)}{\sin(30)}$$

Therefore, $x = r$. In order for the chord to be longer than one of the sides of the big equilateral triangle the line has to be coinciding with the lines x . Since the diameter of the circle is $2r$ and the length of x is r , the probability that a random chord is longer than a side of the triangle is $\frac{1}{2}$.

This exploration of the Bertrand Paradox was an eye-opener for me on how ambiguous probability questions can be, especially if they involve infinity. But it also sparked my interest in a more core exploration: the chords in a unit circle. More specifically: *the average chord length of a unit circle*. Obviously, after the exploration of Bertrand Paradox, lesson learned, so I am going to specify how the randomization will be defined for the infinite number of chords. It will be closest to the method 1 of Bertrand Paradox, where two random points will be selected on the circumference of the circle and they will be joint to form a chord.

Average Chord Length

The Equations That Will Be Used:

E1³: Cosine rule

For triangle abc:

$$c = \sqrt{a^2 + b^2 - 2 \cdot a \cdot b \cdot \cos\theta}$$

E2⁴: Sine half angle formula

$$\sin\left(\frac{\theta}{2}\right) = \sqrt{\frac{1 - \cos\theta}{2}}$$

E3⁵: Average value of a function

$$\frac{1}{(b-a)(d-c)} \int_a^b \int_c^d f(x,y) \, dx \, dy$$

E4⁶: Fundamental theorem of calculus

$$\int_b^a f(x) \, dx = F(a) - F(b)$$

3 The Law of Cosines, <https://www.mathsisfun.com/algebra/trig-cosine-law.html>.

4 "Half Angle Formulas - Examples: Half Angle Identities: Proof." Cuemath, <https://www.cuemath.com/half-angle-formula/>.

5 "Calculus." Average Value of a Function, <https://math24.net/average-value-function.html>.

6 "Fundamental Theorems of Calculus." From Wolfram MathWorld, <https://mathworld.wolfram.com/FundamentalTheoremsOfCalculus.html>.

First of all, the problem can be deduced to just the average chord length on a half-circle because any chord can be represented symmetrically on the other half of the circle. If the calculations are made using the whole circle, then twice as much chords would be drawn and then divided by twice as much to find the average which is just unnecessary work. Deducing the problem to a half circle allows for rotation of the circle in order to place one of the points on to a middle line that is passing through the center of the circle.

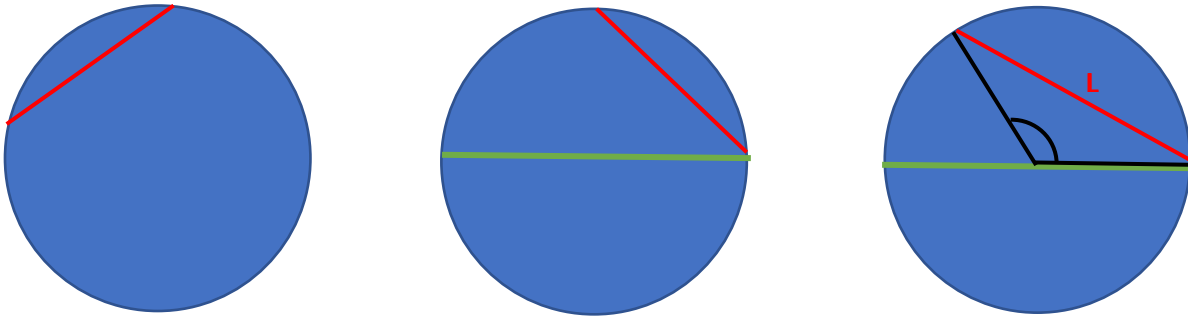


Figure 8. A sketch showing how the function for the chord length is chosen

When the circle is rotated in order to place one of the points on to the middle line, then two lines can be drawn from the two points to the center (Figure 9) and simply the cosine rule (E1) can be used to calculate the length between these two points. ($r = 1$ since it is a unit circle)

$$L = \sqrt{1^2 + 1^2 - 2 \cdot 1 \cdot 1 \cdot \cos\theta}$$

$$L = \sqrt{2(1 - \cos\theta)}$$

$$L = 2 \cdot \sqrt{\frac{(1 - \cos\theta)}{2}}$$

In order to further simplify this function, the sine half angle formula (E2) is needed.

$$L = 2 \cdot \sin\left(\frac{\theta}{2}\right)$$

The formula for finding the average value of a function is E3. Suppose the average value of L is called A.

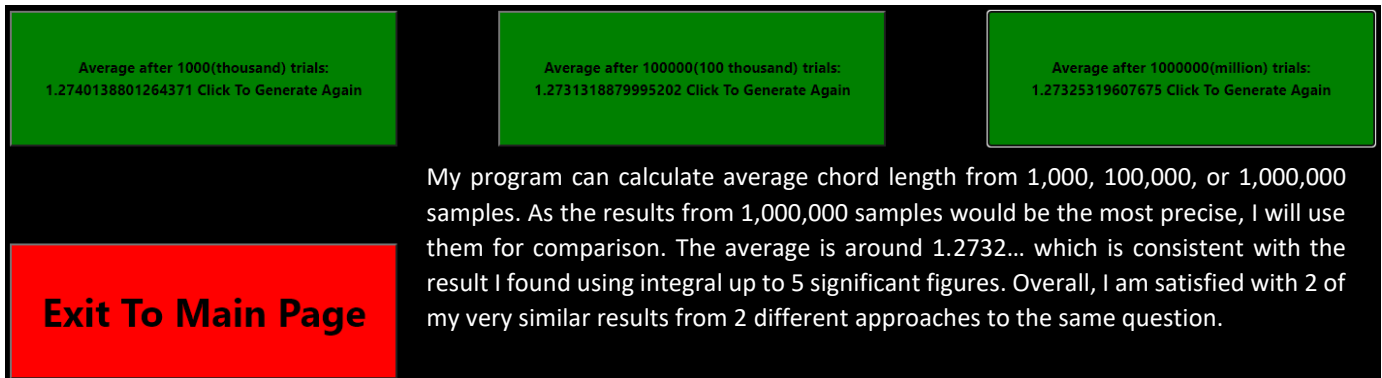
$$A = \frac{1}{\pi - 0} \cdot 2 \cdot \int_0^{\pi} \sin\left(\frac{\theta}{2}\right) d\theta$$

In order to evaluate this integral the fundamental theorem of calculus (E4) is needed. Therefore, the antiderivative of $\sin\left(\frac{\theta}{2}\right)$ is necessary which is $-2 \cdot \cos\left(\frac{\theta}{2}\right)$.

$$\begin{aligned} A &= \frac{2}{\pi} \cdot \left(-2 \cdot \cos\left(\frac{\theta}{2}\right)\right) \Big|_0^{\pi} \\ &= \frac{2}{\pi} \cdot \left(-2 \cdot \cos\left(\frac{\pi}{2}\right) + 2 \cdot \cos\left(\frac{0}{2}\right)\right) \\ &= \frac{4}{\pi} \cdot \left(\cos(0) - \cos\left(\frac{\pi}{2}\right)\right) \\ &= \frac{4}{\pi} (1 - 0) = \frac{4}{\pi} \\ &\approx 1.2732 \end{aligned}$$

Checking the Answer with Programming

I will construct a program that will create 2 random points on a unit circle then calculate the distance between them and repeat this process thousands or millions of times, and finally it will calculate the average. Afterwards, I will compare the answer I get with the answer I got by using integral calculation.



The screenshot shows a dark-themed web interface. At the top, there are three green rectangular boxes, each containing text about the average chord length after a certain number of trials and a 'Click To Generate Again' link. The first box is for 1,000 trials, the second for 100,000 trials, and the third for 1,000,000 trials. Below these boxes is a large red rectangular button with the text 'Exit To Main Page' in white. To the right of the button, there is a paragraph of text explaining the program's purpose and results.

Average after 1000(thousand) trials:
1.2740138801264371 Click To Generate Again

Average after 100000(100 thousand) trials:
1.2731318879995202 Click To Generate Again

Average after 1000000(million) trials:
1.27325319607675 Click To Generate Again

My program can calculate average chord length from 1,000, 100,000, or 1,000,000 samples. As the results from 1,000,000 samples would be the most precise, I will use them for comparison. The average is around 1.2732... which is consistent with the result I found using integral up to 5 significant figures. Overall, I am satisfied with 2 of my very similar results from 2 different approaches to the same question.

Exit To Main Page

Figure 9. The output of the program used to randomly generate chords

```
<script>
var trial;
function calculate() {
  var x1;
  var y1;
  var x2;
  var y2;
  var a1;
  var a2;
  var total = 0;
  for (var i = 0; i < trial; i++) {
    a1 = Math.random()*2*Math.PI;
    a2 = Math.random()*2*Math.PI;
    x1 = Math.cos(a1);
    y1 = Math.sin(a1);
    x2 = Math.cos(a2);
    y2 = Math.sin(a2);
    total += Math.sqrt(Math.pow((x1-x2), 2) + Math.pow((y1-y2), 2));
  }
  total /= trial;
  if (trial == 1000) {
    document.getElementById('thousand').innerHTML = 'Average after 1000(thousand) trials: ' + total + ' Click To Generate Again';
  }
  else if (trial == 100000) {
    document.getElementById('100thousand').innerHTML = 'Average after 100000(100 thousand) trials: ' + total + ' Click To Generate Again';
  }
  else if (trial == 1000000) {
    document.getElementById('million').innerHTML = 'Average after 1000000(million) trials: ' + total + ' Click To Generate Again';
  }
}
document.addEventListener('DOMContentLoaded', function() {
  document.getElementById('thousand').addEventListener('click', function() {
    trial = 1000;
    calculate();
  });
  document.getElementById('100thousand').addEventListener('click', function() {
    trial = 100000;
    calculate();
  });
  document.getElementById('million').addEventListener('click', function() {
    trial = 1000000;
    calculate();
  });
  document.getElementById('exit').addEventListener('click', function() {
    window.location.href = 'https://1de-ebaec32673104e9fa6444948cd91b11-8080.cs90.ws/eindex.html';
  });
});
</script>
```

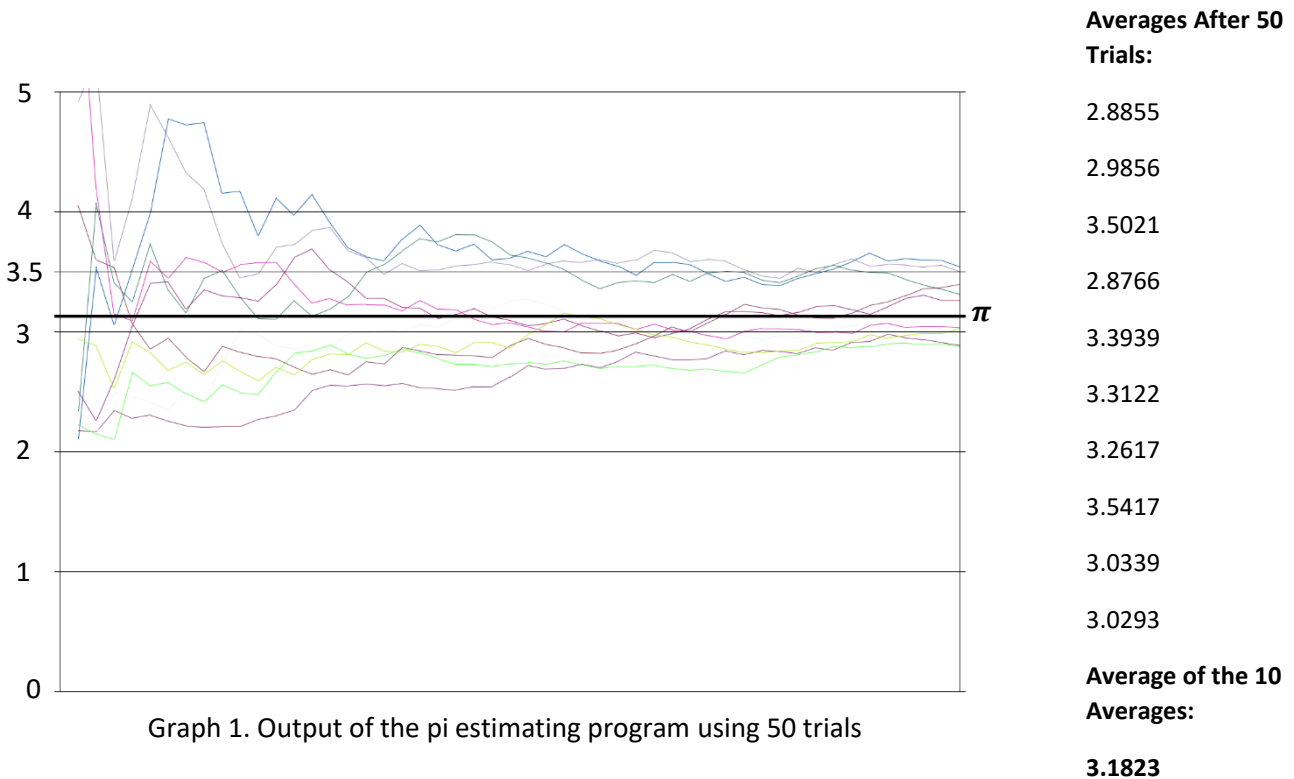
Figure 10. A portion of the code used to randomly generate chords

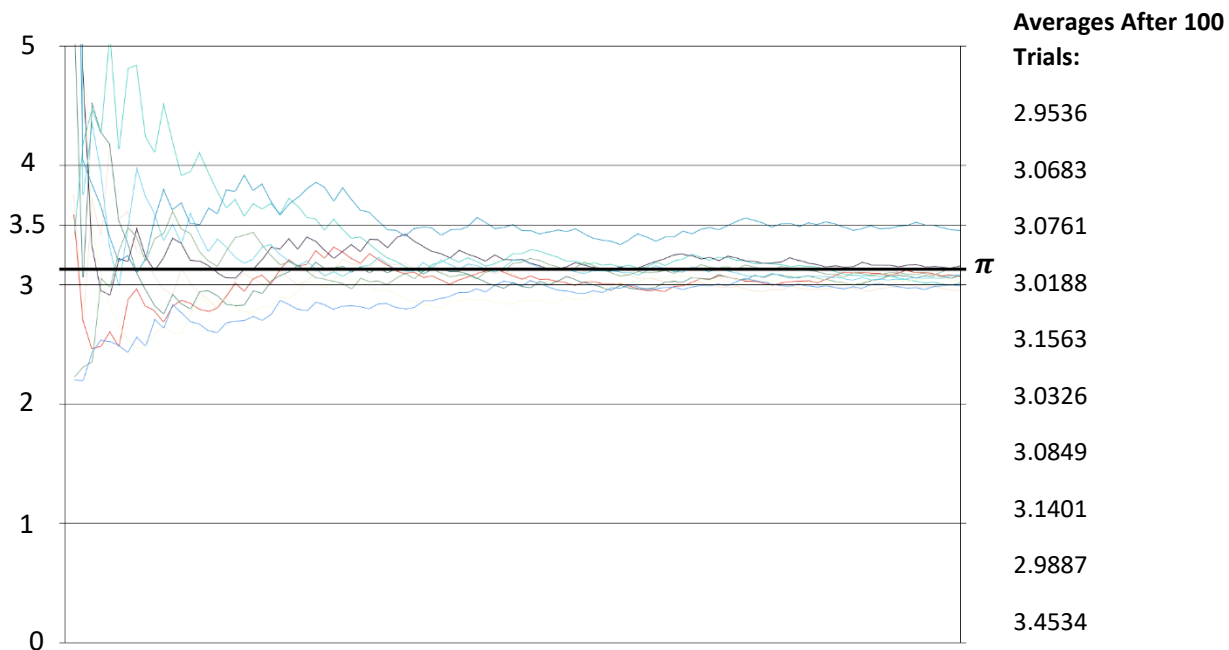
(Link to the full code: https://github.com/DurukanBTN/Extended-Essay/blob/main/Random_Chord_Length.html)

Application of Findings About the Average Chord Length of a Unit Circle

Estimating the Value of Pi

It has been demonstrated that the average chord length of a unit circle is $\frac{4}{\pi}$. Therefore, if the average chord length of a unit circle is called A, then $\frac{4}{A} = \pi$. This equality can be used to compute the value of pi. By generating randomly selected chords, averaging their length and substituting the result to A in order estimate the value of π . The bigger the sample size gets the closer the estimation will be to the actual value of pi. The procedure is similar to Monte Carlo Simulations. To illustrate the phenomenon, I wrote a small graphing program where the **x variable is the number of trials** and the **y variable is $\frac{4}{A}$** , where A is the average chord length after x many trials. The graphs will consist of 50, 100, 200, and 300 trials. Each of them repeated 10 times.

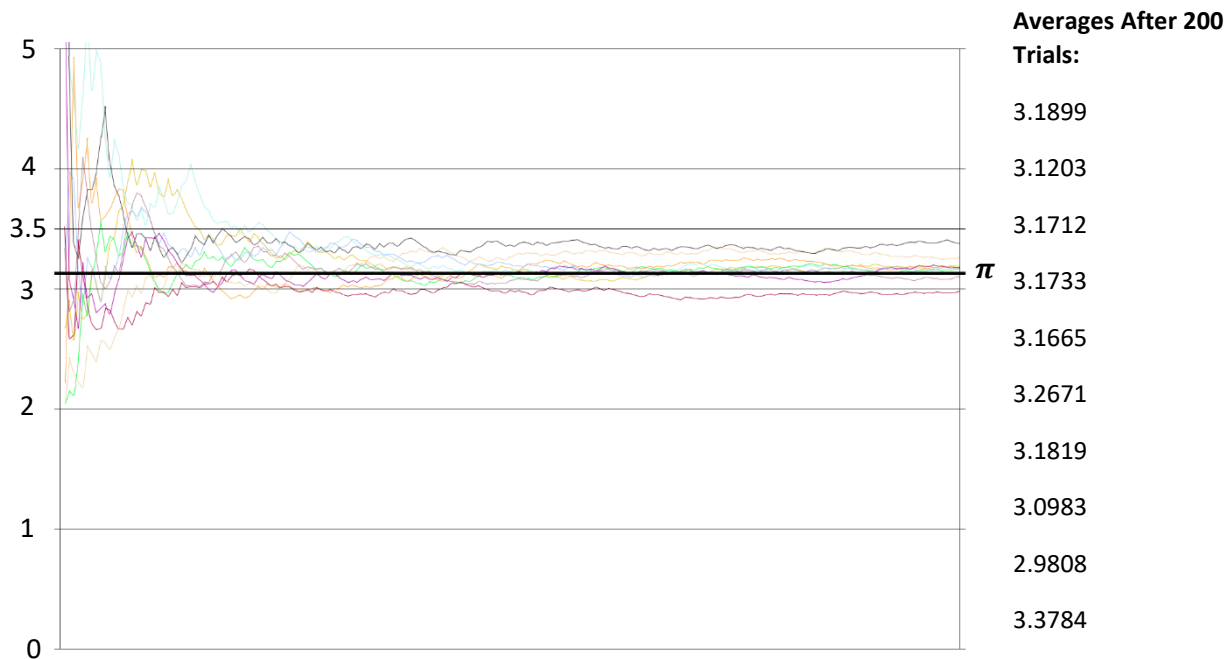




Graph 2. Output of the pi estimating program using 100 trials

Average of the 10 Averages:

3.0973



Graph 3. Output of the pi estimating program using 200 trials

Average of the 10 Averages:

3.1728

As it can be observed, the bigger the trial number gets, the narrower and closer to pi the estimation is. Even with only the average of 10 attempts with 300 trials, the estimation got as close to pi as **3.1149**. If the same procedure was to be repeated again, possibly with a super computer which can generate and average billions, trillions, or possibly even more number of trials, then the estimation would be significantly closer to pi. If a super computer was programmed in such a way to generate random chord lengths continuously, multiply the trial number by 4, and divide the result by the total sum of the randomly generated chord lengths, then more and more digits of pi would be calculated continuously. This is because of the following expression mentioned earlier (A=average length of chords):

$$\frac{4}{A} = \pi$$

$$A = \frac{\textit{total value of the length of randomly generated chord lengths}}{\textit{total trial number}}$$

Matter In Circular Orbit^{7,8}

The average chord length of a circle can be used to estimate the average effect uniformly distributed matters in a circular orbit have on each other. Let me explain what I mean by the help of an example. Every mass on the universe applies gravitational force to each other depending on the distance and the mass of the two objects. The formula for the gravitational force that two masses apply to each other is:

$$F = G \frac{M \cdot m}{d^2}$$

Where $G = \text{gravitational constant} \approx 6.674 \cdot 10^{-11}$

$M = \text{mass of object 1 (kilograms)}$

$m = \text{mass of object 2 (kilograms)}$

$d = \text{the distance between objects (meter)}$

Take Saturn's rings for example, the rings consist of countless differently massed particles. These particles depend primarily on the gravitational force of the Saturn, its moons, and each other to stay in orbit. I am going to attempt to estimate the average gravitational force between any of these two particles, that construct the rings. There is no exact data on the average volume of these particles but it is known that the average ranges from 1cm^3 to 5cm^3 , for the sake of pure estimation and making an example for the application of average chord lengths, the average volume can be

⁷ "Rings of Saturn." Wikipedia, Wikimedia Foundation, 25 Feb. 2022, https://en.wikipedia.org/wiki/Rings_of_Saturn#Physical_characteristics.

⁸ "Rings of Saturn." Wikipedia, Wikimedia Foundation, 25 Feb. 2022, https://en.wikipedia.org/wiki/Rings_of_Saturn#Major_subdivisions.

simply estimated to be 3cm^3 . It is also known that 99.9% of these particles compose of pure water ice, which has a density of 1g/cm^3 . Therefore, it can be estimated that the average mass of a particle which composes Saturn's rings is 3grams. I will use the G ring of Saturn for context as it is pretty thin, which means that the chord length will not be affected by much uncertainty. The G ring has a radius of 170,000 km. Therefore, according to the average chord length calculation the average distance between two random particles in Saturn's G ring is:

$$\frac{4}{\pi} \cdot 170,000$$

$$= 216,450 \text{ kilometers}$$

Using this, the average gravitational force between the particles in the G ring can be estimated as:

$$F = 6.674 \cdot 10^{-11} \cdot \frac{3 \cdot 10^{-3} \cdot 3 \cdot 10^{-3}}{(216,450 \cdot 10^3)^2}$$

$$= \mathbf{1.282 \cdot 10^{-32} \text{ Newton}}$$

The scope of this application can include any type of matter in circular orbits that are randomly or uniformly distributed. Also, the interaction does not have to be gravitational force either, it can be any type of physical interaction, for example electric force. The example of Saturn's rings was pure estimation to make an example of how the knowledge of average chord lengths can be applied.

Average Distance Between Two Randomly Selected Points on the Perimeter of a Unit Square

The average chord length calculation of a unit circle revealed surprising results. They also left an open door to me about whether similar estimations are possible for other geometrical shapes as well. To answer this question, I am going to take another generic shape for context: square. To be consistent with the meaning of *randomly selected chords* in the previous exploration, I will once again define the question as: *What is the average distance between two randomly selected points on the perimeter of a unit square?*

The Equations That Will Be Used:

E1⁹: Average Value of a Double Integral

$$\frac{1}{(b-a)(d-c)} \int_a^b \int_c^d f(x,y) \, dx \, dy$$

E2: Expressions like $\sqrt{bx^2 + a}$ are usually substituted as:

$$x = \frac{\sqrt{a}}{\sqrt{b}} \tan(u).$$

⁹ "Calculus Early Transcendentals: Integral & Multi-Variable Calculus for Social Sciences." Double Integrals: Volume and Average Value, https://www.sfu.ca/math-coursenotes/Math%20158%20Course%20Notes/sec_VolumeAvgHeight.html.

E3¹⁰:

$$\int \sec^n(x) dx = \frac{\sec^{n-1}(x)\sin(x)}{n-1} + \frac{n-2}{n-1} \int \sec^{n-2}(x) dx.$$

E4:

$$\int \sec(x) dx = \ln|\tan(x) + \sec(x)|.$$

E5:

$$\sec(\arctan(x)) = \sqrt{1+x^2}$$

E6:

$$1 + \tan^2(u) = \sec^2(u)$$

E7⁶: Fundamental Theorem of Calculus

$$\int_b^a f(x) dx = F(a) - F(b)$$

E8: Integration by parts

$$\int uv' = uv - \int u'v$$

10 "How to Derive Power Reducing Formula for $\int (\sec^N x) dx$ and $\int (\tan^N x) dx$ for Integration?: Socratic."
Socratic.org, 19 May 2018, <https://socratic.org/questions/how-to-derive-power-reducing-formula-for-int-sec-nx-dx-and-int-tan-nx-dx-for-int>.

1. The Approach

The average length has to be calculated by averaging the 3 different scenarios: 2 points on the same side, 2 points on opposite sides, and 2 points on adjacent sides. If one of the points is based on a side, then there are 2 possible adjacent sides. Therefore, when averaging the 3 scenarios the result of 2 adjacent points' average will be added twice and the whole total will be divided to the total number of possible outcomes, which is 4.

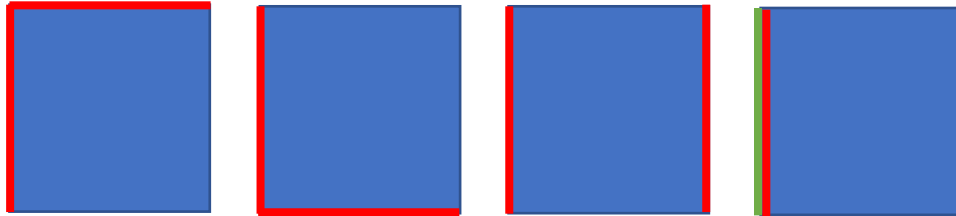


Figure 12. All possible variations of how the two randomly selected points can be oriented

2. Adjacent Sides

First the average distance between two random points on adjacent sides of a square has to be calculated. If y component of the point on the vertical side is y and the x component of the point on the horizontal side is x then the equation for the distance between them would be based off of the Pythagorean Theorem and simply: $\sqrt{x^2 + y^2}$. Because of E1 the average value of two random points on adjacent sides of a unit square would be:

$$\begin{aligned} & \frac{1}{(1-0)(1-0)} \int_0^1 \int_0^1 \sqrt{x^2 + y^2} \, dx dy \\ & = \int_0^1 \int_0^1 \sqrt{x^2 + y^2} \, dx dy \end{aligned}$$

The integral inside which is $\int_0^1 \sqrt{x^2 + y^2} dx$ will be solved first. y^2 can be treated as a constant because we are integrating over x . Trigonometric substitution has to be applied to simplify the expression and get rid of the square root. Because of E2 the substitution will be:

$$x = y \tan(u)$$

$$\frac{dx}{du} = y \sec^2(u)$$

$$dx = y \sec^2(u) du$$

Therefore, the expression is:

$$\begin{aligned} & \int \sqrt{y^2 + (y \tan(u))^2} y \sec^2(u) du \\ &= y^2 \int \sqrt{1 + \tan^2(u)} \sec^2(u) du \end{aligned}$$

Because of E6:

$$y^2 \int \sec^3(u) du$$

To calculate the definite integral, the integral boundaries have to be adjusted.

$$x = y \tan(u)$$

$$u = \arctan\left(\frac{x}{y}\right)$$

x is integrated from 0 to 1 therefore 1 and 0 have to be substituted for x thus:

$$u = \arctan\left(\frac{1}{y}\right)$$

$$\text{upper boundary: } u = \arctan\left(\frac{1}{y}\right)$$

$$u = \arctan\left(\frac{0}{y}\right)$$

$$\text{lower boundary: } u = 0$$

In conclusion, the new boundaries are 0 and $\arctan\left(\frac{1}{y}\right)$. So, the definite integral is:

$$y^2 \int_0^{\arctan\left(\frac{1}{y}\right)} \sec^3(u) du$$

Because of E3 the expression is:

$$y^2 \left(\frac{\sec^2(u) \sin(u)}{2} \Big|_0^{\arctan\left(\frac{1}{y}\right)} + \frac{1}{2} \int_0^{\arctan\left(\frac{1}{y}\right)} \sec(u) du \right).$$

E4, E5, and E7 have to be used to calculate the right side of the equation:

$$\begin{aligned} & \frac{1}{2} \int_0^{\arctan\left(\frac{1}{y}\right)} \sec(u) du \\ &= \frac{1}{2} \left(\ln |\tan(u) + \sec(u)| \Big|_0^{\arctan\left(\frac{1}{y}\right)} \right) \\ &= \frac{1}{2} \left(\ln \left| \tan\left(\arctan\left(\frac{1}{y}\right)\right) + \sec\left(\arctan\left(\frac{1}{y}\right)\right) \right| - \ln |\tan(0) + \sec(0)| \right) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \left(\ln \left| \frac{1}{y} + \sqrt{\frac{y^2 + 1}{y^2}} \right| - 0 \right) \\
&= \frac{1}{2} \left(\ln \left| \frac{1}{y} + \frac{\sqrt{y^2 + 1}}{|y|} \right| \right) \\
&= \frac{1}{2} \left(\ln \left| \frac{1 + \sqrt{y^2 + 1}}{|y|} \right| \right)
\end{aligned}$$

The expression has come down to:

$$y^2 \left(\frac{\sec^2(u) \sin(u)}{2} \Big|_0^{\arctan(\frac{1}{y})} + \frac{1}{2} \ln \left| \frac{1 + \sqrt{y^2 + 1}}{|y|} \right| \right)$$

Now, the left side of the expression has to be calculated using E7:

$$\begin{aligned}
&\frac{\sec^2(u) \sin(u)}{2} \Big|_0^{\arctan(\frac{1}{y})} \\
&= \frac{1}{2} \sec(u) \tan(u) \Big|_0^{\arctan(\frac{1}{y})} \\
&= \frac{1}{2} \sec \left(\arctan \left(\frac{1}{y} \right) \right) \tan \left(\arctan \left(\frac{1}{y} \right) \right) - \frac{1}{2} \sec(0) \tan(0) \\
&= \frac{1}{2} \cdot \frac{1}{y} \cdot \sqrt{1 + \left(\frac{1}{y} \right)^2} \\
&= \sqrt{\frac{y^2 + 1}{y^2}} \cdot \frac{1}{2} \cdot \frac{1}{y}
\end{aligned}$$

$$= \frac{\sqrt{y^2 + 1}}{y} \cdot \frac{1}{2y}$$

$$= \frac{\sqrt{y^2 + 1}}{2y^2}$$

So now the inside integral of the starting double integral equation is evaluated as:

$$\int_0^1 \sqrt{x^2 + y^2} \, dx = y^2 \left(\frac{\sqrt{y^2 + 1}}{2y^2} + \frac{1}{2} \ln \left| \frac{1 + \sqrt{y^2 + 1}}{|y|} \right| \right)$$

Now, this expression has to be integrated with respect to y (because the outside part of the double integral is still present):

$$\int_0^1 y^2 \left(\frac{\sqrt{y^2 + 1}}{2y^2} + \frac{1}{2} \ln \left(\frac{1 + \sqrt{y^2 + 1}}{|y|} \right) \right) dy$$

First, this expression has to be divided into 2 separate integrals to simplify the problem:

$$\int_0^1 \frac{\sqrt{y^2 + 1}}{2} dy + \frac{1}{2} \int_0^1 y^2 \ln \left(\frac{1 + \sqrt{y^2 + 1}}{|y|} \right) dy$$

Now, the integral on the left has to be solved using E2 and substitution:

$$y = \frac{\sqrt{1}}{\sqrt{1}} \tan(u)$$

$$dy = \sec^2(u) \, du$$

The integral boundaries have to be adjusted.

$$u = \arctan(1)$$

$$\text{Upper boundary: } u = \frac{\pi}{4}$$

$$u = \arctan(0)$$

$$\text{Lower boundary: } u = 0$$

So, the integral is:

$$\begin{aligned} & \frac{1}{2} \int_0^{\frac{\pi}{4}} \sqrt{\tan^2(u) + 1} \sec^2(u) du \\ &= \frac{1}{2} \int_0^{\frac{\pi}{4}} \sec^3(u) du \text{ because of E6} \end{aligned}$$

Using E3 and E4, the expression is:

$$\begin{aligned} & \frac{1}{2} \left(\left. \frac{\sec^2(u) \sin(u)}{2} \right|_0^{\frac{\pi}{4}} + \frac{1}{2} \int_0^{\frac{\pi}{4}} \sec(u) du \right) \\ & \frac{1}{2} \left(\left[\frac{\sec^2(u) \sin(u)}{2} \right]_0^{\frac{\pi}{4}} + \frac{1}{2} \left[\frac{\ln |\tan(u) + \sec(u)|}{1} \right]_0^{\frac{\pi}{4}} \right) \end{aligned}$$

E7 needs to be used to calculate the expression:

$$\begin{aligned} & \frac{1}{2} \left(\frac{\sec^2\left(\frac{\pi}{4}\right) \sin\left(\frac{\pi}{4}\right)}{2} - \frac{\sec^2(0) \sin(0)}{2} + \frac{1}{2} \left(\ln \left| \tan\left(\frac{\pi}{4}\right) + \sec\left(\frac{\pi}{4}\right) \right| - \ln |\tan(0) + \sec(0)| \right) \right) \\ &= \frac{1}{2} \left(\frac{\sec^2\left(\frac{\pi}{4}\right) \sin\left(\frac{\pi}{4}\right)}{2} + \frac{1}{2} \ln \left| \tan\left(\frac{\pi}{4}\right) + \sec\left(\frac{\pi}{4}\right) \right| \right) \end{aligned}$$

$$= \frac{1}{2} \left(\frac{2\sqrt{2}}{2} + \frac{1}{2} \ln \left| 1 + \frac{2}{\sqrt{2}} \right| \right)$$

$$= \frac{1}{2} \left(\frac{\sqrt{2}}{2} + \frac{1}{2} (0.8814) \right)$$

$$\approx \mathbf{0.5739}$$

This was the solution to only one of the two integrals in the expression:

$$\int_0^1 \frac{\sqrt{y^2+1}}{2} dy + \frac{1}{2} \int_0^1 y^2 \ln \left(\frac{1+\sqrt{y^2+1}}{|y|} \right) dy$$

$$= 0.5739 + \frac{1}{2} \int_0^1 y^2 \ln \left(\frac{1+\sqrt{y^2+1}}{|y|} \right) dy$$

Now to solve the remaining integral, E8 is needed:

$$u = \ln \left(\frac{1+\sqrt{y^2+1}}{|y|} \right)$$

$$v' = y^2$$

$$u' = \frac{d}{dy} \ln \left(\frac{1+\sqrt{y^2+1}}{|y|} \right)$$

$$= \frac{|y|}{1+\sqrt{y^2+1}} \left(\left(\frac{1+\sqrt{y^2+1}}{|y|} \right) \frac{d}{dy} \right)$$

$$= \frac{|y|}{1+\sqrt{y^2+1}} \frac{\left(\frac{y|y|}{(y^2+1)^{\frac{1}{2}}} - \frac{y}{|y|} (1+\sqrt{y^2+1}) \right)}{y^2}$$

$$\begin{aligned}
&= \frac{|y|}{1 + \sqrt{y^2 + 1}} \left(\frac{y^2}{|y|\sqrt{y^2 + 1}} - \frac{\sqrt{y^2 + 1}(1 + \sqrt{y^2 + 1})}{|y|\sqrt{y^2 + 1}} \right) \\
&= \frac{|y|}{1 + \sqrt{y^2 + 1}} \frac{\left(\frac{-\sqrt{y^2 + 1} - 1}{|y|\sqrt{y^2 + 1}} \right)}{y} \\
&= \frac{1}{1 + \sqrt{y^2 + 1}} \left(\frac{-\sqrt{y^2 + 1} - 1}{y\sqrt{y^2 + 1}} \right)
\end{aligned}$$

$$u' = -\frac{1}{y\sqrt{y^2 + 1}}$$

$$v = \int y^2 dy = \frac{y^3}{3}$$

Now, the integral can be deduced to:

$$\begin{aligned}
&\frac{1}{2} \left(\ln \left(\frac{1 + \sqrt{y^2 + 1}}{|y|} \right) \frac{y^3}{3} - \int -\frac{1}{y\sqrt{y^2 + 1}} \frac{y^3}{3} dy \right) \\
&= \frac{1}{2} \left(\frac{1}{3} \ln \left(\frac{1 + \sqrt{y^2 + 1}}{|y|} \right) y^3 - \int -\frac{1}{\sqrt{y^2 + 1}} \frac{y^2}{3} dy \right)
\end{aligned}$$

As a final step the integral on the right has to be solved:

$$\begin{aligned}
&\int -\frac{1}{\sqrt{y^2 + 1}} \frac{y^2}{3} dy \\
&= -\frac{1}{3} \int \frac{y^2}{\sqrt{y^2 + 1}} dy
\end{aligned}$$

Once again, E2 will be used:

$$y = \frac{\sqrt{1}}{\sqrt{1}} \tan(u)$$

$$dy = \sec^2(u) du$$

So, the integral is:

$$\begin{aligned} & -\frac{1}{3} \int \frac{\tan^2(u)}{\sqrt{\tan^2(u) + 1}} \sec^2(u) du \\ &= -\frac{1}{3} \int \frac{\tan^2(u)}{\sqrt{\sec^2(u)}} \sec^2(u) du \\ &= -\frac{1}{3} \int \tan^2(u) \sec(u) du \\ &= -\frac{1}{3} \int (-1 + \sec^2(u)) \sec(u) du \\ &= -\frac{1}{3} \int \sec^3(u) - \sec(u) du \\ &= -\frac{1}{3} \left(\int \sec^3(u) du - \int \sec(u) du \right) \\ &= -\frac{1}{3} \left(\int \sec^3(u) du - \ln|\tan(u) + \sec(u)| \right) \end{aligned}$$

The integral on the left has to be solved using E3:

$$\begin{aligned} & \int \sec^3(u) du \\ &= \frac{\sec^2(u) \sin(u)}{2} + \frac{1}{2} \int \sec(u) du \end{aligned}$$

$$= \frac{\sec^2(u) \sin(u)}{2} + \frac{1}{2} \ln |\tan(u) + \sec(u)|$$

$$= \frac{1}{2} \sec(u) \tan(u) + \frac{1}{2} \ln |\tan(u) + \sec(u)|$$

The solved part has to be substituted back into the equation:

$$= -\frac{1}{3} \left(\frac{1}{2} \sec(u) \tan(u) + \frac{1}{2} \ln |\tan(u) + \sec(u)| - \ln |\tan(u) + \sec(u)| \right)$$

$$= -\frac{1}{3} \left(\frac{1}{2} \sec(\arctan(y)) \tan(\arctan(y)) + \frac{1}{2} \ln |\tan(\arctan(y)) + \sec(\arctan(y))| - \ln |\tan(\arctan(y)) + \sec(\arctan(y))| \right)$$

E5 has to be used to continue:

$$= -\frac{1}{3} \left(\frac{1}{2} y \sqrt{1+y^2} - \frac{1}{2} \ln |y + \sqrt{1+y^2}| \right)$$

$$= -\frac{1}{6} \left(y \sqrt{1+y^2} - \ln |y + \sqrt{1+y^2}| \right)$$

This expression has to be substituted back into the whole equation:

$$\frac{1}{2} \left(\frac{1}{3} \ln \left(\frac{1 + \sqrt{y^2 + 1}}{|y|} \right) y^3 + \frac{1}{6} \left(y \sqrt{1 + y^2} - \ln |y + \sqrt{1 + y^2}| \right) \right) + C$$

Now this value has to be evaluated from 1 to 0:

$$\frac{1}{2} \left(\frac{1}{3} \ln \left(\frac{1 + \sqrt{y^2 + 1}}{|y|} \right) y^3 + \frac{1}{6} \left(y \sqrt{1 + y^2} - \ln |y + \sqrt{1 + y^2}| \right) \right) \Bigg|_0^1$$

$$= \frac{1}{2} \left(\frac{1}{3} \ln \left(\frac{1 + \sqrt{1^2 + 1}}{|1|} \right) 1^3 + \frac{1}{6} \left(1 \sqrt{1 + 1^2} - \ln |1 + \sqrt{1 + 1^2}| \right) \right) - \frac{1}{2} \left(0 + \frac{1}{6} \left(0 \sqrt{1 + 0^2} - \ln |0 + \sqrt{1 + 0^2}| \right) \right)$$

$$\approx \mathbf{0.1913}$$

Now this number has to be plugged into the previous equation:

$$\begin{aligned} 0.5739 + \frac{1}{2} \int_0^1 y^2 \ln \left(\frac{1 + \sqrt{y^2 + 1}}{|y|} \right) dy \\ \approx 0.5739 + 0.1913 \\ \approx \mathbf{0.7652} \end{aligned}$$

3. Same Side

Now that the average distance between two random points on adjacent sides are calculated, it is time to calculate the average distance between two random points on the same side. If the x coordinate of the first point is called “ x ” and the y coordinate of the second point is called “ y ”, then the distance between these two points would be $|x - y|$. Because of E1 the average value of two random points on the same side of a unit square would be:

$$\begin{aligned} \frac{1}{(1-0)(1-0)} \int_0^1 \int_0^1 |x - y| dx dy \\ = \int_0^1 \int_0^1 |x - y| dx dy \end{aligned}$$

First, the integral inside has to be calculated. Substitution needs to take place:

$$u = x - y, \frac{du}{dx} = 1$$

Which would yield:

$$\int |u| du.$$

This is a common integral¹¹ and the solution is:

$$= \frac{u|u|}{2} + C$$

If u is replaced with $x-y$ and integrated from 0 to 1 then the following would be obtained:

$$\begin{aligned} & \left. \frac{(x-y)|x-y|}{2} \right|_0^1 \\ &= \frac{(1-y)|1-y|}{2} - \frac{(0-y)|0-y|}{2} \\ &= \frac{(1-y)|1-y|}{2} + \frac{y^2}{2} \\ &= \frac{(1-y)|1-y| + y^2}{2} \end{aligned}$$

Now that the inside integral is calculated it is time to calculate the outer integral.

$$\begin{aligned} & \int_0^1 \int_0^1 |x-y| dx dy \\ &= \int_0^1 \frac{(1-y)|1-y| + y^2}{2} dy \end{aligned}$$

Because the expression will be integrated from 0 to 1, the expression $|1-y|$ will never be negative; therefore, the absolute value can be eliminated and the expression can be simplified to:

$$\frac{1}{2} \int_0^1 (1-y)^2 + y^2 dy$$

¹¹ "What Is the Integral of the Absolute Value of X." Cuemath, <https://www.cuemath.com/questions/what-is-the-integral-of-the-absolute-value-of-x/>.

$$= \frac{1}{2} \left(\int_0^1 (1-y)^2 dy + \int_0^1 y^2 dy \right)$$

First, the integral on the right has to be calculated:

$$\int_0^1 y^2 dy$$

$$= \frac{y^3}{3} \Big|_0^1$$

$$= \frac{1^3}{3} - \frac{0^3}{3}$$

$$= \frac{1}{3}$$

So now the expression is:

$$\frac{1}{2} \left(\int_0^1 (1-y)^2 dy + \int_0^1 y^2 dy \right)$$

$$= \frac{1}{2} \left(\int_0^1 (1-y)^2 dy + \frac{1}{3} \right)$$

Now, the remaining integral has to be solved:

$$= \int_0^1 (1-y)^2 dy$$

$$= \int_0^1 1 - 2y + y^2 dy$$

$$= \int_0^1 1 dy - \int_0^1 2y dy + \int_0^1 y^2 dy$$

$$\begin{aligned}
&= y \Big|_0^1 - y^2 \Big|_0^1 + \frac{y^3}{3} \Big|_0^1 \\
&= (1 - 0) - (1 - 0) + \left(\frac{1}{3} - \frac{0}{3} \right) \\
&= 1 - 1 + \frac{1}{3} \\
&= \frac{1}{3}
\end{aligned}$$

This result has to be substituted back into the main equation:

$$\frac{1}{2} \left(\frac{1}{3} + \frac{1}{3} \right) = \frac{1}{3}$$

4. Opposite Sides

Now that the average distance of two random points on adjacent sides and on the same side have been calculated, it is time to calculate the average distance between two random points when they are at opposite sides. If the y component of the first point is called y and the y component of the second point is called x than according to the Pythagorean Theorem the distance between these points should be:

$$\sqrt{1 + (x - y)^2}$$

Then the average distance between the randomly selected points would be:

$$\begin{aligned}
&\frac{1}{(1 - 0)} \frac{1}{(1 - 0)} \int_0^1 \int_0^1 \sqrt{1 + (x - y)^2} dx dy \\
&= \int_0^1 \int_0^1 \sqrt{1 + (x - y)^2} dx dy
\end{aligned}$$

The inside integral has to be solved first, therefore, substitution needs to take place:

$$u = x - y$$

$$du = dx$$

$$\text{when } x = 1 \rightarrow u = 1 - y$$

$$\text{when } x = 0 \rightarrow u = -y$$

Which would yield:

$$\int_{-y}^{1-y} \sqrt{u^2 + 1} du$$

Using E2:

$$u = \tan(v)$$

$$du = \sec^2(v)dv$$

$$\text{when } u = 1 - y \rightarrow v = \arctan(1 - y)$$

$$\text{when } u = -y \rightarrow v = \arctan(-y)$$

The expression has come down to:

$$\int_{\arctan(-y)}^{\arctan(1-y)} \sqrt{\tan^2(v) + 1} \sec^2(v) dv$$

$$\int_{\arctan(-y)}^{\arctan(1-y)} \sec^3(v) dv$$

E3 needs to be used:

$$\frac{\sec^2(v) \sin(v)}{2} \Big|_{\arctan(-y)}^{\arctan(1-y)} + \frac{1}{2} \int_{\arctan(-y)}^{\arctan(1-y)} \sec(v) dv$$

First the integral on the right has to be calculated using E4 and E5:

$$\begin{aligned} & \ln |\tan(v) + \sec(v)| \Big|_{\arctan(-y)}^{\arctan(1-y)} \\ &= \ln \left| (1-y) + \sqrt{1 + (1-y)^2} \right| - \ln \left| -y + \sqrt{1 + (-y)^2} \right| \\ &= \ln \left| (1-y) + \sqrt{y^2 - 2y + 2} \right| - \ln \left| -y + \sqrt{1 + y^2} \right| \end{aligned}$$

So now the expression is:

$$\begin{aligned} & \frac{\sec^2(v) \sin(v)}{2} \Big|_{\arctan(-y)}^{\arctan(1-y)} + \frac{1}{2} \int_{\arctan(-y)}^{\arctan(1-y)} \sec(v) dv \\ &= \frac{\sec^2(v) \sin(v)}{2} \Big|_{\arctan(-y)}^{\arctan(1-y)} + \frac{1}{2} \left(\ln \left| (1-y) + \sqrt{y^2 - 2y + 2} \right| - \ln \left| -y + \sqrt{1 + y^2} \right| \right) \end{aligned}$$

Now, left side of the equation has to be solved:

$$\begin{aligned} & \frac{\sec^2(v) \sin(v)}{2} \Big|_{\arctan(-y)}^{\arctan(1-y)} \\ &= \frac{\sec(v) \tan(v)}{2} \Big|_{\arctan(-y)}^{\arctan(1-y)} \\ &= \frac{\sec(\arctan(1-y)) (1-y)}{2} - \frac{\sec(\arctan(-y)) (-y)}{2} \end{aligned}$$

$$\begin{aligned}
&= \frac{\sqrt{1 + (1 - y)^2} (1 - y)}{2} - \frac{\sqrt{1 + y^2} (-y)}{2} \\
&= \frac{\sqrt{y^2 - 2y + 2} (1 - y) + y\sqrt{1 + y^2}}{2}
\end{aligned}$$

Now the final expression is:

$$\begin{aligned}
&\int_0^1 \left(\frac{\sqrt{y^2 - 2y + 2} (1 - y) + y\sqrt{1 + y^2}}{2} + \frac{1}{2} (\ln |(1 - y) + \sqrt{y^2 - 2y + 2}| - \ln |-y + \sqrt{1 + y^2}|) \right) dy \\
&= \int_0^1 \frac{\sqrt{y^2 - 2y + 2} (1 - y) + y\sqrt{1 + y^2} + (\ln |(1 - y) + \sqrt{y^2 - 2y + 2}| - \ln |-y + \sqrt{1 + y^2}|)}{2} dy \\
&= \mathbf{1.0766}
\end{aligned}$$

The integral was too complicated to calculate without the help of a calculator. So, a calculator was used to evaluate it.

5.Averaging all the possible outcomes:

Now that the average distance between two random points on adjacent sides, on opposite sides, and on the same side are known, all of them should be averaged to find the distance between two randomly selected points on the perimeter of a unit square. As mentioned before, the result from the adjacent sides has to be multiplied by two because there are two possible ways of obtaining two points on adjacent sides.

$$\begin{aligned}
&\frac{0.7652 \times 2 + \frac{1}{3} + 1.0766}{4} \\
&\approx \mathbf{0.7351}
\end{aligned}$$

6. Double-checking the result using computation:

I wrote a small code that takes two random points on the perimeter of a 1 by 1 square and calculates the distance between them. The program then repeats this process a thousand, a hundred thousand, and a million times and adds the results together. It then takes the average of the results by dividing the total value to the trial number. The result after a million tries, which is 0.7350, is accurate with the result found by integration for up to 3 significant figures. Which is beyond enough to confirm that the calculations were indeed correct.

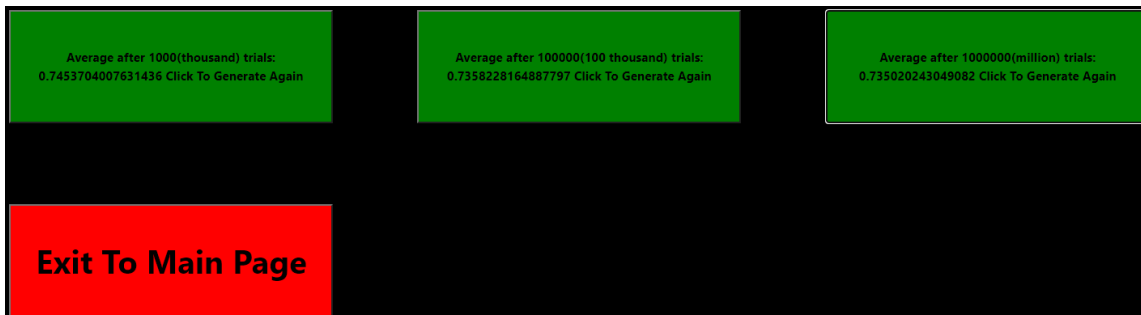


Figure 13. The output of the program used to randomly generate two points on a unit square and take the distance between them

```
for (var i = 0; i < trial; i++) {
  ran = Math.ceil(Math.random() * 4);
  ran2 = Math.ceil(Math.random() * 4);
  if (ran == 1) {
    y1 = Math.random();
    x1 = 0;
  }
  else if (ran == 2) {
    y1 = 1;
    x1 = Math.random();
  }
  else if (ran == 3) {
    y1 = Math.random();
    x1 = 1;
  }
  else if (ran == 4) {
    y1 = 0;
    x1 = Math.random();
  }

  if (ran2 == 1) {
    y2 = Math.random();
    x2 = 0;
  }
  else if (ran2 == 2) {
    y2 = 1;
    x2 = Math.random();
  }
  else if (ran2 == 3) {
    y2 = Math.random();
    x2 = 1;
  }
  else if (ran2 == 4) {
    y2 = 0;
    x2 = Math.random();
  }

  total += Math.sqrt(Math.pow((x1-x2), 2) + Math.pow((y1-y2), 2));
}
total /= trial;
```

Figure 14. A portion of the code used to generate 2 random points on the unit square

(Link to the full code: https://github.com/DurukanBTN/Extended-Essay/blob/main/Average_Length_Between_Two_Points_On_A_Unit_Square.html)

Conclusion

The journey that began with me stumbling upon the Bertrand Paradox ended with the calculation of the average distance between two randomly selected points on a unit square. In the end it was concluded that the average chord length of a unit circle is approximately 1.2732 units and that the average distance between two randomly selected points on the perimeter of a unit square is 0.7351 units. Throughout the essay, calculation and significance of average distances in geometrical shapes have been discussed. The sheer scarceness of resources on the topic came as a surprise to me. But this scarceness pushed me to develop individual mathematical thinking and problem-solving skills, so in a sense it was an unexpected positive aspect.

This research taught me that seemingly simple questions can make you dive deep into complex calculations and various topics of mathematics. I needed to learn topics that were not even in my yearly curriculum like double integrals. I learned how to effectively research on mathematical subjects and apply the knowledge I acquired to my calculations and approaches towards my problems. I also got experienced in conjoining my knowledge in computer science (to check my findings by using average of randomly generated chords) and mathematics, or in sub-topics of mathematics like integration and geometry.

This research can be thought as the basis of a future larger scale investigation. A larger variety of polygons can be investigated like pentagons or hexagons. This way, the question of whether a pattern in terms of the distance between randomly selected points on the perimeter is prominent amongst polygons can be answered.

Bibliography

- “Bertrand Paradox (Probability).” *Wikipedia*, Wikimedia Foundation, 27 Dec. 2021, [https://en.m.wikipedia.org/wiki/Bertrand_paradox_\(probability\)](https://en.m.wikipedia.org/wiki/Bertrand_paradox_(probability)).
- “Calculus Early Transcendentals: Integral & Multi-Variable Calculus for Social Sciences.” *Double Integrals: Volume and Average Value*, https://www.sfu.ca/math-coursenotes/Math%20158%20Course%20Notes/sec_VolumeAvgHeight.html.
- “Calculus.” *Average Value of a Function*, <https://math24.net/average-value-function.html>.
- “Centroid of a Triangle.” *Centroid*, https://www.varsitytutors.com/hotmath/hotmath_help/topics/centroid#:~:text=The%20centroid%20theorem%20states%20that,midpoint%20of%20the%20opposite%20side.
- “Fundamental Theorems of Calculus.” *From Wolfram MathWorld*, <https://mathworld.wolfram.com/FundamentalTheoremsOfCalculus.html>.
- “Half Angle Formulas - Examples: Half Angle Identities: Proof.” *Cuemath*, <https://www.cuemath.com/half-angle-formula/>.
- “How to Derive Power Reducing Formula for $\int(\sec^N x)dx$ and $\int(\tan^N x)dx$ for Integration?: Socratic.” *Socratic.org*, 19 May 2018, <https://socratic.org/questions/how-to-derive-power-reducing-formula-for-int-sec-nx-dx-and-int-tan-nx-dx-for-int>.
- “The Law of Cosines”, <https://www.mathsisfun.com/algebra/trig-cosine-law.html>.
- “Rings of Saturn.” *Wikipedia*, Wikimedia Foundation, 25 Feb. 2022, https://en.wikipedia.org/wiki/Rings_of_Saturn#Major_subdivisions.
- “Rings of Saturn.” *Wikipedia*, Wikimedia Foundation, 25 Feb. 2022, https://en.wikipedia.org/wiki/Rings_of_Saturn#Physical_characteristics.
- “What Is the Integral of the Absolute Value of X.” *Cuemath*, <https://www.cuemath.com/questions/what-is-the-integral-of-the-absolute-value-of-x/>.