Scatter Distribution of Balls after Colliding with a Single Peg

Introduction

The scatter pattern of particles from a target has always been a point of interest in physics. It helped to establish the Rutherford model of the atom and the basis of many microscopy techniques is the analysis of photon scatter patterns [1-2].

In Rutherford's famous gold foil experiment (or the Geiger-Marsden experiment), a stream of alpha particles was launched at a gold foil, and Rutherford mathematically modeled the scatter pattern of the particles [1]. Fig. 1 below shows its apparatus.

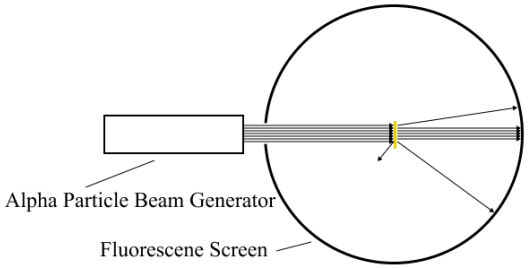

Figure 1: Simplified geometry of Rutherford's experiment.

A circular, fluorescent screen was set up around the gold foil to detect the scatter pattern. As the particles encountered the gold foils, most went through, some were slightly deflected, and others bounced back [1]. This helped Rutherford to conclude that an atom consists mostly of empty spaces.

We wished to conduct a Monte-Carlo simulation on a similar apparatus. We made the simplifying assumption that both the alpha particles and the gold foil behave like circular, neutral point charges, with some radius. In essence, the alpha particles became balls, and the gold foil became a peg.

This begs the research question: what is the probability density function (PDF) of the scatter pattern of a ball colliding with a circular peg? We aim to deterministically derive a PDF that describes the scatter pattern, then compare it with the numerical results generated by the Monte Carlo simulation of a ball colliding with a peg.

Methodology

The apparatus is constructed in Python, and all objects are mathematical objects plotted on a cartesian plane. This simulation ran via Jupyter Notebook, a web-based Python code compiler and development environment. We imported Python's default math module and NumPy to perform simple mathematical operations. We also imported Python's default random package for its uniform random number generator. The geometry of the apparatus is somewhat similar to Rutherford's experiment, as shown in Fig. 2a below.

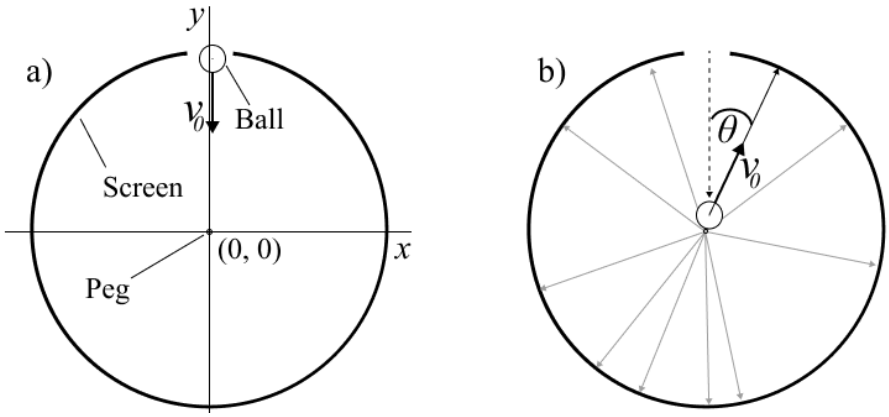

Figure 2: Geometry of the simulation apparatus.

The peg with a radius of $r = 0.001$, is at the origin. The ball has a radius of $R = 0.2$. We assumed both the ball and the peg are ideal: they are perfectly circular, have uniform density, do not deform upon impact, and their collisions are elastic.

The ball is assumed to enter the apparatus at some velocity v_0, moving towards the peg. The ball collides with the peg elastically. It reflects off the peg at some angle θ. The gray arrows in Fig. 2b above show some example trajectories of the ball if we repeat this simulation. We set up a symbolic screen to detect the locations of the balls. We divided the screen into 100 "chutes" with identical sizes, as shown in Fig. 3 below, where the size of each chute is $\frac{\pi}{50}$ radians.

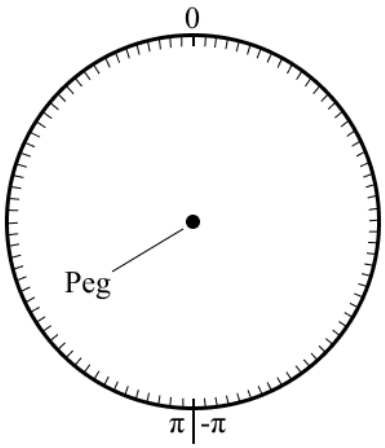

Figure 3: Location of the figurative chutes on the screen.

We recorded the angle of reflection of the balls and stored them in an array. These angles are then sorted into bins and plotted in a histogram. The chutes in Fig. 3 each represent a bin in the output histogram. We then compared the output histogram with the deterministically derived PDF.

We built two versions of this model. Case 1 represents the version where the ball is launched at a starting point determined by a uniform random distribution and undergoes perfect elastic collision. The apparatus of Case 2 is identical to Case 1, except the initial x-coordinate of the ball is determined by a gaussian distribution.

Case 1

Fig. 4 below shows the collision geometry between the ball and the peg, using the assumptions we made for Case 1.

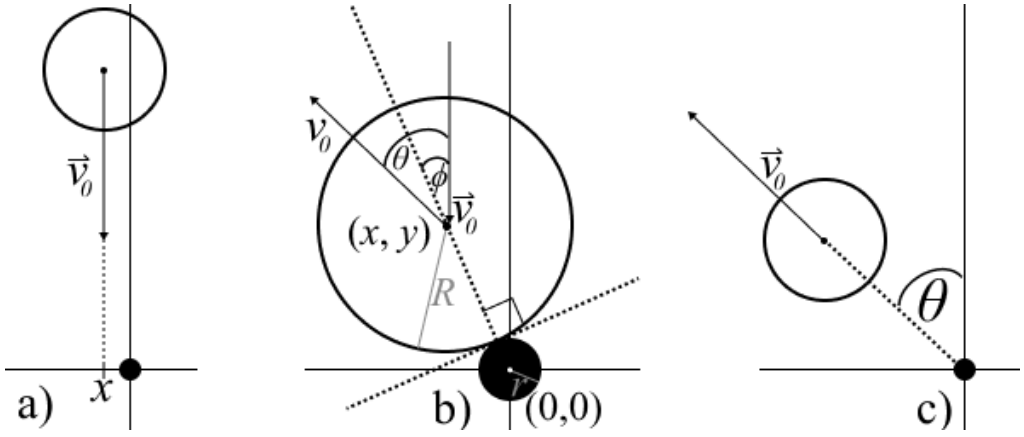

Figure 4: Collision geometry between the ball and the peg.

The peg with a radius of $r = 0.001$, is at the origin. The ball has a radius of $R = 0.2$. In Fig. 4a, the ball has an initial velocity of v_0 as it travels toward the peg. The x-coordinate of the ball's center point is at *x*, where *x* is generated via the uniform random number generator. We restricted its range to $(-R - r, R + r)$ to ensure the ball and the peg collide. When the ball collides with the peg, as shown in Fig. 4b, its center point is at (x, y). It undergoes perfect elastic collision, as shown in Fig. 4c. Therefore:

$$\tan\phi = \frac{x}{y} \tag{1}$$

where ϕ is the incident angle of the collision. Since the collision is perfectly elastic, the magnitude of the final velocity equals the initial velocity. Since the elastic collision follows the law of reflection, we conclude:

$$\theta = 2\phi$$

where θ is the angle of reflection, relative to the y-axis. To find θ in relation to *x*, we must find the value of *y* when the ball and the peg are tangent and not overlapping. We know that the distance between the center points of the ball and peg is $R + r$. Therefore, using the Pythagorean theorem, we find:

$$y = \sqrt{(R + r)^2 - x^2}$$

Using the value of *y* and the inverse tangent function, we find:

$$f(x) = \theta = 2\arctan\left(\frac{x}{\sqrt{(R+r)^2 - x^2}}\right) \tag{2}$$

Eq. 2 above gives the relationship between θ and x. We constructed the Monte-Carlo simulation using it. The code block below shows the main loop in the algorithm.

```
r = 0.001   #radius of the peg
R = 0.2     #radius of the ball
for i in range(1000000):
  x_ball = random.uniform(-R-r, R+r)
  y_ball = math.sqrt(collected_rad**2 - x_ball**2)
  ball_coord = [x_ball, y_ball]
  slope_norm = (ball_coord[1]-peg_coord[1])/(ball_coord[0]-peg_coord[0])
  phi = math.atan(1/slope_norm)
  reflected_angle = 2* phi
  reflected_angles.append(reflected_angle)
```

We defined the coordinates of the ball as lists with the suffix "coord" and all other computations are based on Eq. 2 above. The resultant angles θ of the balls after the collision were stored in a list called "reflected_angle".

Theory

Since we are trying to deterministically derive a PDF of θ using the PDF of *x*, we must first establish a relationship between them. If we recall Eq. 1:

$$\tan\phi = \frac{x}{y} \tag{1}$$
$$x = y\ \tan\frac{\theta}{2}$$

If there is a one-to-one relationship between θ and *x*, there must be a PDF of θ directly in terms of the PDF of x [4]. The relationship between the two PDF is described by Eq. 3 below [4]:

$$p(\theta) = U(f^{-1}(x))\left|\frac{d}{d\theta}f^{-1}(x)\right| \tag{3}$$

where $U(x)$ represents the PDF of *x*. We see that, if we attempt to map a range of θ onto *x*, as the location of this range changes (its width remains constant), the width of the range of *x* changes, as shown in Fig. 5 below. Therefore, to map θ onto *x*, we must differentiate $f^{-1}(x)$ to find its rate of change [4].

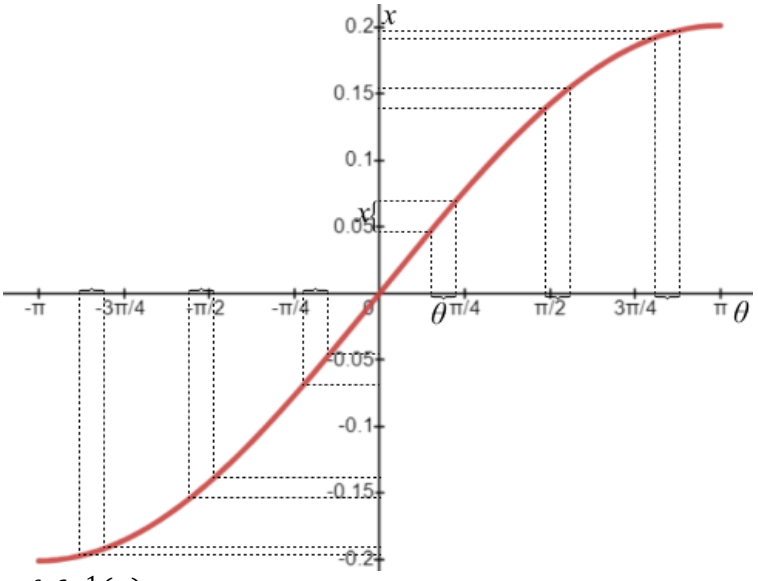

Figure 5: Graph of $f^{-1}(x)$.

The equation below is the PDF of *x*, a uniform random distribution that outputs in the range of (a, b) [5]:

$$U(x; a, b) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & elsewhere \end{cases}$$

Since Eq. 3 states that to find the PDF of θ, we must find both the inverse of $f(x)$ and its derivative, we obtained them through mathematically manipulating $f(x)$. Let us recall Eq. 2:

$$f(x) = \theta = 2\arctan\left(\frac{x}{\sqrt{(R+r)^2 - x^2}}\right) \tag{2}$$

Inverting Eq. 2, we get:

$$\tan^2\left(\frac{\theta}{2}\right) = \frac{x^2}{(R+r)^2 - x^2}$$

$$x^2 = (R+r)^2 \tan^2\left(\frac{\theta}{2}\right) - x^2\tan^2\left(\frac{\theta}{2}\right)$$

$$x = \frac{(R+r)\tan\left(\frac{\theta}{2}\right)}{\sec\left(\frac{\theta}{2}\right)} \tag{4}$$

Differentiating Eq. 4, we get:

$$\frac{d}{d\theta}\frac{(R+r)\tan\left(\frac{\theta}{2}\right)}{\sec\left(\frac{\theta}{2}\right)} = (R+r)\left(\frac{\left(\tan\left(\frac{\theta}{2}\right)\right)'\left(\sec\left(\frac{\theta}{2}\right)\right) - \left(\sec\left(\frac{\theta}{2}\right)\right)'\left(\tan\left(\frac{\theta}{2}\right)\right)}{\sec^2\left(\frac{\theta}{2}\right)}\right)$$

$$\frac{d}{d\theta}f^{-1}(x) = (R+r)\left(\frac{\sec^2\left(\frac{\theta}{2}\right) - \tan^2\left(\frac{\theta}{2}\right)}{2\sec\left(\frac{\theta}{2}\right)}\right) \tag{5}$$

Substituting Eq. 4 and Eq. 5 into Eq. 3, we find that the PDF, $p(\theta)$, between the domain of $(-\pi, \pi)$ is:

$$p(\theta) = \frac{\sec^2\left(\frac{\theta}{2}\right) - \tan^2\left(\frac{\theta}{2}\right)}{4\sec\left(\frac{\theta}{2}\right)} \tag{6}$$

Interestingly, the factor of $(R+r)$ disappeared, suggesting that the shape of the distribution is not dependent on the size of the ball and the peg. We overlaid this function on the normalized output histogram.

Results

After running 1,000,000 trials, we sorted all the values in the list "reflected angles" into 100 bins, where each bin has a bin width of $\frac{\pi}{50}$. We plotted them on a histogram. We then

normalized the histogram such that the total area underneath all the bins equals 1 by solving the following equation:

$$\gamma \sum_{i=1}^{m} bk_i = 1$$

where γ is the normalizing factor, m is the number of bins, b is the bin width, and k_i is the number of balls in each chute. After solving this equation, we find γ equals 1.59×10^{-5}. Fig. 6 below shows the normalized histogram overlaid with the derived PDF.

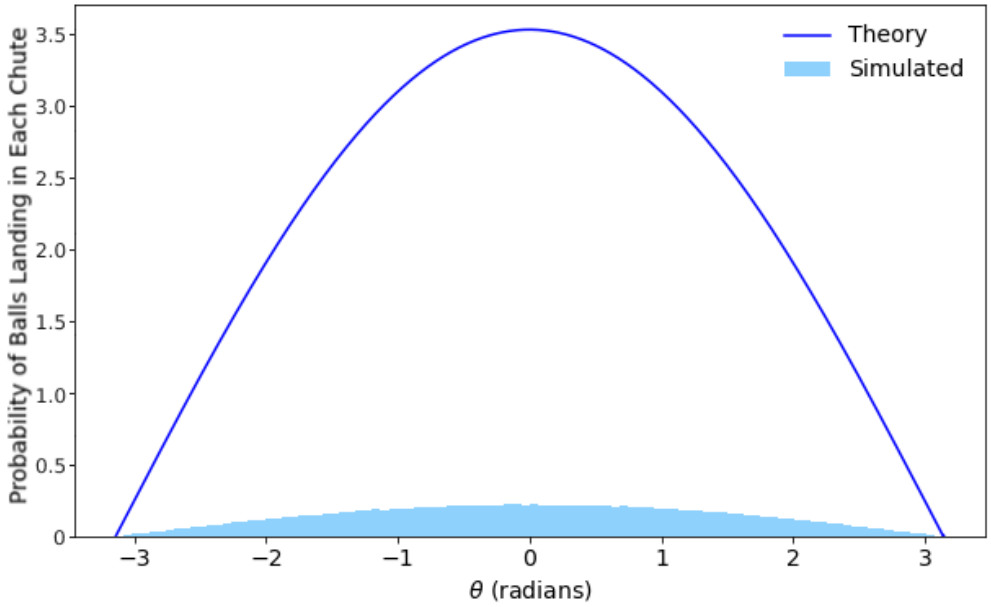

Figure 6: Simulation results for Case 1.

We see the simulation results and the theoretical PDF in Fig. 6 have a similar shape, but there is a huge difference in their scale factor. We also see that the area under the PDF must be larger than 1, because its maximum indicates the probability of $\theta = 0$ is 3.5, whereas a probability, by definition, must be in the range of 0 and 1. This suggests that the PDF is missing a scale factor. We denoted this scale factor as α. We were unsure why $p(\theta)$ is missing a scale factor because when we transformed its variable, we carried forward a scaling factor in $(R + r)$. Nonetheless,

since the domain of $p(\theta)$ is defined as $(-\pi, \pi)$, we tried to find α, the scale factor, using a finite integral:

$$\int_{-\pi}^{\pi} \alpha \cdot p(\theta) d\theta = 1$$

$$\alpha \int_{-\pi}^{\pi} \frac{\sec^2\left(\frac{\theta}{2}\right) - \tan^2\left(\frac{\theta}{2}\right)}{4\sec\left(\frac{\theta}{2}\right)} d\theta = 1$$

$$\alpha = \frac{1}{\int_{-\pi}^{\pi} \frac{\sec^2\left(\frac{\theta}{2}\right) - \tan^2\left(\frac{\theta}{2}\right)}{4\sec\left(\frac{\theta}{2}\right)} d\theta}$$

We could not evaluate this integral because it is intractable. Although all continuous functions can be differentiated, they may not be able to be integrated. The more complicated the function is, the less likely that it will have an antiderivative. Considering that $p(\theta)$ has a complex structure and no traditional, non-numerical techniques could be applied to evaluate it, we concluded that we could not find a numerical value of α through integrating $p(\theta)$.

Then, we tried to find it numerically by manually searching for the optimal value of α, the scale factor. We first tried a variety of values of α to fit the PDF to the histogram and found that the optimal value of α is in the range of 0.06 and 0.064. We then calculated the mean squared error (MSE) for each proposed value of α and conducted least squares fit. The equation below shows the equation for calculating MSE:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{k}_i - k_i\right)^2$$

where n is the total number of bins, k_i is the number of balls in each chute, and $\hat{k}_i$ is the predicted number of balls in each chute. We plotted the MSE values as a function of the scale factor, as shown in Fig. 7 below.

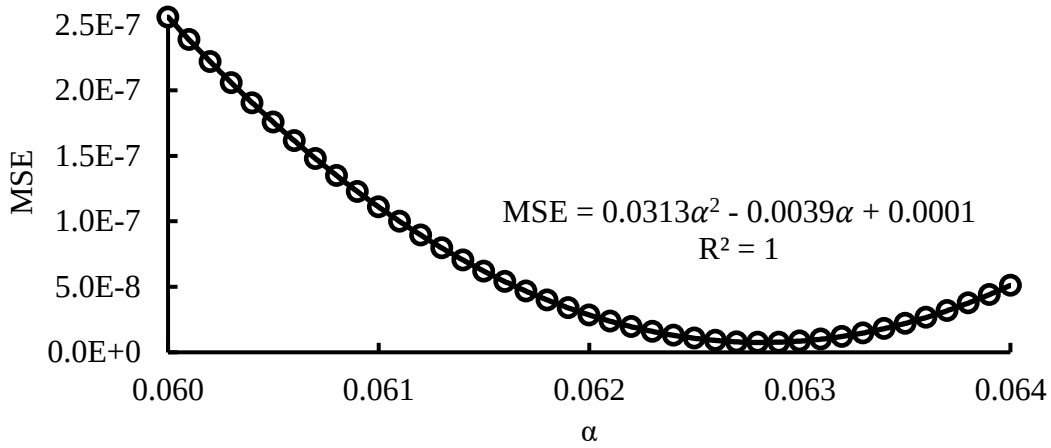

Figure 7: Scale factor versus its MSE.

The optimal value of α occurs when the absolute minimum point of the quadratic best fit line of Fig. 7 occurs. This happens when the derivative of the quadratic equals 0. Therefore:

$$MSE = 0.0313\alpha^2 - 0.0039\alpha + 0.0001$$
$$(MSE)' = 0.0626\alpha - 0.0039$$
$$0.0626\alpha - 0.0039 = 0$$
$$\alpha = 0.0628$$

Therefore, the optimal value of α is 0.0628. We then compared the scaled PDF with the simulated probability distribution, as shown in Fig. 8 below.

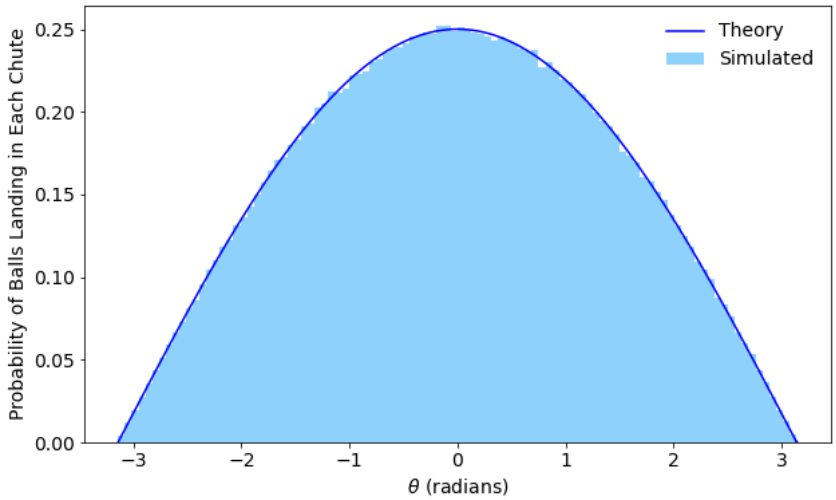

Figure 8: Comparison between the theoretical PDF and the simulation results.

We see that the scaled theoretical function is essentially a smoothed function of the simulation results, supporting the initial claim that the original theoretical PDF has the same shape

as the simulation outcome. Therefore, we concluded that the equation below is the approximate PDF of Case 1.

$$p_{1,exp}(\theta) = \alpha\left(\frac{\sec^2\left(\frac{\theta}{2}\right)-\tan^2\left(\frac{\theta}{2}\right)}{4\sec\left(\frac{\theta}{2}\right)}\right)$$

where α is 0.0628.

Case 2

We then replaced the uniform random number generator with a gaussian random number generator to generate the initial x-coordinate of the balls.

Theory

For Case 2, the geometry of the collision remained the same but the PDF for determining the initial x-coordinate of the ball has changed. The only change we made to the Python algorithm was to transform the uniform random generator into a normal random generator.

The Box-Mueller transformation [6] is a method to generate a pair of standard normal samples, Z_1 and Z_2, from a pair of uniform samples, U_1 and U_2. To understand the process of the Box-Mueller transformation, we must first digress into the properties of a bivariate normal distribution. A bivariate normal distribution is constructed by two normally distributed variables, Z_1 and Z_2 [7]. Each of them has a mean and a standard deviation. For a bivariate normal distribution whose variables have identical means and standard deviations, it is also called a circular normal distribution, because it has a perfectly symmetrical bell-shaped curve in three dimensions [7]. For distributions whose distributed variables have different means and standard deviations, their cross-sections are not radially symmetrical; their shape can be described as a slanted oval [7].

The bivariate normal density is a quantitative metric to measure the radial symmetry of the cross-sections, and it shows circular symmetry, as shown in Fig. 9 below [6].

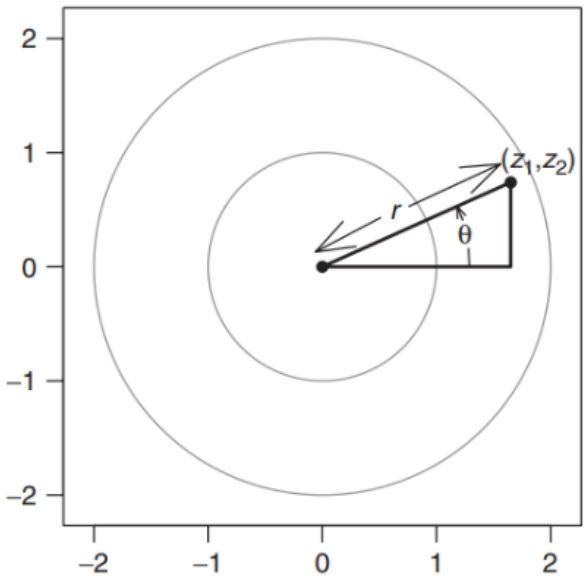

Figure 9: The bivariate normal density, represented in the polar form [6].

If we express the bivariate normal density in the polar form where its magnitude is r_i and its modulus is θ_i. Therefore, using this circular symmetry, we find [6]:

$$Z_1 = r_i \cos\theta_i = \sqrt{-2\log_{10}(1-U_2)}\cos(2\pi U_1)$$
$$Z_2 = r_i \sin\theta_i = \sqrt{-2\log_{10}(1-U_2)}\sin(2\pi U_1)$$

where Z_1 and Z_2 represent elements from two identical standard normal distributions. Even though for our simulation, we only needed one normal distribution, the Box-Mueller transformation is required to generate a second distribution to take advantage of the circular symmetry of a bivariate normal distribution. Nonetheless, it was not needed, so we only solved for Z_1 in our algorithm.

Since the distribution for Case 2 is not a standard normal distribution, we used the following equation to transform the distribution.

$$Z(x) = \sigma X + \mu$$

where σ is the standard deviation and μ is the mean. We defined σ as $\frac{R+r}{3}$ and μ as 0 because we want most of the values generated by the distribution to be within the range of $\pm(R+r)$ to ensure that the ball collided with the peg while making the distribution as wide as possible. The code block below shows the algorithm used to generate the x-coordinates of the ball.

```
r = 0.001    #radius of the peg
R = 0.2      #radius of the ball
sigma = (R+r)/3
mu = 0
Xmax = mu + 3*sigma
Xmin = mu - 3*sigma
Zmin = (Xmax - mu) / sigma
Zmax = (Xmin - mu) / sigma

for i in range(1000000):
  U1 = random.random()
  U2 = random.random()
  Ri = np.sqrt(-2 * np.log(U1))
  Thetai = 2 * np.pi * U2
  X = Ri * np.cos(Thetai)
  x_ball_coord = X*sigma + mu
  if (x_ball_coord > Zmax) or (x_ball_coord < Zmin):
     i = i-1
```

We also confirmed that the output from the code block above formed a normal distribution by running it in Jupyter Notebook. We plotted the output x-coordinates, as represented by the variable "x_ball_coord", in both a histogram with 100 bins (with a bin width of 0.002) and a quantile-quantile plot, as shown in Fig. 10 below.

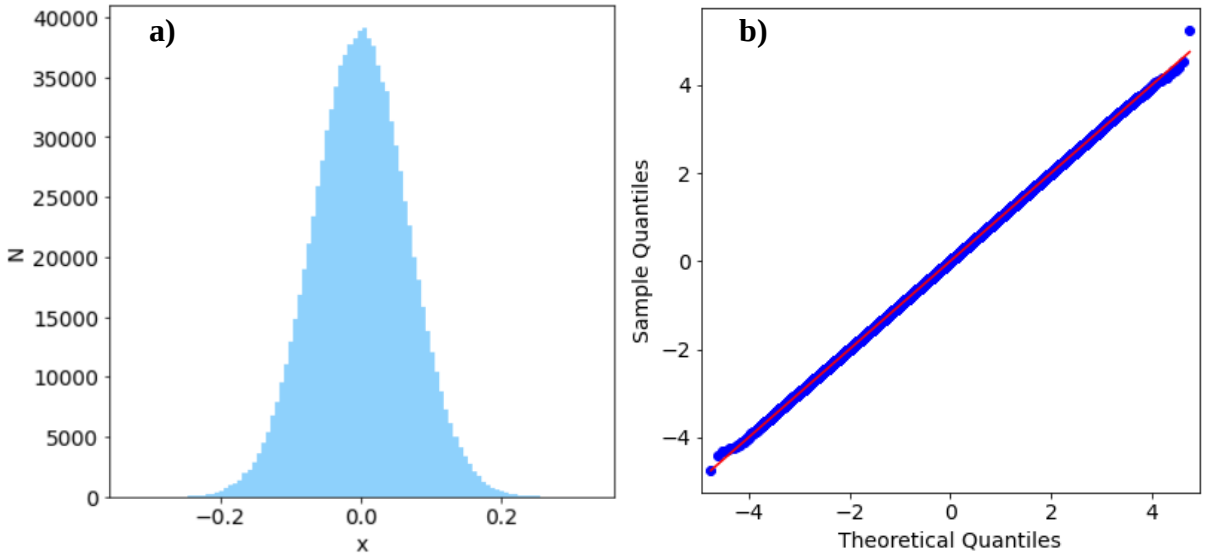

Figure 10: Distribution histogram and Quantile-Quantile plot of the output x-coordinate.

Fig. 10a shows a distribution that visually resembles a normal distribution. To confirm this, we then tested it by plotting it into a quantile-quantile plot. A quantile-quantile plot is a scatter plot that compares the sample and theoretical quantile [8]. If the distribution is normal, the data points

should approximately form a line, define by $y = x$, as shown in Fig. 10b. Therefore, the PDF of the x-coordinates is:

$$Z(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

where σ equals to $\frac{R+r}{3}$. Let us recall Eq. 3, which transforms $U(x)$ into $p(\theta)$:

$$p(\theta) = U(f^{-1}(x)) \left| \frac{d}{d\theta} f^{-1}(x) \right| \tag{3}$$

Therefore, for Case 2:

$$p_2(\theta) = Z(f^{-1}(x)) \left| \frac{d}{d\theta} f^{-1}(x) \right|$$

We must also recall both $f^{-1}(x)$ (see Eq. 4 below) and $\frac{d}{d\theta} f^{-1}(x)$ (see Eq. 5 below) because Eq. 3 uses them to transform $Z(x)$ into $p_2(\theta)$:

$$x = \frac{(R+r)\tan\left(\frac{\theta}{2}\right)}{\sec\left(\frac{\theta}{2}\right)} \tag{4}$$

$$\frac{d}{d\theta} f^{-1}(x) = (R+r)\left(\frac{\sec^2\left(\frac{\theta}{2}\right) - \tan^2\left(\frac{\theta}{2}\right)}{2\sec\left(\frac{\theta}{2}\right)}\right) \tag{5}$$

Substituting σ, Eq. 4, and Eq. 5 into Eq. 3, we find the PDF of Case 2 is:

$$p_2(\theta) = \frac{1}{\left(\frac{R+r}{3}\right)\sqrt{2\pi}} e^{-\frac{\left((R+r)\tan\left(\frac{\theta}{2}\right)\right)^2}{2\left(\frac{R+r}{3}\right)^2 \sec\left(\frac{\theta}{2}\right)}} \left| (R+r)\left(\frac{\sec^2\left(\frac{\theta}{2}\right) - \tan^2\left(\frac{\theta}{2}\right)}{2\sec\left(\frac{\theta}{2}\right)}\right)\right|$$

$$p_2(\theta) = \frac{3}{\sqrt{2\pi}} e^{-\frac{9\tan^2\left(\frac{\theta}{2}\right)}{2\sec\left(\frac{\theta}{2}\right)}} \left| \left(\frac{\sec^2\left(\frac{\theta}{2}\right) - \tan^2\left(\frac{\theta}{2}\right)}{2\sec\left(\frac{\theta}{2}\right)}\right)\right| \tag{7}$$

where the domain of $p_2(\theta)$ is $(-\pi, \pi)$. Note that since the mean value of the distribution is 0, all terms multiplied by μ went to 0. Interestingly, similar to the PDF of Case 1, the factor of $(R+r)$ disappeared, again suggesting that the shape of the distribution is not dependent on the size of the ball and the peg. We then compared Ep. 7, the PDF of Case 2 with the algorithm's output histogram.

Results

Fig. 11 below shows the simulation outcome after running 1,000,000 trials. Note the bin width in Fig. 11 is $\frac{\pi}{50}$. We see that the distribution of the balls in each chute somewhat resembles a normal distribution.

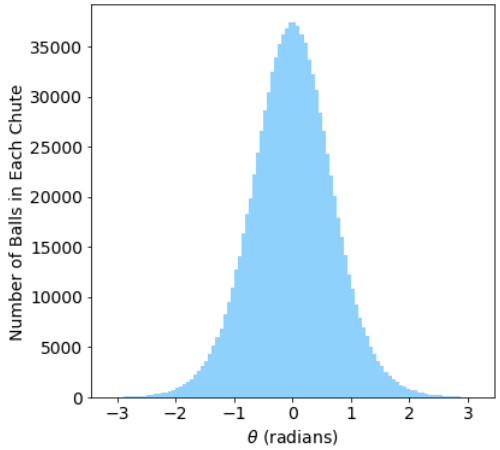

Figure 11: Simulation output distribution for Case 2.

We attempted to validate this hypothesis by plotting this distribution on a quantile-quantile plot, as shown in Fig. 12 below.

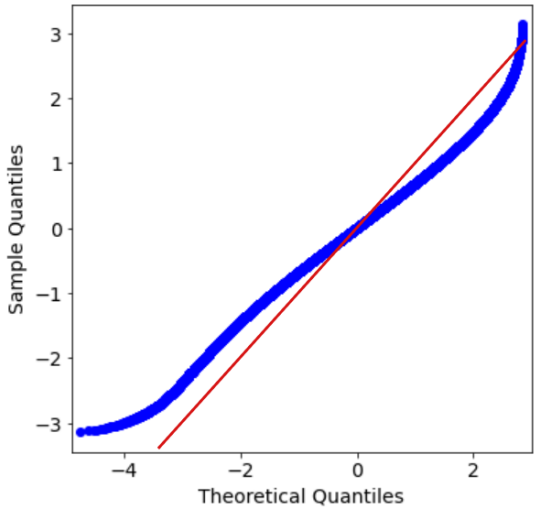

Figure 12: Quantile-Quantile plot of the output distribution.

We see that the line formed by the blue data points, representing the simulation outputs, is non-linear. There also are no segments of the blue line that overlaps with the red trendline, showing that the distribution is not normal. This is consistent with the theoretical PDF, $p_2(\theta)$, which is not

a gaussian, because the initial gaussian PDF, $Z(x)$, has been transformed by $f^{-1}(x)$ and $\frac{d}{d\theta}f^{-1}(x)$. Then, we overlaid the graph of $p_2(\theta)$ onto the normalized version of the histogram shown in Fig. 11, as shown in Fig. 13 below.

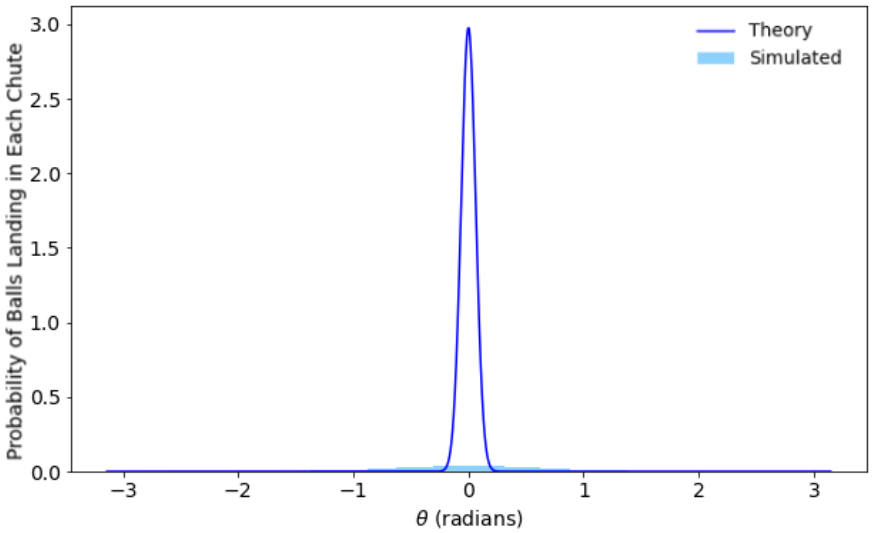

Figure 13: The comparison between the simulation output and the PDF.

We see that the theoretical PDF is on a completely different scale from the simulation outcomes, on both axes. However, the general shape of the theoretical distribution shows some semblance to the shape of the simulation outcome. Therefore, we decided to normalize it such that the area underneath it is 1, then compare it with the simulation outcome. We can express the scaled $p_2(\theta)$ as:

$$p_{2,exp}(\theta) = \alpha_2 p(\beta_2 \theta)$$

where α_2 is the vertical scale factor and β_2 is the horizontal scale factor. We decided to find β_2 numerically first, then integrate the PDF to find α_2, such that the area underneath it becomes 1.

We found $\beta_2$using the same method we used to find α for Case 1. We first narrowed the range of possible values of β_2 to between 0.09 and 0.1. Then, we calculated the MSE between the theoretical and the simulated distribution and plotted it as a function of β_2, as shown in Fig. 14 below.

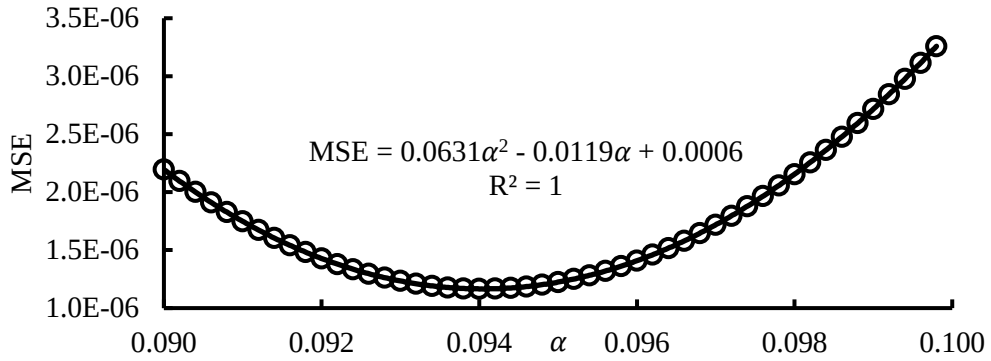

Figure 14: Horizontal scale factor versus its MSE.

We found the absolute minimum of the quadratic best fit line is:

$$MSE = 0.0631\alpha^2 - 0.0119\alpha + 0.0006$$

$$(MSE)' = 0.1262\alpha - 0.0119$$

$$0.1262\alpha - 0.0119 = 0$$

$$\alpha = 0.0942$$

Therefore, the optimal value of β_2 is 0.0942. We then set up the definite integral for finding the value of α_2 to normalize the area underneath $p_2(\theta)$:

$$\int_{-\pi}^{\pi} \alpha_2 p_2(\beta_2\theta)\, d\theta = 1$$

$$\alpha_2 \int_{-\pi}^{\pi} \frac{3}{\sqrt{2\pi}} e^{-\frac{9\tan^2\left(\frac{\beta_2\theta}{2}\right)}{2\sec\left(\frac{\beta_2\theta}{2}\right)}} \left|\left(\frac{\sec^2\left(\frac{\beta_2\theta}{2}\right)-\tan^2\left(\frac{\beta_2\theta}{2}\right)}{2\sec\left(\frac{\beta_2\theta}{2}\right)}\right)\right| d\theta = 1$$

$$\alpha_2 = \frac{1}{\frac{3}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{-\frac{9\tan^2\left(\frac{\beta_2\theta}{2}\right)}{2\sec\left(\frac{\beta_2\theta}{2}\right)}} \left|\left(\frac{\sec^2\left(\frac{\beta_2\theta}{2}\right)-\tan^2\left(\frac{\beta_2\theta}{2}\right)}{2\sec\left(\frac{\beta_2\theta}{2}\right)}\right)\right| d\theta}$$

We concluded that the integral cannot be evaluated to produce a numerical solution for α_2 because it is intractable. The function has too many components to be integrated through traditional, non-numerical techniques. Since we could not find α_2 through evaluating the integrand, we repeated the same process of finding the optimal value of β_2: we conducted least squares fit of the MSE to find α_2 and concluded that its optimal value is 0.1985. Fig. 15 below shows the comparison between the scaled PDF and the simulation output distribution.

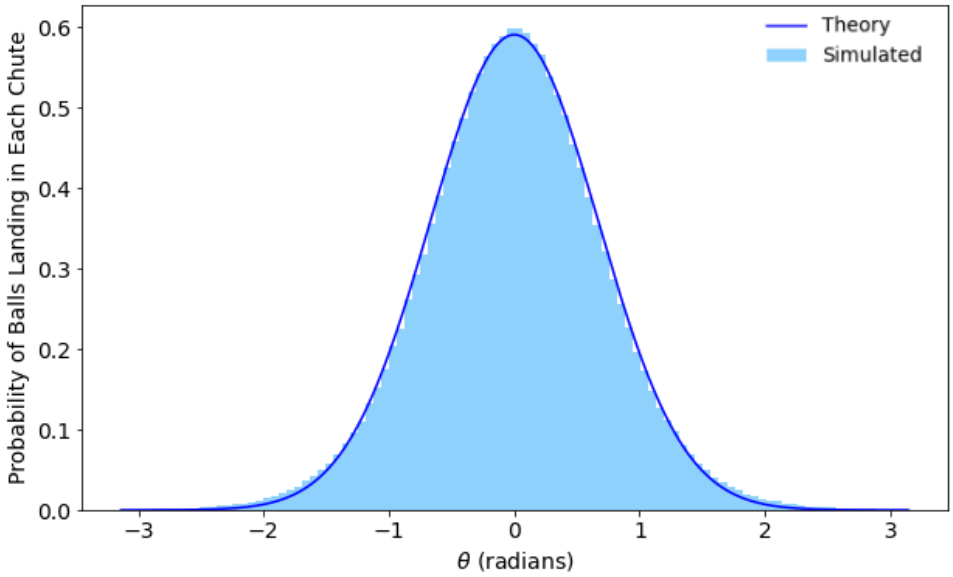

Figure 15: Comparison between the scaled PDF and the simulation output.

We see that the scaled PDF and the simulation output have the same shape. This suggests that our original PDF also has the same shape as the simulation output, but they are on a different scale. This is strange because when we transformed the variable of our PDF from *x* to θ, we carried forward the scaling factors. We can only speculate that when we found the inverse function of $f(x)$ or when we differentiated the inverse function, we must have changed the scale factors of the PDF. We believe this is unlikely to be caused by the differentiation of the inverse function because differentiation retains the scale of the original function. Therefore, through process of elimination, we conclude that this loss of the scale factor must be caused by inverting $f(x)$. Also, for Case 1, since the PDF of a uniform distribution is independent of the value of *x*, therefore, it is only vertically stretched. However, for Case 2, since the PDF of a gaussian is dependent on the value of *x*, the distribution is both horizontally and vertically stretched.

Nonetheless, the scaled, approximate PDF for Case 2 is:

$$p_{2,exp}(\theta) = \frac{3\alpha_2}{\sqrt{2\pi}} e^{-\frac{9\tan^2\left(\frac{\beta_2\theta}{2}\right)}{2\sec\left(\frac{\beta_2\theta}{2}\right)}} \left| \left(\frac{\sec^2\left(\frac{\beta_2\theta}{2}\right) - \tan^2\left(\frac{\beta_2\theta}{2}\right)}{2\sec\left(\frac{\beta_2\theta}{2}\right)} \right) \right|$$

where β_2 equals 0.0942 and α_2 equals 0.1985.

Conclusion

We constructed a Monte Carlo simulation for a ball colliding with a peg. In both cases, we deterministically derived the PDF of the ball's scatter pattern and conducted least squares fit of their MSE to numerically find their scale factors. From graphical evidence, we concluded that these derived PDFs fit the simulation outcome well. However, we could not explain why we needed to numerically determine scaling factors for the PDF, as the formula for transforming variables in distribution functions should have allowed us to carry forward the scaling factors. We also found that for both cases, after simplifying the PDF, the factor of $(R + r)$ disappeared. This suggests that the shape of the distribution is not dependent on the size of the ball and the peg.

Extensions

The original theoretical PDF of both Case 1 and Case 2 was a scaled version of their simulation outcomes. We suspect that a scaling factor is lost during the process of transforming the PDF in terms of x to a PDF in terms of θ. Specifically, we believe it is caused by inverting the function $f(x)$ because we eliminated the only other likely cause: the factor is lost when we differentiated $f^{-1}(x)$, since a function retains its scale factor when it is differentiated. However, this is only a speculation. Therefore, an extension could be to explore why we had to numerically determine the scale factor of the PDF and during which specific step did the PDF's scale change.

In addition, since the factor of $(R + r)$ canceled itself out in the PDF of both Case 1 and Case 2, this led us to conclude that the shape of the distribution is not dependent on the size of the ball and the peg. Therefore, an extension is to validate this theory by simulating the collision of a different-sized peg and a different-sized ball.

Also, we had to derive the scaling factors of the PDF numerically. We tried to derive them by evaluating a definite integral that symbolizes the area underneath the PDF, but for both cases, the integrands simplified into an indeterminate form, meaning they did not produce a numerical solution. Therefore, another improvement would be finding a way deterministically derive the scaling factors.

We made some starting assumptions, such as the collisions are elastic, the ball is not spinning, and the balls and peg do not deform upon impact. These assumptions idealized the model, which in turn simplified the mathematics associated with it. However, in reality, the collision between a ball and a peg would result in a loss in the kinetic energy of the ball. Since the conservation of energy must occur, the trajectory of the ball must change to accommodate for this energy loss. Therefore, we could include a factor that symbolizes the energy loss and compare the distribution produced by this model to both Case 1 and Case 2.

Another possible extension is to add a spin rate to the ball and compare the ball's trajectory to the idealized model. We could also make the balls and pegs charged objects and compare the resultant distributions to the current models. With these additions, we could also record the final velocity of the ball and its spin rate and compare these probability distributions as well.

There are also many other possible geometrical setups to a ball colliding with a peg. For example, what if the peg is elliptical and its major axis is not parallel to the horizontal axis? Will the distribution of the ball be skewed? Or will it somehow maintain its symmetry? There are an infinite number of possible geometries that can be explored, and all of them are possible further extensions.

Works Cited

[1] Bishop, C. B. (1990). Simulation of Rutherford's Experiment. *Journal of Chemical Education, 67*(10), 889. https://doi.org/10.1021/ed067p889.

[2] Bunaciu, A. A., Udriştioiu, E. Gabriela, & Aboul-Enein, H. Y. (2015). X-ray diffraction: Instrumentation and applications. *Critical Reviews in Analytical Chemistry*, *45*(4), 289–299. https://doi.org/10.1080/10408347.2014.949616.

[3] *Project Jupyter*. Jupyter. (2022). Retrieved December 3, 2022, from https://jupyter.org/.

[4] Siegrist, K. (2022, April 24). *3.7: Transformations of random variables*. Statistics LibreTexts. Retrieved October 28, 2022, from https://stats.libretexts.org/Bookshelves/Probability_Theory/Probability_Mathematical_Statistics_and_Stochastic_Processes_(Siegrist)/03%3A_Distributions/3.07%3A_Transformations_of_Random_Variables.

[5] Canavos, G. C. (1984). *Applied probability and statistical methods*. Little, Brown and Company.

[6] Scott, D. W. (2011). Box-Muller Transformation. *Wiley Interdisciplinary Reviews: Computational Statistics*, *3*(2), 177–179. https://doi.org/10.1002/wics.148.

[7] Pennsylvania State University. (2022). *4.2 - bivariate normal distribution: Stat 505*. Eberly College of Science. Retrieved December 3, 2022, from https://online.stat.psu.edu/stat505/lesson/4/4.2.

[8] Ford, C. (2015, August 26). *Understanding Q-Q Plots*. University of Virginia Library. Retrieved October 29, 2022, from https://data.library.virginia.edu/understanding-q-q-plots/.