

S2 Coursework Report

Zhimei Liu (zl474)

March 28, 2025

Word count: 1636

Declaration: I, Zhimei Liu, hereby declare that the work presented in this report and the GitLab repository is entirely my own. I have only used ChatGPT for some debugging to speed up the process.

Contents

0	Introduction	2
1	Loading the data	3
2	The Model	3
3	The Likelihood Function	5
3.1	First Choice of Covariance Matrix	6
3.2	Second Choice of Covariance Matrix	7
4	Computing the Gradient of Log-likelihood	8
4.1	Gradient for the First Model	9
4.2	Gradient for the Second Model	10
4.3	Automatic Differentiation with JAX	11
5	Maximum Likelihood Estimators	12
6	Sample the Posterior Distribution using HMC	13
7	Analysis and Comparison between Two Models	14

0 Introduction

The *Antikythera mechanism* is a mechanical calculator discovered from a shipwreck close to the Greek island of Antikythera in 1901. It is believed that it can be used to predict astronomical phenomena. However, its original purpose and functionality are not yet fully understood. The key to understanding its purpose is to count the number of holes in it.

The Antikythera mechanism included a calendar ring with holes punched around its circumference (see Figure 1). However, only a part of the full ring survives, and it is fractured into several sections. Careful measurements of the positions of closely spaced holes beneath the ring have been made by Budiselic et al. These holes are thought to have been used to rotationally align the calendar ring, and their number is important for understanding the ring's function.

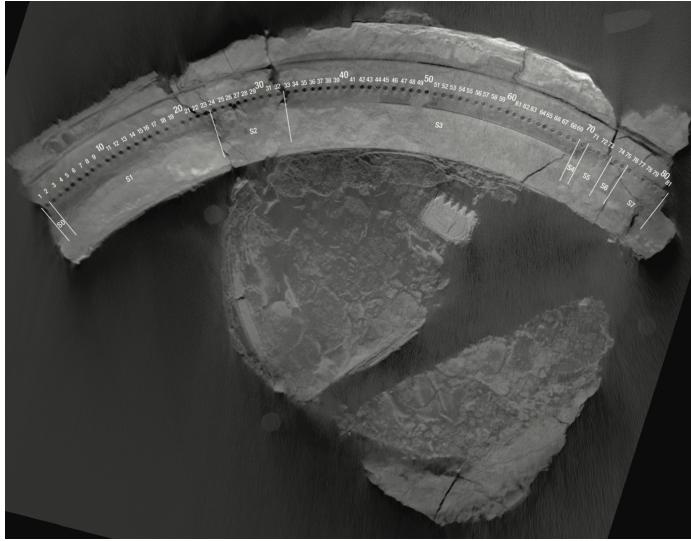


Figure 1: The calendar ring. Figure reproduced from Budiselic et al. (2020).

The goal of this report is to infer the number of holes that were present in the complete calendar ring, given the X-ray measurements, some reasonable assumptions and using Bayesian inference with Hamiltonian Monte Carlo (HMC).

1 Loading the data

This section answers the part (a) of the coursework.

After loading the data from Ref [1], I plotted the measured hole locations in the x - y plane. There are 81 holes, dividing into 8 sections (see Figure (2)).

2 The Model

This subsection answers the part (b) of the coursework.

The method presented in Ref [2] will be used in this section to model the hole positions. The following assumptions will be made:

1. Originally, N holes were **evenly spaced** around a circle of radius r .
2. Today, the ring exists as s fractured sections that are slightly displaced and rotated relative to each other.

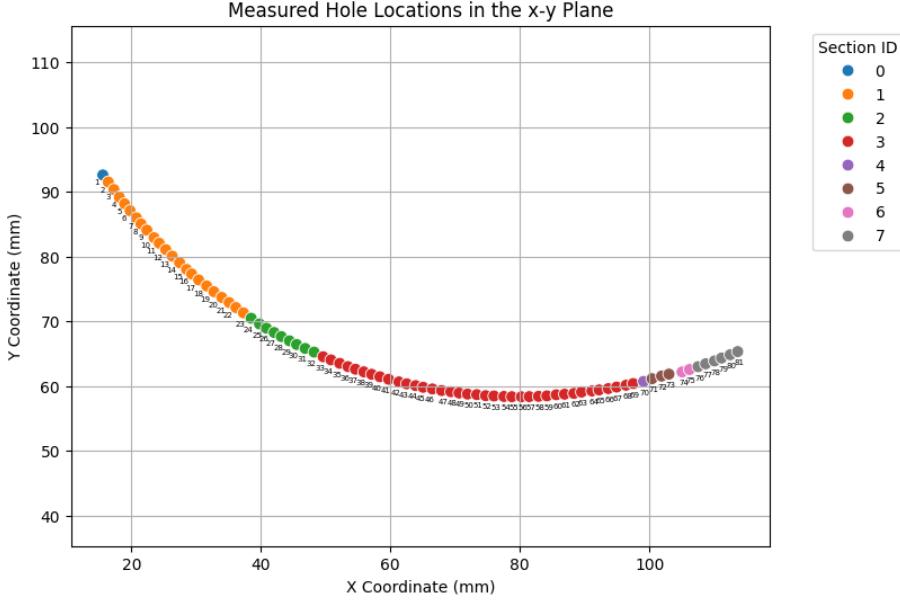


Figure 2: The measured hole locations in the x - y plane from the X-ray image, labelled by the hole ID and fractured section ID.

- 3. The relative shifts (translations) and rotations of each fractured section are **unknown**.
- 4. There are no internal distortions of the sections and they lie in the (x, y) plane defined by the dataset.

Let $\mathbf{d}_i = (x_i, y_i)$, $1 \leq i \leq n = 81$ denote the Cartesian coordinates of the 81 points from the X-ray image. Let $\mathbf{r}_0 = (x_0, y_0)$ denote the centre of the original, unbroken ring, while $\mathbf{r}_{0j} = (x_{0j}, y_{0j})$, $0 \leq j \leq s - 1$, where $s = 8$, denote the centre of each individual fractured section. These eight section centres are not exactly the same, but we expect them to be relatively close to each other near the original ring's centre, because there are only slight displacements and rotations for each section.

To measure the amount of rotation after the break from each section, let's denote α_j as the angular position of the first hole (the measured hole with ID=1) of the full circle, when this section j is extended to that point. Since each section shows minimal relative rotation, we expect these α_j values to be similar. With this definition of α_j , the angular position of the i th hole in the j th section with respect to its arc-centre is therefore

$$\phi_{ij} = 2\pi \frac{(i-1)}{N} + \alpha_j , \quad 1 \leq i \leq 81, 0 \leq j \leq 7. \quad (1)$$

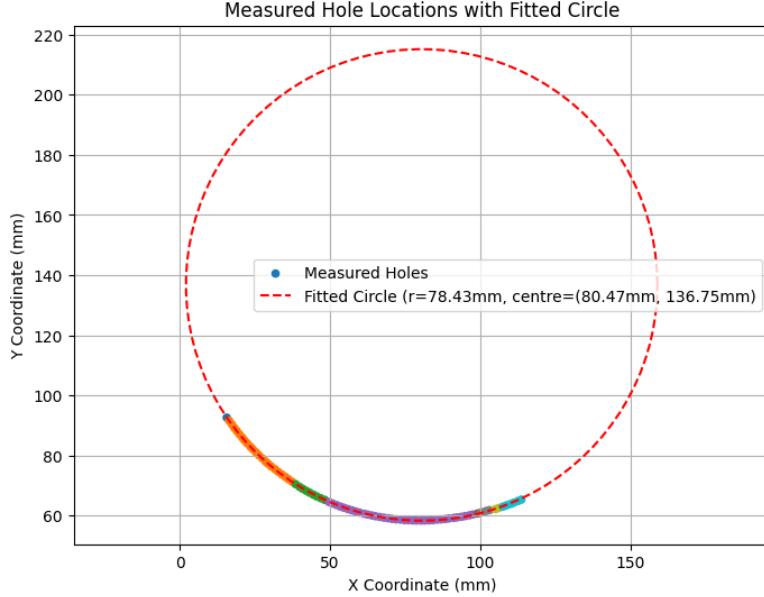


Figure 3: The fitted circle of the measure 81 holes.

To get started, let's fit a circle to the 81 holes. I obtained a fitted circle with radius 78.43 mm, centre at (80.47 mm, 136.75 mm), as shown in Figure (3). The holes from different fractured sections are labelled with different points. The fitted results give us a rough estimates of the radius and the location of the centre, which can be used for later comparison.

3 The Likelihood Function

This subsection answers the part (c) of the coursework.

To estimate the parameters of the model and their uncertainty, we first need to write down the likelihood of the model, so that we can write down the posterior distribution of the parameters using Bayesian's Theorem:

$$p(\boldsymbol{\theta}|D) = \frac{\pi(\boldsymbol{\theta})\mathcal{L}(D|\boldsymbol{\theta})}{p(D)} \propto \pi(\boldsymbol{\theta})\mathcal{L}(D|\boldsymbol{\theta}), \quad (2)$$

where $D = \{\mathbf{d}_i | i = 1, \dots, 81\}$ and $\boldsymbol{\theta}$ is a parameter vector that contains all the parameters of the model.

Let's assume a Gaussian likelihood function for each hole location and assume that the error in the placement of different holes is independent. There-

fore, we can write the overall likelihood function as a product of the individual likelihood function for each hole \mathbf{d}_i :

$$\mathcal{L}(\mathbf{d}_i|\boldsymbol{\theta}) = \frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{d}_i - \mathbf{m}_i)^T \boldsymbol{\Sigma}^{-1} (\mathbf{d}_i - \mathbf{m}_i)\right), \quad (3)$$

where $\boldsymbol{\Sigma}$ is a 2-dimensional covariant matrix and \mathbf{m}_i denotes the model prediction (a function of $\boldsymbol{\theta}$) for the location of hole i .

Let's write out the explicit form of the likelihood function for different choices of $\boldsymbol{\Sigma}$.

3.1 First Choice of Covariance Matrix

Firstly, we consider a 2-dimensional isotropic Gaussian distribution with standard deviation σ , i.e.,

$$\boldsymbol{\Sigma} = \sigma^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (4)$$

Then Eq. (3) becomes

$$\mathcal{L}(\mathbf{d}_i|\boldsymbol{\theta}_i) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{d}_i - \mathbf{m}_i)^T (\mathbf{d}_i - \mathbf{m}_i)\right). \quad (5)$$

Let's count the total number of parameters in this model first: the ring radius r , the number of holes in the original ring N , σ in the covariance matrix, and 3 parameters $(x_{0j}, y_{0j}, \alpha_j)$ for each section of the ring. However, Section 0 and 4 each contain only one hole. These sections do not constrain any of our parameters and have therefore been omitted from the analysis. Hence, there are 21 unknown quantities constrained by the data.

To express \mathbf{m}_i explicitly in terms of the parameters, suppose \mathbf{d}_i is the i th hole in section j , where its centre is $\mathbf{r}_{0j} = (x_{0j}, y_{0j})$. From Eq. (1), we know that this hole subtends an angle ϕ_{ij} about the centre. See Fig. (4) for more detail. Therefore, the predicted i th hole location is

$$\mathbf{m}_i = \begin{pmatrix} x_{0j} + r \cos \phi_{ij} \\ y_{0j} + r \sin \phi_{ij} \end{pmatrix}. \quad (6)$$

We can see that \mathbf{m}_i is a function of 5 parameters, i.e., $\mathbf{m}_i = \mathbf{m}_i(x_{0j}, y_{0j}, \alpha_j, r, N)$.

Hence, the overall likelihood function is

$$\begin{aligned} \mathcal{L}(D|\boldsymbol{\theta}) &= \frac{1}{(2\pi\sigma^2)^n} \prod_{i=1}^n \exp\left(-\frac{1}{2\sigma^2}(\mathbf{d}_i - \mathbf{m}_i)^T (\mathbf{d}_i - \mathbf{m}_i)\right) \\ &= \frac{1}{(2\pi\sigma^2)^n} \prod_{j=1}^{s-1} \prod_{i \text{ in } j} \exp\left(-\frac{(x_i - x_{0j} - r \cos \phi_{ij})^2 + (y_i - y_{0j} - r \sin \phi_{ij})^2}{2\sigma^2}\right) \end{aligned} \quad (7)$$

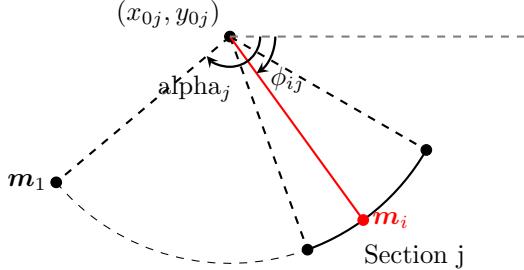


Figure 4: Diagram showing the relationship between points \mathbf{m}_j and the angles ϕ_{ij} and α_j .

where now $n = 79$ as we have ignored the single hole from Section 0 and 4.

3.2 Second Choice of Covariance Matrix

Secondly, we consider a 2-dimensional Gaussian distribution with principle axes aligned with the radial and tangential direction for each hole and with standard deviations σ_r and σ_t , describing the radial and tangential accuracies of the hole placement. In other words, the covariance matrix has the form

$$\Sigma = \begin{pmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_t^2 \end{pmatrix}. \quad (8)$$

We further assume that σ_r , σ_t are the same for every hole and that the hole-to-hole errors are uncorrelated.

With this choice of covariance matrix, there is a total of 22 unknown parameters to estimate.

Let's define the radial and tangential unit vector at each hole location. Fig. (5) labels the radial and tangential unit vectors $\hat{\mathbf{r}}_{ij}$ and $\hat{\mathbf{t}}_{ij}$ for the i th hole in Section j . With some algebra, we have

$$\begin{aligned} \hat{\mathbf{r}}_{ij} &= (\cos \phi_{ij}, \sin \phi_{ij}), \\ \hat{\mathbf{t}}_{ij} &= (\sin \phi_{ij}, -\cos \phi_{ij}). \end{aligned} \quad (9)$$

Let's further define an error vector, which is defined as the difference between the measured value of the hole location and its predicted value:

$$\mathbf{e}_{ij} = \mathbf{d}_i - \mathbf{m}_i = \begin{pmatrix} x_i - x_{0j} - r \cos \phi_{ij} \\ y_i - y_{0j} - r \sin \phi_{ij} \end{pmatrix}. \quad (10)$$

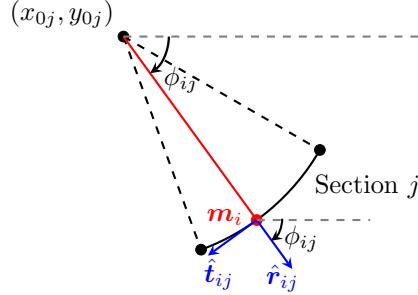


Figure 5: Diagram showing the position of the tangential unit vector \hat{t}_{ij} and \hat{r}_{ij} .

Then the Gaussian likelihood for the i th hole at j th section is given by

$$\begin{aligned}\mathcal{L}(\mathbf{d}_i|\boldsymbol{\theta}_i) &= \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}\begin{pmatrix} \mathbf{e}_{ij} \cdot \hat{r}_{ij} \\ \mathbf{e}_{ij} \cdot \hat{t}_{ij} \end{pmatrix}^T \Sigma^{-1} \begin{pmatrix} \mathbf{e}_{ij} \cdot \hat{r}_{ij} \\ \mathbf{e}_{ij} \cdot \hat{t}_{ij} \end{pmatrix}\right) \\ &= \frac{1}{2\pi\sigma_r\sigma_t} \exp\left(-\frac{1}{2}\left[\frac{(\hat{e}_{ij} \cdot \hat{r}_{ij})^2}{\sigma_r^2} + \frac{(\hat{e}_{ij} \cdot \hat{t}_{ij})^2}{\sigma_t^2}\right]\right).\end{aligned}\quad (11)$$

The overall likelihood function is therefore

$$\begin{aligned}\mathcal{L}(D|\boldsymbol{\theta}) &= \frac{1}{(2\pi\sigma_r\sigma_t)^n} \prod_{i=1}^n \exp\left(-\frac{1}{2}\left[\frac{(\hat{e}_{ij} \cdot \hat{r}_{ij})^2}{\sigma_r^2} + \frac{(\hat{e}_{ij} \cdot \hat{t}_{ij})^2}{\sigma_t^2}\right]\right) \\ &= \frac{1}{(2\pi\sigma_r\sigma_t)^n} \prod_{i=1}^n \exp\left(-\frac{1}{2}\left[\frac{(x_i - x_{0j}) \cos \phi_{ij} + (y_i - y_{0j}) \sin \phi_{ij} - r)^2}{\sigma_r^2} + \frac{((x_i - x_{0j}) \sin \phi_{ij} - (y_i - y_{0j}) \cos \phi_{ij})^2}{\sigma_t^2}\right]\right),\end{aligned}\quad (12)$$

where $n = 79$.

4 Computing the Gradient of Log-likelihood

This subsection answers the part (d) of the coursework.

Since the scenario and the models are relatively simple, we can compute the gradients of the log-likelihood with respect to each parameter analytically.

	mean	sd	hdi_3%	hdi_97%	mcse_mean	mcse_sd	ess_bulk	ess_tail	r_hat
N	355.172	2.566	350.336	360.065	0.039	0.026	4303.0	6145.0	1.0
alpha[0]	-145.709	0.156	-145.993	-145.404	0.002	0.001	10126.0	11843.0	1.0
alpha[1]	-145.659	0.515	-146.616	-144.684	0.005	0.004	11000.0	12096.0	1.0
alpha[2]	-145.556	0.371	-146.258	-144.857	0.006	0.004	4453.0	6387.0	1.0
alpha[3]	-145.504	1.343	-148.010	-142.943	0.013	0.010	10380.0	11045.0	1.0
alpha[4]	-145.089	1.437	-147.663	-142.296	0.014	0.011	10744.0	11987.0	1.0
alpha[5]	-146.968	0.946	-148.740	-145.197	0.010	0.007	8521.0	11668.0	1.0
r	77.320	0.553	76.282	78.372	0.008	0.006	4296.0	6065.0	1.0
sigma	0.095	0.006	0.084	0.106	0.000	0.000	20969.0	13042.0	1.0
x0[0]	79.672	0.410	78.880	80.421	0.006	0.004	4783.0	6881.0	1.0
x0[1]	79.874	0.643	78.667	81.076	0.006	0.005	10436.0	12044.0	1.0
x0[2]	79.864	0.095	79.682	80.039	0.001	0.001	10724.0	12439.0	1.0
x0[3]	79.867	1.674	76.714	83.008	0.016	0.013	11408.0	11198.0	1.0
x0[4]	79.931	1.779	76.662	83.330	0.017	0.013	11346.0	12443.0	1.0
x0[5]	82.154	1.016	80.265	84.038	0.009	0.007	12688.0	12704.0	1.0
y0[0]	136.044	0.417	135.261	136.840	0.006	0.004	4558.0	6576.0	1.0
y0[1]	135.710	0.567	134.616	136.764	0.008	0.005	5315.0	7900.0	1.0
y0[2]	135.682	0.561	134.616	136.733	0.009	0.006	4299.0	6150.0	1.0
y0[3]	135.618	0.771	134.179	137.055	0.010	0.006	5694.0	8847.0	1.0
y0[4]	135.260	0.875	133.576	136.857	0.011	0.007	6223.0	9127.0	1.0
y0[5]	135.979	0.679	134.667	137.210	0.009	0.006	5464.0	8587.0	1.0

Table 1: Summary statistics from MCMC for Model 1.

4.1 Gradient for the First Model

The log-likelihood for the first model is

$$\log \mathcal{L}(D|\boldsymbol{\theta}) = -2n \log \sigma - \sum_{i=1}^n \frac{(x_i - x_{0j} - r \cos \phi_{ij})^2 + (y_i - y_{0j} - r \sin \phi_{ij})^2}{2\sigma^2} + \text{constant} \quad (13)$$

where $n = 79$. Then the gradients with respect to the 21 parameters are:

$$\frac{\partial}{\partial r} \log \mathcal{L}(D|\boldsymbol{\theta}) = \sum_{i=1}^n \frac{(x_i - x_{0j}) \cos \phi_{ij} + (y_i - y_{0j}) \sin \phi_{ij} - r}{\sigma^2}, \quad (14)$$

$$\frac{\partial}{\partial N} \log \mathcal{L}(D|\boldsymbol{\theta}) = \sum_{i=1}^n r \left(\frac{2\pi(i-1)}{N^2} \right) \left(\frac{(x_i - x_{0j}) \sin \phi_{ij} - (y_i - y_{0j}) \cos \phi_{ij}}{\sigma^2} \right), \quad (15)$$

$$\frac{\partial}{\partial \sigma} \log \mathcal{L}(D|\boldsymbol{\theta}) = -\frac{2n}{\sigma} + \sum_{i=1}^n \frac{(x_i - x_{0j} - r \cos \phi_{ij})^2 + (y_i - y_{0j} - r \sin \phi_{ij})^2}{\sigma^3}. \quad (16)$$

For hole i in Section j , where $j = 1, 2, 3, 5, 6, 7$, we have the following gradients

$$\frac{\partial}{\partial x_{0j}} \log \mathcal{L}(D|\boldsymbol{\theta}) = \sum_{i \text{ in } j} \frac{x_i - x_{0j} - r \cos \phi_{ij}}{\sigma^2}, \quad (17)$$

	mean	sd	hdi_3%	hdi_97%	mcse_mean	mcse_sd	ess_bulk	ess_tail	r_hat
N	355.419	1.287	353.011	357.831	0.021	0.012	3724.0	6974.0	1.0
alpha[0]	-145.721	0.062	-145.828	-145.595	0.001	0.000	6512.0	9724.0	1.0
alpha[1]	-145.658	0.185	-145.998	-145.311	0.002	0.001	8984.0	11755.0	1.0
alpha[2]	-145.514	0.185	-145.844	-145.155	0.003	0.002	3807.0	7436.0	1.0
alpha[3]	-146.270	0.772	-147.735	-144.833	0.007	0.006	11285.0	11955.0	1.0
alpha[4]	-145.465	1.156	-147.707	-143.346	0.011	0.009	11226.0	11406.0	1.0
alpha[5]	-147.655	0.406	-148.400	-146.872	0.005	0.003	6706.0	9847.0	1.0
r	77.362	0.261	76.858	77.839	0.004	0.003	3459.0	6279.0	1.0
sigma_r	0.028	0.003	0.024	0.033	0.000	0.000	17797.0	13170.0	1.0
sigma_t	0.130	0.011	0.110	0.151	0.000	0.000	19521.0	13062.0	1.0
x0[0]	79.707	0.187	79.367	80.068	0.003	0.002	3654.0	6461.0	1.0
x0[1]	79.918	0.222	79.505	80.340	0.003	0.002	7820.0	11135.0	1.0
x0[2]	79.864	0.034	79.799	79.926	0.000	0.000	6081.0	9953.0	1.0
x0[3]	80.915	0.954	79.027	82.645	0.008	0.007	13796.0	12462.0	1.0
x0[4]	80.461	1.452	77.787	83.238	0.013	0.011	11862.0	11837.0	1.0
x0[5]	83.073	0.379	82.355	83.787	0.003	0.003	15054.0	13336.0	1.0
y0[0]	136.049	0.193	135.692	136.415	0.003	0.002	3517.0	6572.0	1.0
y0[1]	135.738	0.249	135.278	136.213	0.004	0.002	4023.0	7241.0	1.0
y0[2]	135.732	0.265	135.246	136.239	0.005	0.003	3468.0	6354.0	1.0
y0[3]	135.977	0.373	135.287	136.684	0.005	0.003	5653.0	9292.0	1.0
y0[4]	135.498	0.570	134.440	136.572	0.007	0.004	7673.0	11123.0	1.0
y0[5]	136.391	0.284	135.891	136.953	0.004	0.002	4396.0	8000.0	1.0

Table 2: Summary statistics from MCMC for Model 2.

$$\frac{\partial}{\partial y_{0j}} \log \mathcal{L}(D|\boldsymbol{\theta}) = \sum_{i \text{ in } j} \frac{y_i - y_{0j} - r \sin \phi_{ij}}{\sigma^2}, \quad (18)$$

$$\frac{\partial}{\partial \alpha_j} \log \mathcal{L}(D|\boldsymbol{\theta}) = -r \sum_{i \text{ in } j} \frac{(x_i - x_{0j}) \sin \phi_{ij} - (y_i - y_{0j}) \cos \phi_{ij}}{\sigma^2}. \quad (19)$$

4.2 Gradient for the Second Model

The log-likelihood for the second model is

$$\begin{aligned} \log \mathcal{L}(D|\boldsymbol{\theta}) = -n \log \sigma_r - n \log \sigma_t - \sum_{i=1}^n & \left[\frac{((x_i - x_{0j}) \cos \phi_{ij} + (y_i - y_{0j}) \sin \phi_{ij} - r)^2}{2\sigma_r^2} + \right. \\ & \left. \frac{((x_i - x_{0j}) \sin \phi_{ij} - (y_i - y_{0j}) \cos \phi_{ij})^2}{2\sigma_t^2} \right] \end{aligned} \quad (20)$$

where $n = 79$. Then the gradients with respect to the 22 parameters are:

$$\frac{\partial}{\partial r} \log \mathcal{L}(D|\boldsymbol{\theta}) = \sum_{i=1}^n \frac{(x_i - x_{0j}) \cos \phi_{ij} + (y_i - y_{0j}) \sin \phi_{ij} - r}{\sigma_r^2} \quad (21)$$

The rest of the gradients for remaining parameters follow the same steps as for the previous model.

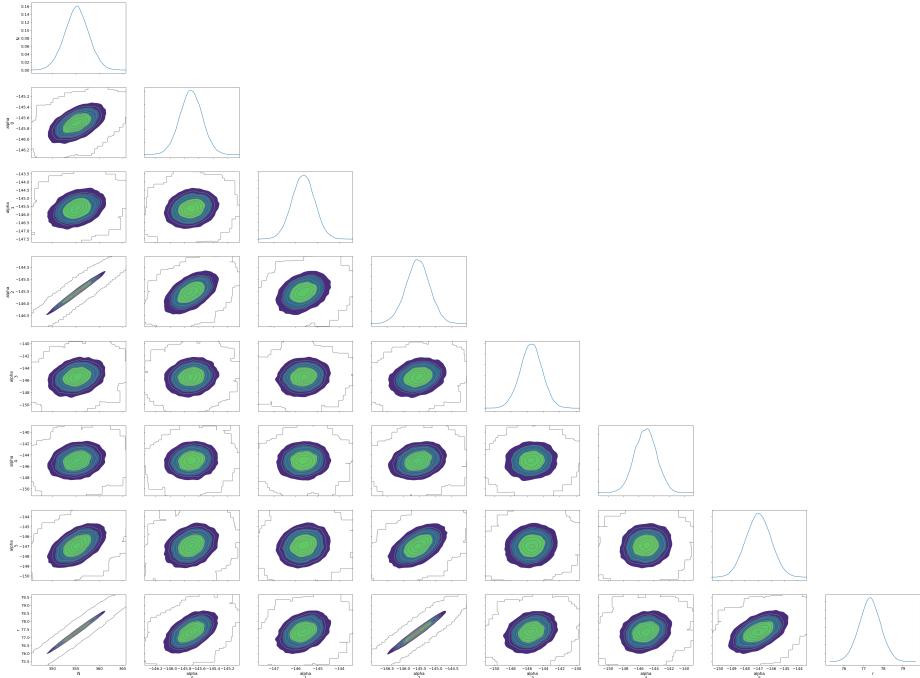


Figure 6: Corner plot which shows the joint distributions and marginal distributions of multiple parameters for Model 1.

4.3 Automatic Differentiation with JAX

However, instead of manually deriving and coding the gradient expressions, we can also use the automatic differentiation in JAX. By defining the log-likelihood function as a JAX computation and enabling gradient tracking for model parameters, JAX automatically computes the gradients using reverse-mode differentiation. This ensures both correctness and efficiency.

JAX enables automatic differentiation through its `grad` and `jacfwd/jacrev` functions, which compute gradients and Jacobians efficiently and accurately by leveraging reverse-mode and forward-mode differentiation.

Table (1) is the output from MCMC analysis with

`num_warmup=5000, num_samples=10000, num_chains=2.`

Table (2) is the output from MCMC analysis with the same input parameters.

Figure (6) and (7) are the corner plots which show the joint distributions and

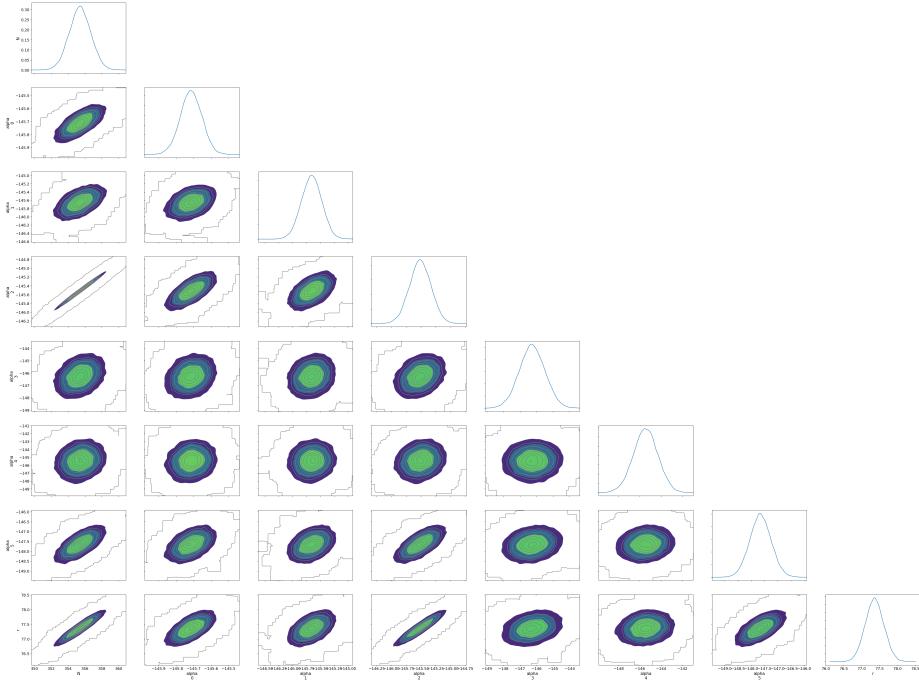


Figure 7: Corner plot which shows the joint distributions and marginal distributions of multiple parameters for Model 2.

marginal distributions of multiple parameters for Model 1 and 2, respectively. The diagonal plots are marginal distributions of each parameters. The marginal distributions are symmetric and smooth, which suggest that both models are well-constrained. For off-diagonal contour plots, circular shapes suggest no correlations between parameters. However, the contour plots between N and r and between N and $\text{alpha}[2]$ are highly skewed, which suggest a strong correlation between parameters.

5 Maximum Likelihood Estimators

This section answers part (e) of the coursework.

To find the MLE estimates for the parameters, I used the `minuit` package. Table (3) shows the MLE results for both models. It is clear that these estimates are relatively close to the estimates from MCMC.

Figure (8) and Figure (9) demonstrate the relationship between measured hole locations and the MLE predicted hole locations from Model 1 and Model

	Parameter	Value		Parameter	Value
0	N	355.254	0	N	355.237
1	r	77.343	1	r	77.336
2	sigma	0.088	2	sigma_r	0.026
3	x0[0]	79.686	3	sigma_t	0.123
4	x0[1]	79.904	4	x0[0]	79.689
5	x0[2]	79.868	5	x0[1]	79.907
6	x0[3]	81.505	6	x0[2]	79.860
7	x0[4]	81.508	7	x0[3]	81.436
8	x0[5]	83.213	8	x0[4]	81.519
9	y0[0]	136.063	9	x0[5]	83.220
10	y0[1]	135.723	10	y0[0]	136.031
11	y0[2]	135.705	11	y0[1]	135.715
12	y0[3]	136.127	12	y0[2]	135.706
13	y0[4]	135.846	13	y0[3]	136.101
14	y0[5]	136.422	14	y0[4]	135.843
15	alpha[0]	-145.702	15	y0[5]	136.420
16	alpha[1]	-145.667	16	alpha[0]	-145.721
17	alpha[2]	-145.543	17	alpha[1]	-145.673
18	alpha[3]	-146.755	18	alpha[2]	-145.539
19	alpha[4]	-146.317	19	alpha[3]	-146.704
20	alpha[5]	-147.797	20	alpha[4]	-146.327
21	alpha[5]	-147.805			

Table 3: Maximum likelihood estimators for the model parameters for Model 1 (left) and Model 2 (right).

2, respectively. We can see that both models give relatively accurate MLE predictions, which result in the overlapping between measured and MLE predicted hole locations.

6 Sample the Posterior Distribution using HMC

This section answers part (f) of the coursework.

The functions `sample_posterior_scatter` and `sample_posterior_contour` in `model1.ipynb` and `model2.ipynb` samples the posterior distribution 2000 times with MCMC algorithm. Figure (10) and (11) show the posterior distribution of Hole 31 with Model 1 (isotropic model) in terms of a scatter plot and contour plot, respectively, while Figure (12) and (13) show the same plots but for Model 2 (radial/tangential model).

It is as expected that Model 1 produces plots for hole distributions with rather circular shapes. This is because the model assumes that the x and y coordinates of each hole have the same standard deviation. It is also as expected that Model 2 produces plots for hole distributions with more elliptical shapes. This is because the model assumes the radial and tangential components of

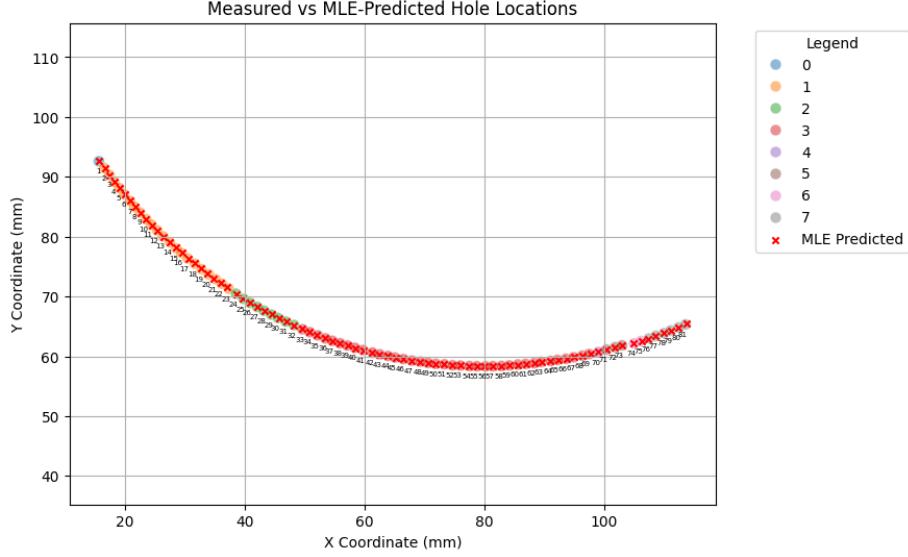


Figure 8: Measured hole locations overlapped with predicted hole locations for Model 1.

the coordinates have different standard deviations. Therefore, for example, in Figure (12) and (13), the shape of the distribution is roughly tangential to the original circle of the calendar ring.

7 Analysis and Comparison between Two Models

This section answers part (g) of the coursework.

To compare the two models, we first analyse the difference between two models. The isotropic model assumes that the x and y components share the same standard deviation (error), while the radial/tangential model assumes that x and y components have different standard deviations. Therefore, to find out the best model, I run another MCMC algorithm, with the 2-dimensional covariant matrix being

$$\Sigma = \begin{pmatrix} \sigma^2 & 0 \\ 0 & (\sigma + \delta)^2 \end{pmatrix},$$

parametrized by two parameters σ and δ . If δ is estimated to be close to 0, then it suggests that the isotropic model is better, and vice versa.

After running the function `find_best_model` in `model1.ipynb`, the result

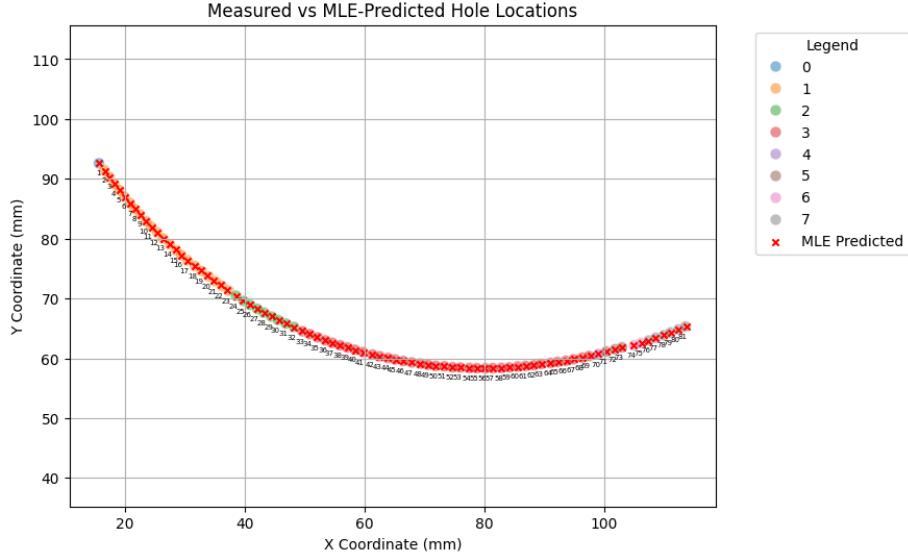


Figure 9: Measured hole locations overlapped with predicted hole locations for Model 2.

for δ is:

	mean	sd
delta	-0.133	0.054

Within one standard deviation, the range of δ does not trap 0. Therefore, we can conclude that the radial/tangential model is better.

References

- [1] Andrew Thoeni, Chris Budiselic, and Andrew Ramsey. *Replication Data for: Antikythera Mechanism Shows Evidence of Lunar Calendar*. Version V3. 2019. DOI: [10.7910/DVN/VJGLVS](https://doi.org/10.7910/DVN/VJGLVS). URL: <https://doi.org/10.7910/DVN/VJGLVS>.
- [2] Graham Woan and Joseph Bayley. *An improved calendar ring hole-count for the Antikythera mechanism*. 2024. arXiv: [2403.00040 \[physics.hist-ph\]](https://arxiv.org/abs/2403.00040). URL: <https://arxiv.org/abs/2403.00040>.

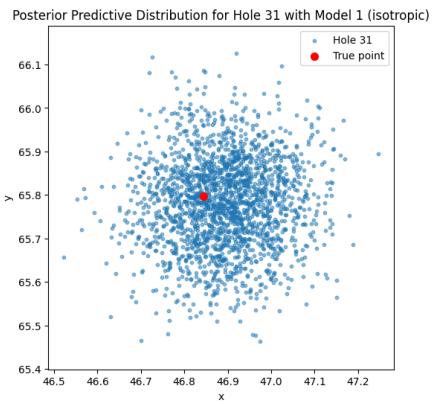


Figure 10: Scatter plot for locations of Hole 31 with 2000 samples using Model 1 (isotropic model).

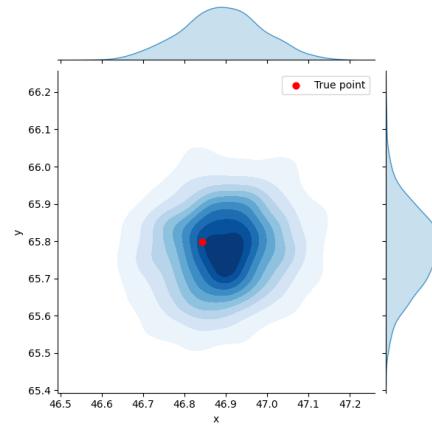


Figure 11: Contour plot with marginal distribution of x and y for Hole 31 using Model 1 (isotropic model).

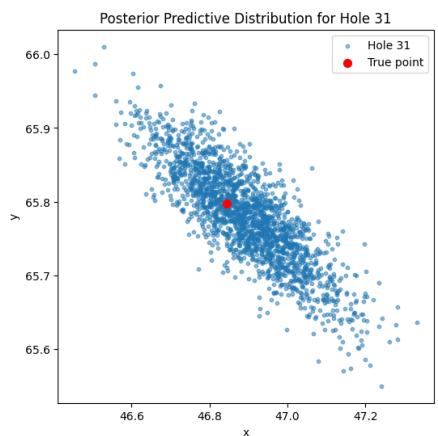


Figure 12: Scatter plot for locations of Hole 31 with 2000 samples using Model 2 (radial/tangential model).

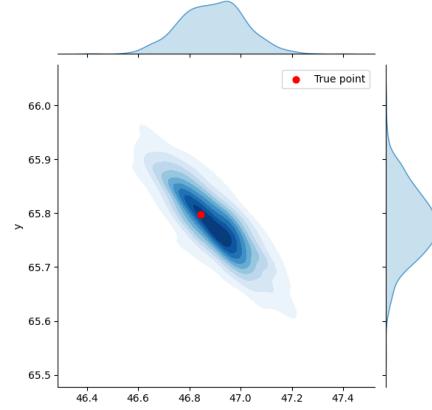


Figure 13: Contour plot with marginal distribution of x and y for Hole 31 using Model 2 (radial/tangential model).