

Data-Driven Methods for Engineers
(MECH0107)
- Coursework 1 -

2025 – 2026

Contents

1	Introduction and Problem Statement	2
2	Methodology	2
2.1	Motion Tracking (Image Processing)	3
2.2	Time–Frequency Analysis (Gabor Transform)	4
2.3	Dimensionality Reduction (SVD)	4
3	Results	5
3.1	Tracked Displacement Data	5
3.2	Spectrograms and Frequency Evolution	6
3.3	SVD Mode Decomposition	7
4	Discussion	9
4.1	Physical Interpretation of Modes	9
4.2	Mode Coupling and Nonlinearity	9
4.3	Simplified Analytical Model	9
4.4	Sources of Error and Improvements	9
5	Conclusion	9

1 Introduction and Problem Statement

The elastic pendulum, also known as the spring pendulum, is a classical example of a nonlinear hybrid dynamical system. It consists of a mass suspended from a spring, free to oscillate both vertically (along the spring axis) and horizontally (in the plane of swing). Unlike a simple pendulum or a pure spring-mass oscillator, the elastic pendulum couples two fundamentally different restoring mechanisms: the elastic spring force, governed by Hooke's law ($F = -kx$), and the gravitational restoring force that drives pendulum-like swinging. This coupling gives rise to rich dynamics. At certain energy levels the system exhibits near-periodic energy exchange between the two modes, and at others it can behave chaotically. The interaction becomes particularly pronounced when the natural frequency of the spring mode is approximately twice that of the pendulum mode, a condition known as autoparametric resonance. Studying this system is therefore not just an exercise in mechanics, but a useful case study for understanding how nonlinear coupling manifests in multi-degree-of-freedom oscillators.

In this coursework, a physical spring-pendulum system was recorded using three cameras placed at different orientations around the mass. The mass was released from an off-centre position, simultaneously exciting both the vertical spring oscillation and the horizontal pendulum swing. The video data from each camera were provided as MATLAB .mat files (`cam1.mat`, `cam2.mat`, `cam3.mat`), each containing a sequence of RGB video frames. Since no frame rate metadata was available, all frequency quantities are expressed in cycles per frame rather than in Hz throughout this report.

The analysis is structured around three main objectives. The first is to extract the position of the oscillating mass from each camera's video using image processing techniques, producing displacement time series for subsequent analysis. The second is to characterise how the frequency content of the motion evolves over time, which requires time-frequency analysis rather than a conventional Fourier transform, since the system's dynamics are non-stationary. The third is to reduce the dimensionality of the multi-camera dataset and identify the dominant modes of motion, linking each mode to its physical origin in the spring-pendulum dynamics.

To achieve these objectives, the following workflow was adopted. In Part 1, image processing techniques, including phase correlation for video stabilisation, brightness-based segmentation, binary thresholding, and connected component analysis, were applied to each camera's video to extract the centroid trajectory of the mass. In Part 2, the Gabor transform (short-time Fourier transform) was applied to the displacement signals to produce spectrograms, revealing how the dominant frequencies change over time. Three different Gaussian window widths were compared to illustrate the trade-off between time and frequency resolution. In Part 3, singular value decomposition (SVD) was applied to the combined multi-camera displacement matrix to extract the principal modes of motion, assess how many modes are needed to represent the system, and interpret each mode in terms of the underlying physics.

2 Methodology

This section describes the methods used to extract, analyse, and interpret the motion data from the three camera recordings. Each method is grounded in the techniques covered in the module lectures, and is chosen based on its suitability for this particular dataset.

2.1 Motion Tracking (Image Processing)

The 2D position of the mass was extracted from each camera's video using the six-step pipeline illustrated in Figure 1.

Video stabilisation. Small frame-to-frame camera shifts were removed by phase correlation. Each frame was converted to grayscale ($I = 0.299R + 0.587G + 0.114B$) and its translational shift estimated from the peak of the inverse FFT of the normalised cross-power spectrum:

$$C = \frac{\hat{f}_1 \cdot \overline{\hat{f}_2}}{|\hat{f}_1 \cdot \hat{f}_2|} \quad (1)$$

where \hat{f} denotes the 2D FFT and \cdot the complex conjugate. Cumulative shifts were applied to align all frames to frame 0.

ROI, brightness, and thresholding. Each stabilised frame was cropped to a region of interest (ROI) around the mass to exclude background clutter. The brightness was defined as the per-pixel channel maximum:

$$B(x, y) = \max(R(x, y), G(x, y), B(x, y)) \quad (2)$$

A fixed threshold θ then produced a binary mask:

$$M(x, y) = \begin{cases} 1 & B(x, y) > \theta \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

Values ($\theta = 240, 245, 200$ for Cameras 1–3) were chosen from the intensity histogram of each camera.

Morphological cleanup and centroid. Connected component labelling isolated distinct blobs; those with area outside [20, 8000] px were discarded, and the blob nearest the previous centroid was retained. The mass position was taken as the centre-of-mass:

$$(c_x, c_y) = \left(\frac{\sum_{(x,y) \in \mathcal{B}} x}{|\mathcal{B}|}, \frac{\sum_{(x,y) \in \mathcal{B}} y}{|\mathcal{B}|} \right) \quad (4)$$

Frames with no valid detection were filled by linear interpolation, and the mean position was subtracted to produce centred displacement signals $\Delta x(t)$ and $\Delta y(t)$. Camera 3 required an additional $90^\circ + 24^\circ$ rotation before processing to correct its mounting angle.

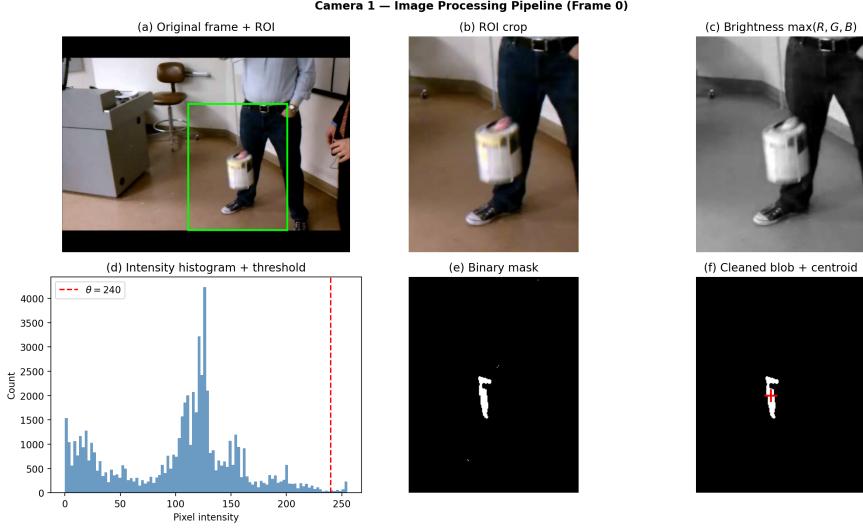


Figure 1: Image processing pipeline for Camera 1, Frame 0. (a) Original RGB frame with ROI (green box). (b) ROI crop. (c) Brightness channel $\max(R, G, B)$. (d) Intensity histogram with threshold $\theta = 240$ (dashed red line). (e) Binary mask. (f) Cleaned blob with detected centroid (red cross).

2.2 Time–Frequency Analysis (Gabor Transform)

The spring-pendulum is non-stationary: the pendulum swing damps out while the spring oscillation persists, so frequencies evolve with time. A global Fourier transform cannot capture this because it averages over the entire recording. Instead, the Gabor transform (short-time Fourier transform) was applied, which localises the spectrum in time by windowing the signal before transforming.

Window function. For each sliding time position τ , a Gaussian window is centred on the signal $S(t)$:

$$g_\tau(t) = \exp\left(-\frac{(t - \tau)^2}{a^2}\right) \quad (5)$$

The windowed signal $S(t) \cdot g_\tau(t)$ is then Fourier-transformed to give the local frequency content at time τ . Repeating across all τ builds the full spectrogram $\mathcal{G}(\tau, \omega)$.

Window width and resolution trade-off. The width a controls the balance between time and frequency resolution, as required by the Heisenberg uncertainty principle: a narrow window (a small) gives good temporal localisation but blurs frequency; a wide window gives sharp frequency peaks but poor time resolution. Three widths were compared — $a = 5, 15, 40$ frames — to illustrate this trade-off. Since no frame-rate metadata was available, all frequencies are reported in cycles per frame throughout.

2.3 Dimensionality Reduction (SVD)

The six displacement signals (x and y from three cameras) were stacked into a data matrix and decomposed using singular value decomposition (SVD) to identify the dominant modes of motion shared across cameras.

Data matrix and mean-centering. The matrix $\mathbf{X} \in \mathbb{R}^{6 \times n}$ was formed by stacking the displacement signals as rows (cam1 x, cam1 y, cam2 x, cam2 y, cam3 x, cam3 y). Row means were subtracted to give the centred matrix $\mathbf{B} = \mathbf{X} - \bar{\mathbf{X}}$.

SVD decomposition. SVD was applied to the scaled matrix:

$$\frac{\mathbf{B}}{\sqrt{n-1}} = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^\top \quad (6)$$

where the columns of \mathbf{U} are the spatial modes (left singular vectors), $\boldsymbol{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_6)$ contains the singular values, and \mathbf{V} holds the temporal coefficients. The scaling by $\sqrt{n-1}$ makes the singular values equivalent to the square roots of eigenvalues of the sample covariance matrix, so that:

$$\lambda_k = \frac{\sigma_k^2}{\sum_j \sigma_j^2} \quad (7)$$

gives the fraction of total variance explained by mode k . The temporal projection of each mode was computed as $\mathbf{Y} = \mathbf{U}^\top \mathbf{B}$, and low-rank reconstructions $\hat{\mathbf{X}}_r = \mathbf{U}_r \boldsymbol{\Sigma}_r \mathbf{V}_r^\top \sqrt{n-1} + \bar{\mathbf{X}}$ were used to assess how many modes are needed to represent the data.

3 Results

3.1 Tracked Displacement Data

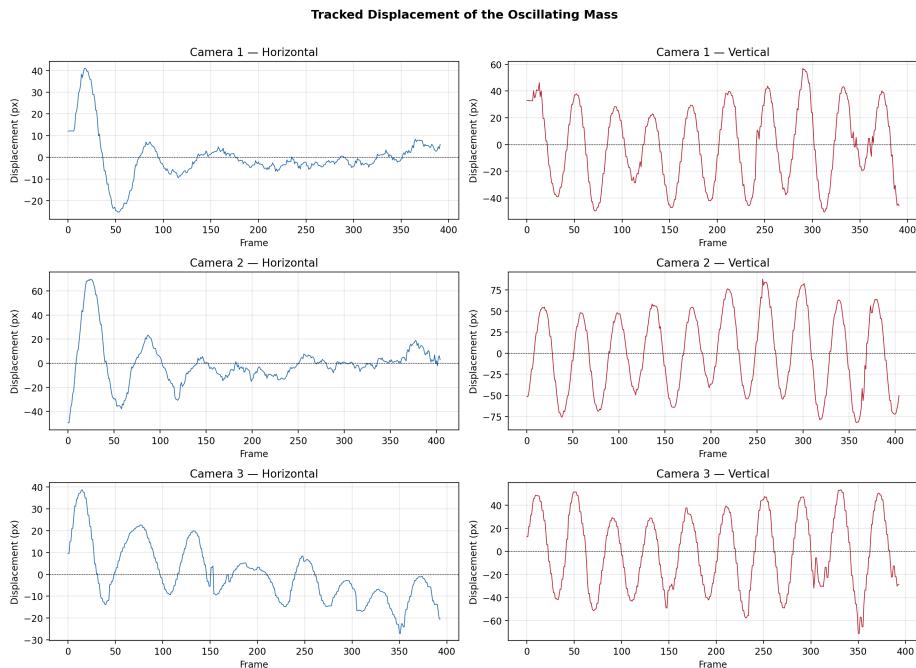


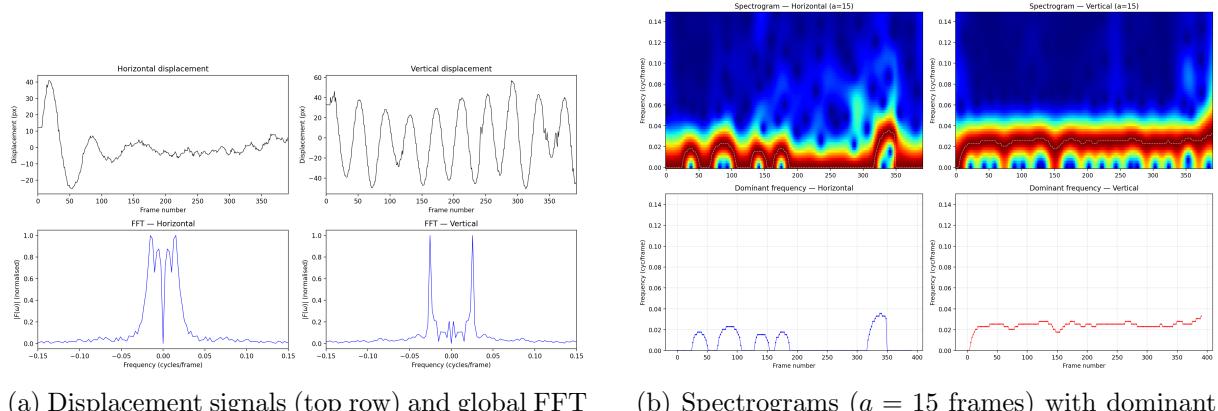
Figure 2: Mean-centred displacement time series from all three cameras (horizontal left, vertical right). The horizontal components show decaying amplitude (pendulum swing); the vertical components show a more persistent higher-frequency oscillation (spring mode).

Figure 2 shows the mean-centred displacement signals extracted from each camera. In Camera 1 (side view), the horizontal component shows a clear oscillation whose amplitude

decays steadily over the recording, consistent with a damped pendulum swing. The vertical component oscillates at a visibly higher frequency and remains largely sustained throughout, indicating the persistent spring-mode vibration. Cameras 2 and 3 capture different projections of the same underlying motion, and all three cameras show this characteristic two-frequency behaviour, with the lower-frequency component damping out faster than the higher-frequency one.

3.2 Spectrograms and Frequency Evolution

The global FFT of the Camera 1 signals (Figure 3, top) shows two distinct frequency peaks. The vertical displacement has a dominant peak at approximately twice the frequency of the horizontal peak, consistent with the 2:1 ratio expected at autoparametric resonance. The spectrograms (Figure 3, bottom) reveal how each component evolves in time: the horizontal frequency band is strong early on but decays and becomes noisy in the second half of the recording as the pendulum damps out, while the vertical frequency band remains coherent throughout. The white dashed line tracking the instantaneous dominant frequency confirms this pattern — a relatively stable peak for the spring mode, and an increasingly scattered dominant frequency for the pendulum mode as its amplitude falls below the noise floor.



(a) Displacement signals (top row) and global FFT magnitude spectra (bottom row) for Camera 1, showing two distinct frequency peaks.

(b) Spectrograms ($a = 15$ frames) with dominant frequency overlay (white dashed), and dominant frequency tracks over time.

Figure 3: Global FFT and time-frequency analysis for Camera 1. The vertical (spring) component is sustained; the horizontal (pendulum) component damps out.

3.3 SVD Mode Decomposition

The singular value spectrum (Figure 4, left) shows a steep drop after the first mode, with Modes 1, 2, and 3 accounting for 59.1%, 26.3%, and 9.3% of total variance respectively. As shown in Figure 5, three modes together capture 94.6% of the variance — just below the conventional 95% threshold, requiring a fourth mode (cumulative 98.0%) to exceed it. The fact that three modes are needed, rather than the two expected for a linear two-degree-of-freedom system, is indicative of nonlinear coupling.

The temporal projection of each mode (Figure 4, right) reveals distinct characters. Mode 1 oscillates at a nearly constant frequency throughout the recording with little decay, consistent with the sustained spring oscillation. Mode 3 oscillates at a lower frequency and its amplitude drops significantly in the first half of the recording, consistent with

the damped pendulum swing. Mode 2 lies between the two and may reflect a coupled interaction or a second plane of oscillation.

The spatial structure of each mode (Figure 5, right) confirms these physical interpretations. Mode 1 has its largest weights on the Camera 1 and Camera 2 vertical channels, consistent with spring-mode (vertical) dominance. Mode 3 is weighted more heavily on the horizontal channels, linking it to the pendulum swing. Mode 2 shows a mixed structure with contributions from multiple cameras and axes.

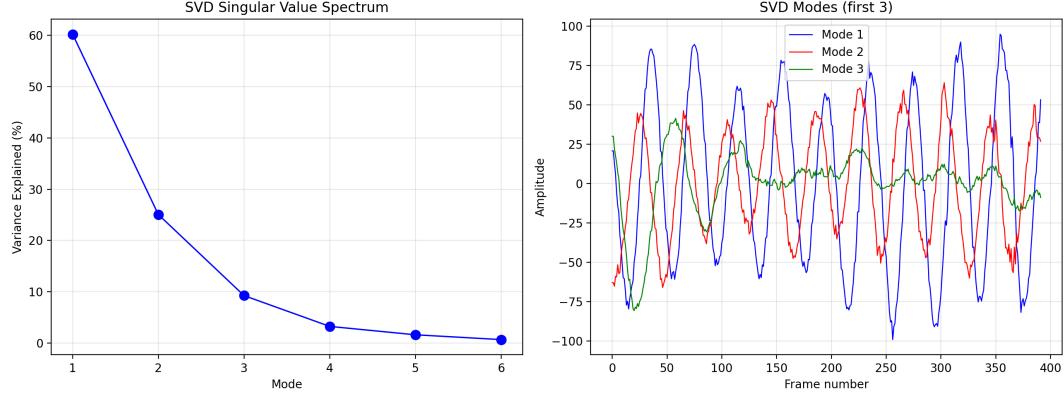
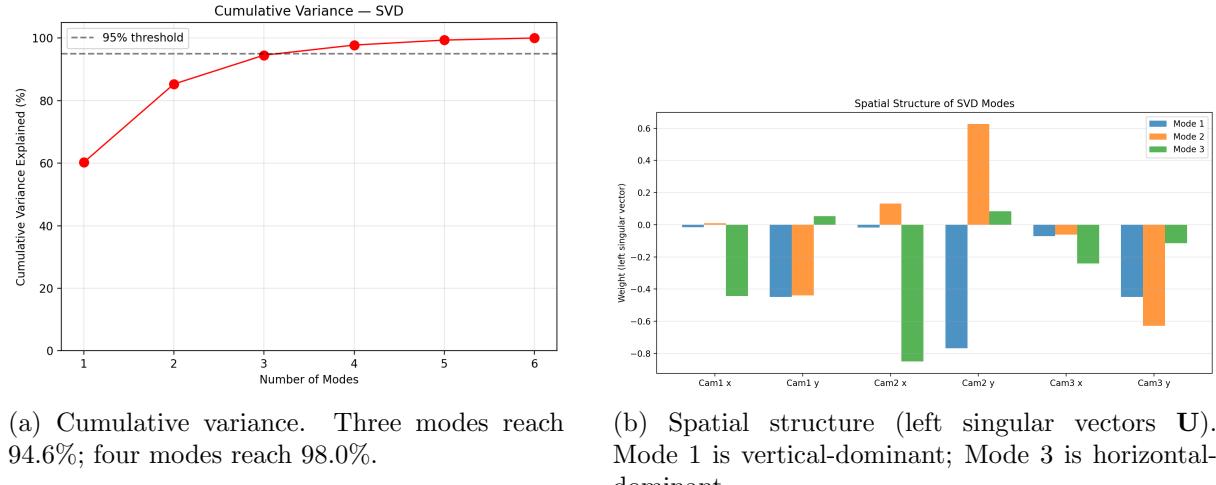


Figure 4: Left: variance explained by each SVD mode. Right: temporal projections of the first three modes. Mode 1 (blue) is sustained; Mode 3 (green) damps out.



(a) Cumulative variance. Three modes reach 94.6%; four modes reach 98.0%.

(b) Spatial structure (left singular vectors \mathbf{U}). Mode 1 is vertical-dominant; Mode 3 is horizontal-dominant.

Figure 5: Cumulative variance and spatial mode structure from the SVD decomposition.

The low-rank reconstruction (Figure 6) shows that a rank-1 approximation captures the dominant spring oscillation but misses the pendulum component entirely. Adding a second mode significantly improves the fit for the horizontal signal, and a rank-3 approximation recovers the main features of both components.

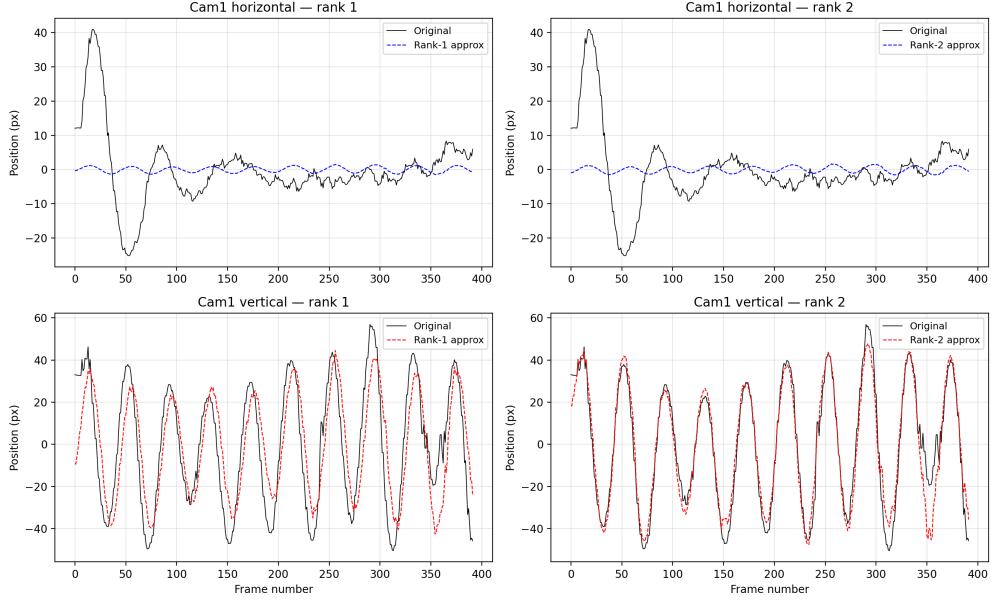


Figure 6: Low-rank reconstruction of Camera 1 horizontal (top) and vertical (bottom) signals for rank-1 (left) and rank-2 (right) approximations.

4 Discussion

4.1 Physical Interpretation of Modes

4.2 Mode Coupling and Nonlinearity

4.3 Simplified Analytical Model

4.4 Sources of Error and Improvements

5 Conclusion