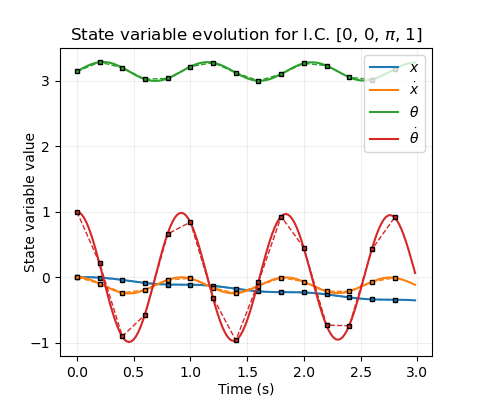
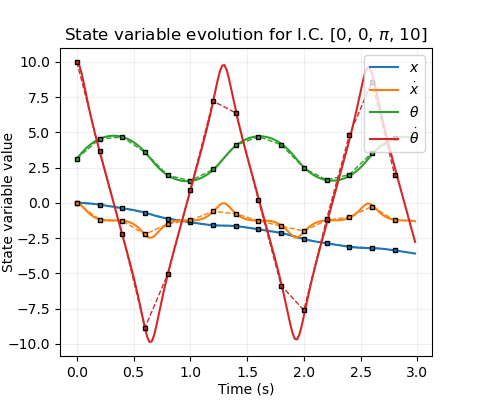
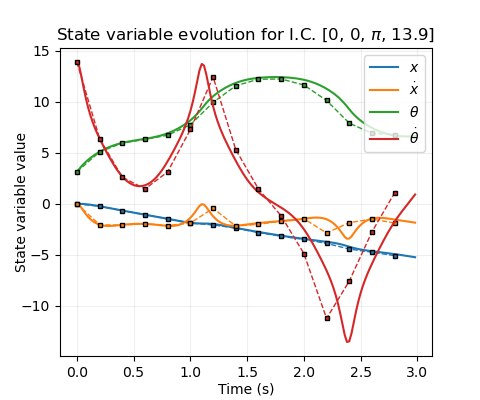
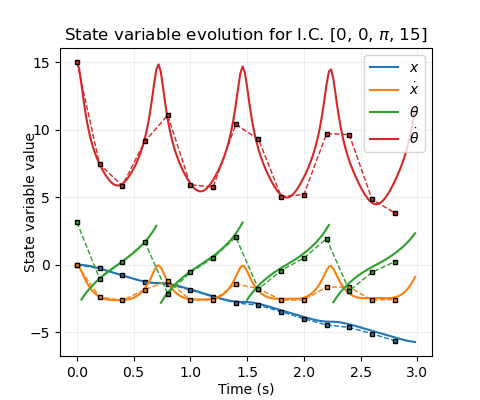
# Introduction

Keep very short. 1 paragraph sufficient.

# General considerations

The model runs 50 Euler steps over a time step of 0.2 seconds. The system has oscillating variables, and so the Nyquist criterion should be considered: if the system is observed at less than half the sampling rate of its maximum oscillation frequency, aliasing artefacts will occur. Such trajectories would be sharply nonlinear, and so would be difficult to predict with any model, linear or nonlinear. Since the sampling rate is fixed, it is therefore best to stick to low frequencies, and assume that useful controllers will keep the system state within this domain.

For the pole angular velocity:



**Fig 1.** Time evolution of system variables with no applied force, starting at the stable equilibrium (, with varying angular velocities. **Solid lines**: Illustrative evolution using a timestep of 0.02 s. **Dashed lines & markers**: Sampled values with the actual timestep (0.2 s). Note that the nature of the Euler scheme causes these pairs to diverge when friction is significant.

So should be satisfactory. Similarly, but more arbitrarily, will be taken. is an angle (centred at the unstable equilibrium) and so . The system dynamics are invariant to shifts in , so can be set initially to zero for all trajectories.

# Rollouts

## Time-trajectories

There are two principal trajectories of the system: *bound*and *unbound*. Bound trajectories occur when the pole oscillates with insufficient energy to reach . In the bound, low-energy limit (fig 1a), the pole oscillates sinusoidally about . The cart’s velocity couples to this, oscillating in phase with the but with a much smaller amplitude. However, it remains of an opposite sign to the initial angular velocity, so that the cart position “drifts” monotonically. Conversely, unbound trajectories have sufficiently high energy that the pole repeatedly “flips”, causing to monotonically increase (sans remapping) (fig 1d). Between these extremes, the behaviour stays broadly similar either side of the critical transition energy (, defining zero potential at [A1]), and the monotonic cart “drifting” remains, though the waveforms change shape in a nontrivial way (fig 1b). In all cases, friction gradually causes the energy to diminish and the oscillation amplitudes to reduce, so that bound states decay to zero and unbound states eventually become bound (fig 1c). This effect of friction can be largely decoupled from the broader system dynamics, especially when a controller is present to offset the energy loss.

Adding a nonzero cart velocity leaves the pole dynamics mostly unchanged (see next section), and shifts in initial cart position are trivial.

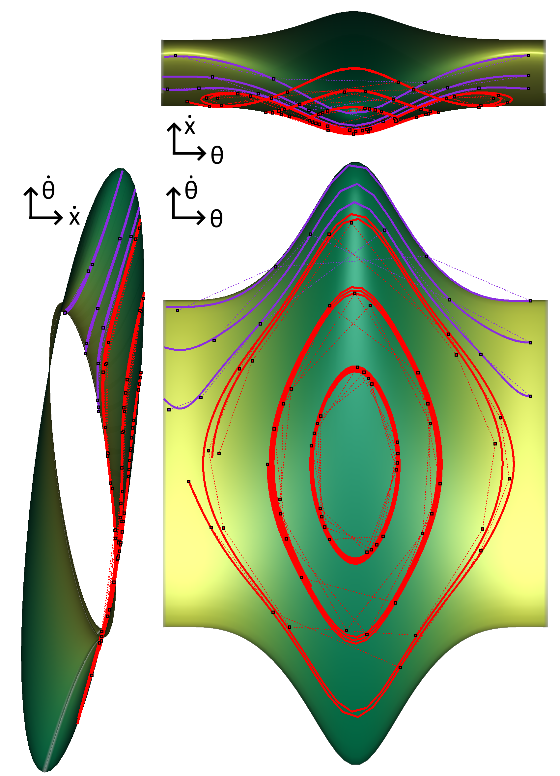
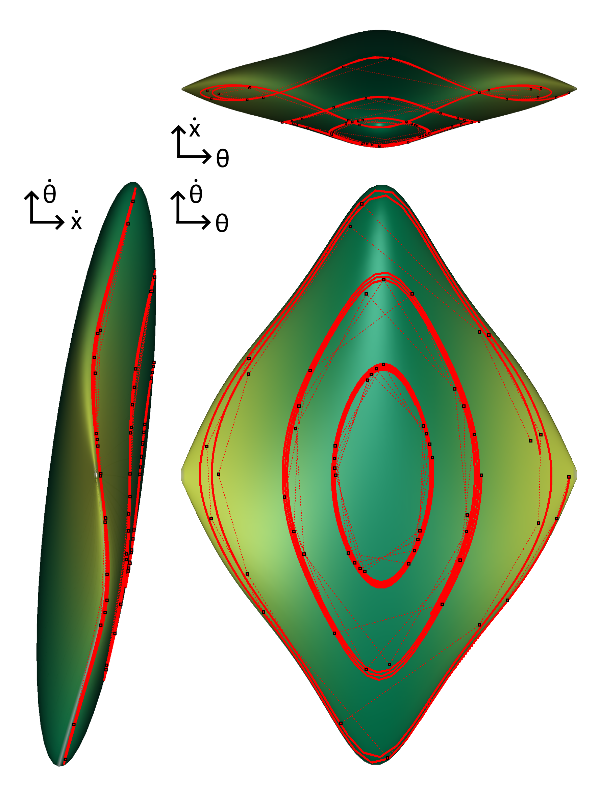
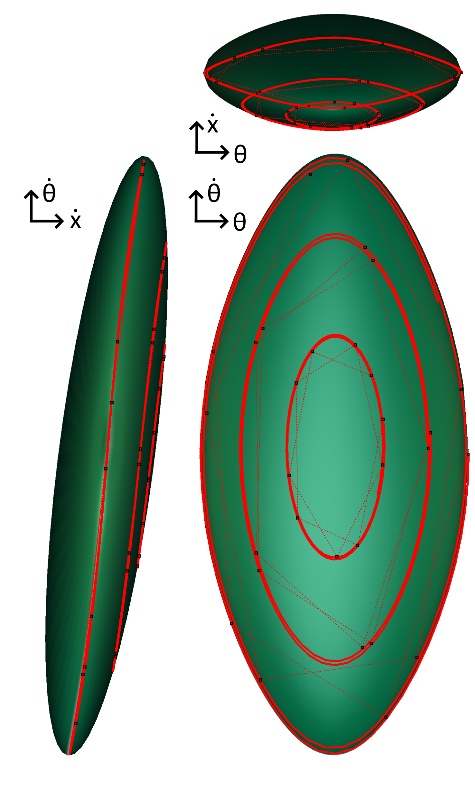
## Waveform evolution, (frequency?)

Come back to this later.

## Phase portraits

Since the cart position is trivial to deduce given its velocity history, and the timescale of energy loss due to friction is much greater than the timescale of the cart dynamics, the a qualitative understanding of the state-space trajectories of this system can be gained entirely via energy isosurfaces embedded in the 3d space formed by [. (Define the energy?)

**Fig 2.** Orthographic views of trajectories in [-space with initial energies equal to 1, 0, and -2 for a), b), c) respectively, superimposed onto energy isosurfaces at those values. **Red lines**: Bound trajectories. **Purple lines:** Unbound. **Solid lines:** Illustrative, with 0.02 s timestep. **Dotted lines & markers:** 0.2 s timestep.



**c)**

**a)**

**b)**

This consists of ellipses in the - plane modulated in size and tilt angle by the value of . When , a continuous, periodic tube is formed. It splits into disjoint surfaces at , which shrink into ellipsoids as reduces to its minimal value. Over longer time periods, trajectories migrate onto lower-energy isosurfaces until they become bound and oscillations decay.

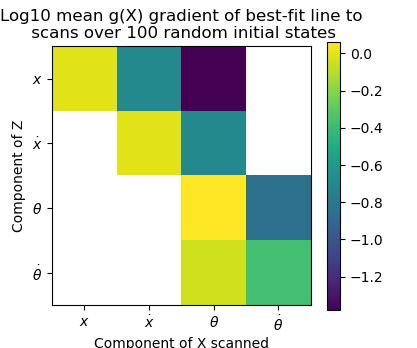
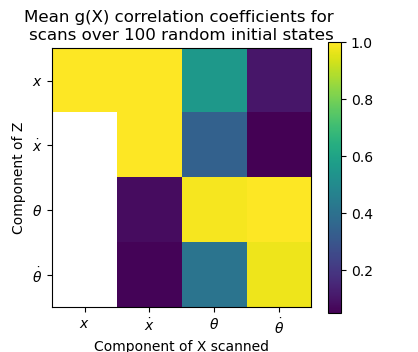
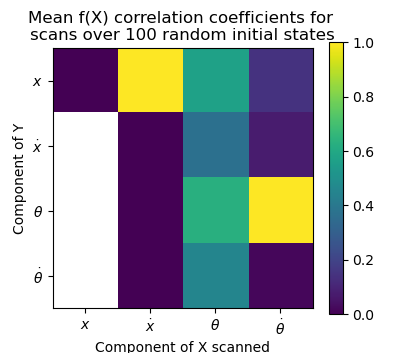
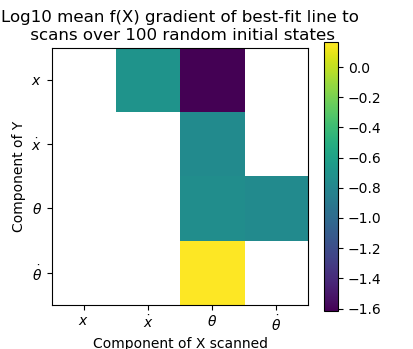
# Changes of State

## Formulation

Suppose is the current state of the system. After one step of time evolution (one “action”), let be the next state, and define . The goal of modelling this system is to learn the mapping , or equivalently .

## General Properties of

Over the relatively small timestep T=0.2 s, variables should generally evolve in a relatively simple way. In fact, many pairs of input-output variables have approximately linear relationships for . This is demonstrable by initialising the system in random states, scanning alternately over the full ranges of state variables, and calculating correlation coefficients with the output state variables. Figure 3a displays the mean of such values over 100 initial states – clearly a number of matrix elements imply linearity. In particular, since each variable does not change much, it is correlated strongly to itself, and the gradient of its best fit line is approximately equal to 1 (fig. 3b).



**Fig 3.** White is np.nan.

Subtracting the initial state gives and removes these diagonal correlations (fig. 3c) – though the others remain – as well removing constants so that correlated variables’ best fit lines have intercepts approximately set to zero. If we fit a linear model to , i.e. for matrix , we can hope to capture these particular relationships. It won’t

Find correlation coefficients for **f** matrix?

To simplify the task of characterising the 16-dimensional mapping of to , it is useful to build an idea of the “average” dependence of system evolution on a particular input parameter over starting states. One way of doing this is to sample random states, evaluate the components of at each sample, and compare by considering their relative summed magnitudes. This matrix is plotted as a heat map in Figure 4.

(To obtain this, the system as initialised in 200 random states, a particular state variable was changed by a uniformly drawn random value up 0.01% of its range (i.e 0.0015 for , and the magnitudes of the difference in evolutions over a timestep between the first and second state were summed and tabulated for each variable.)

Each state variable’s evolution depends fairly strongly on and . Variables only have a very weak dependence on , except of course . The variables are all completely independent of .

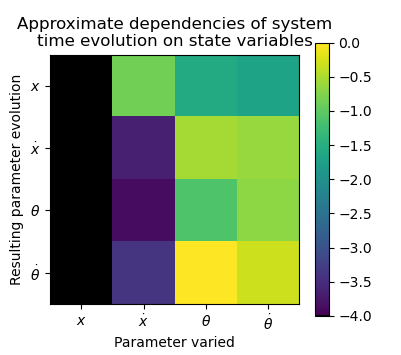
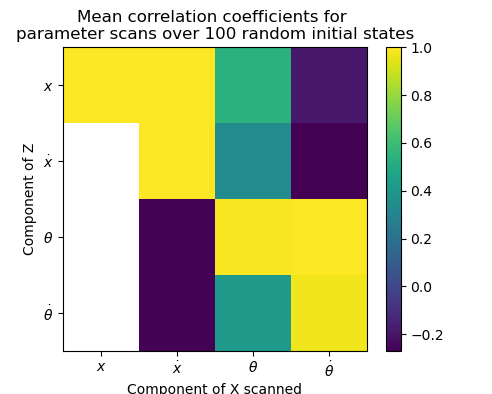
These traits can be understood by considering the dynamical equations of the system [A2], where causal connections for an infinitesimal timestep are depicted schematically in Figure 4. There is a general flow of information downward through the figure, with only affecting the angular variables and itself through the weak effects of friction. During the 50 Euler iterations per timestep, the information flow is diffused through the network, resulting in the causality matrix of Figure 3.

Dependence of each parameter on theta, theta dot.

Dependence on x dot – linear and weak.

Independence on x.

# Linear Model



**Fig 3.**White squares indicate that there is a linear relationship with gradient , so the correlation coefficient is undefined and the output is completely independent of the input

**Fig 4.** Black squares indicate a value numerically zero, i.e. the output is completely independent of the input

Appendix

Energy definition

Equations of motion

**Correlation matrix**

|  |  |  |  |
| --- | --- | --- | --- |
| 1 | 1 | 0.544 | 0.029 |
| n/a | 1 | 0.350 | -0.073 |
| n/a | -0.160 | 0.985 | 0.998 |
| n/a | -0.280 | 0.425 | 0.972 |

**Dependency matrix**

|  |  |  |  |
| --- | --- | --- | --- |
| 1e-15 | 0.14061 | 0.544 | 0.029 |
| 0 | 0.00015 | 0.25843 | -0.073 |
| 0 | 0.00003 | 0.10052 | 0.998 |
| 0 | 0.00030 | 1 | .972 |