

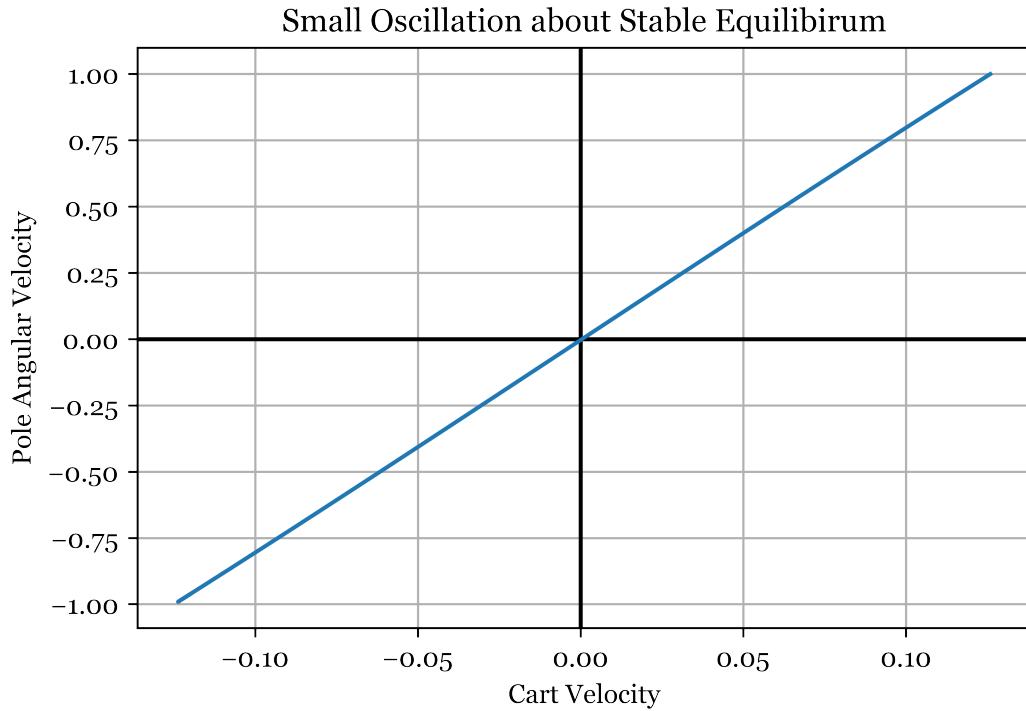
sf3

May 18, 2022

## 1 Task 1.1 - Simulation of Rollout

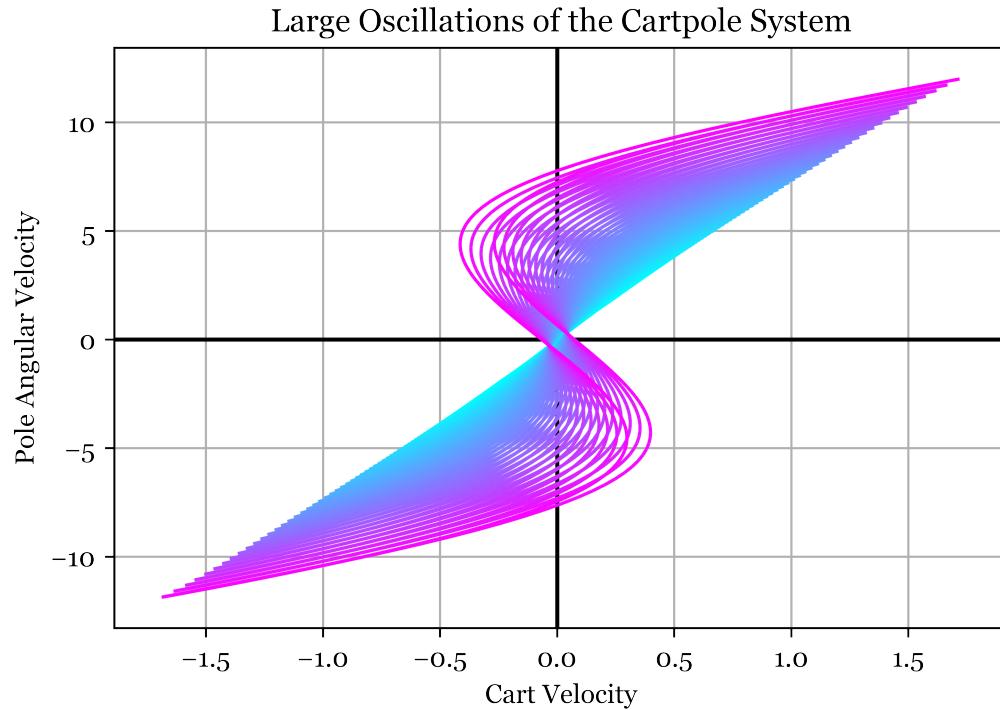
I have established a few different initial conditions and then plotted graphs to represent the resulting motions. I recorded the values of the state vector as the system evolved and then plotted some of the elements of the vector against each other at each timestep to form a line.

### 1.1 Small Oscillation about Stable Equilibrium



The plot above shows pole angular velocity plotted against cart velocity as time progresses through the rollout - the system oscillates up and down the blue line as the pole swings. The linear trend reflects the small angle assumption - the restoring force from the pendulum is a linear function of the deflection. From linear vibration theory we know the cart and pole vibrate at the same frequency so they “move together” and are therefore linearly correlated.

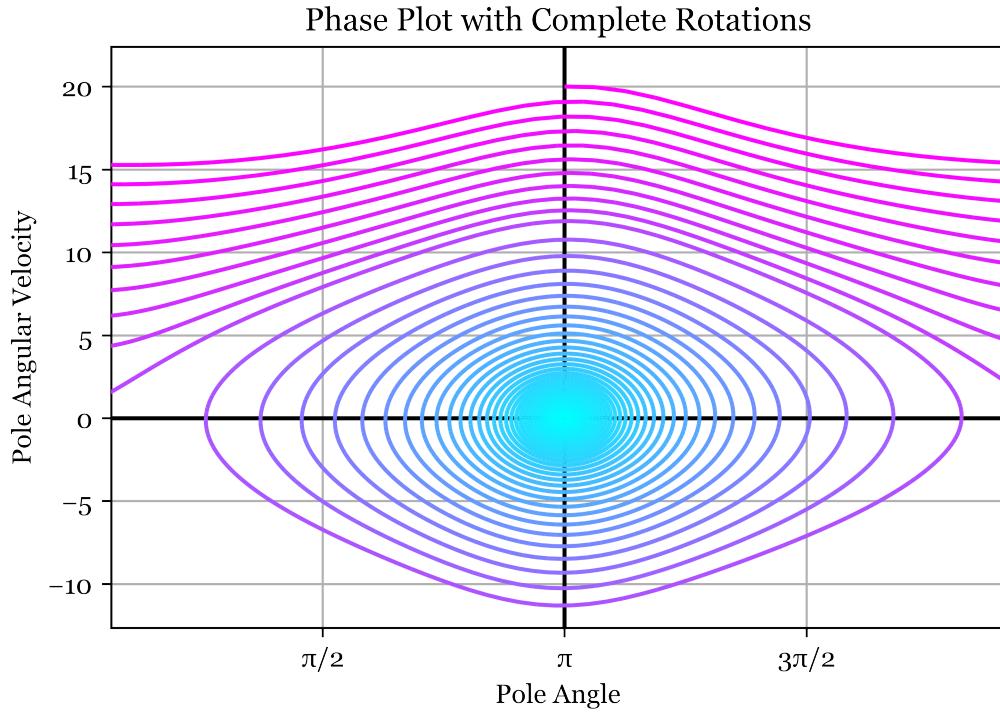
## 1.2 Large Amplitude Oscillations



The plot above shows large amplitude oscillations, with the line turning from magenta to cyan as time progresses. The small angle approximation no longer holds, with the curve of the lines showing the tendency of the cart to change direction during the pole's swing. The plot shows how the amplitude and curvature of the line decreases as energy is dissipated and we return to the small oscillation regime.

## 1.3 Swinging Over the Top

```
Text(0, 0.5, 'Pole Angular Velocity')
```



The plot above is the phase portrait of the pole angle and angular velocity, again starting from magenta and going to cyan. The upper region is where full rotations of the pole occur; then when enough energy is dissipated the system falls into the lower spiral towards the stable equilibrium at a pole angle of  $\pi$ .

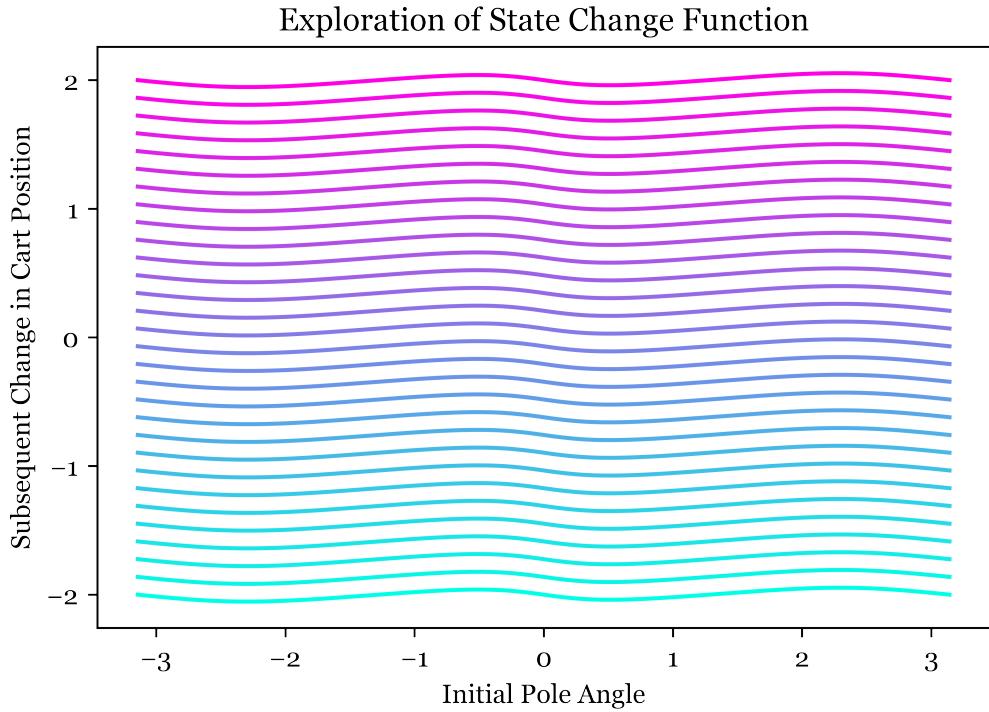
## 2 Task 1.2: Changes in State

The change in state vector over the next time step  $Y \equiv X(T) - X(0)$  is related to the current state vector  $X(0) = [x, \dot{x}, \theta, \dot{\theta}]$  by some complex function which I want to model. I made some plots showing how the change in state vector depends on the current state vector, noting that the change in state vector is not a function of the cart position due to the translational symmetry/invariance of the system.

It is more helpful to investigate the change in state as a function of the initial state rather than the subsequent state as a function of the current state. This is because the subsequent state will be quite similar (almost linearly related) to the initial state, assuming a small timestep. We can capture more meaning in our model by investigating the change in state, which subtracts off the information we already know (the initial state).

### 2.1 Effect of Varying Initial Cart Velocity and Pole Angle

```
Text(0, 0.5, 'Subsequent Change in Cart Position')
```

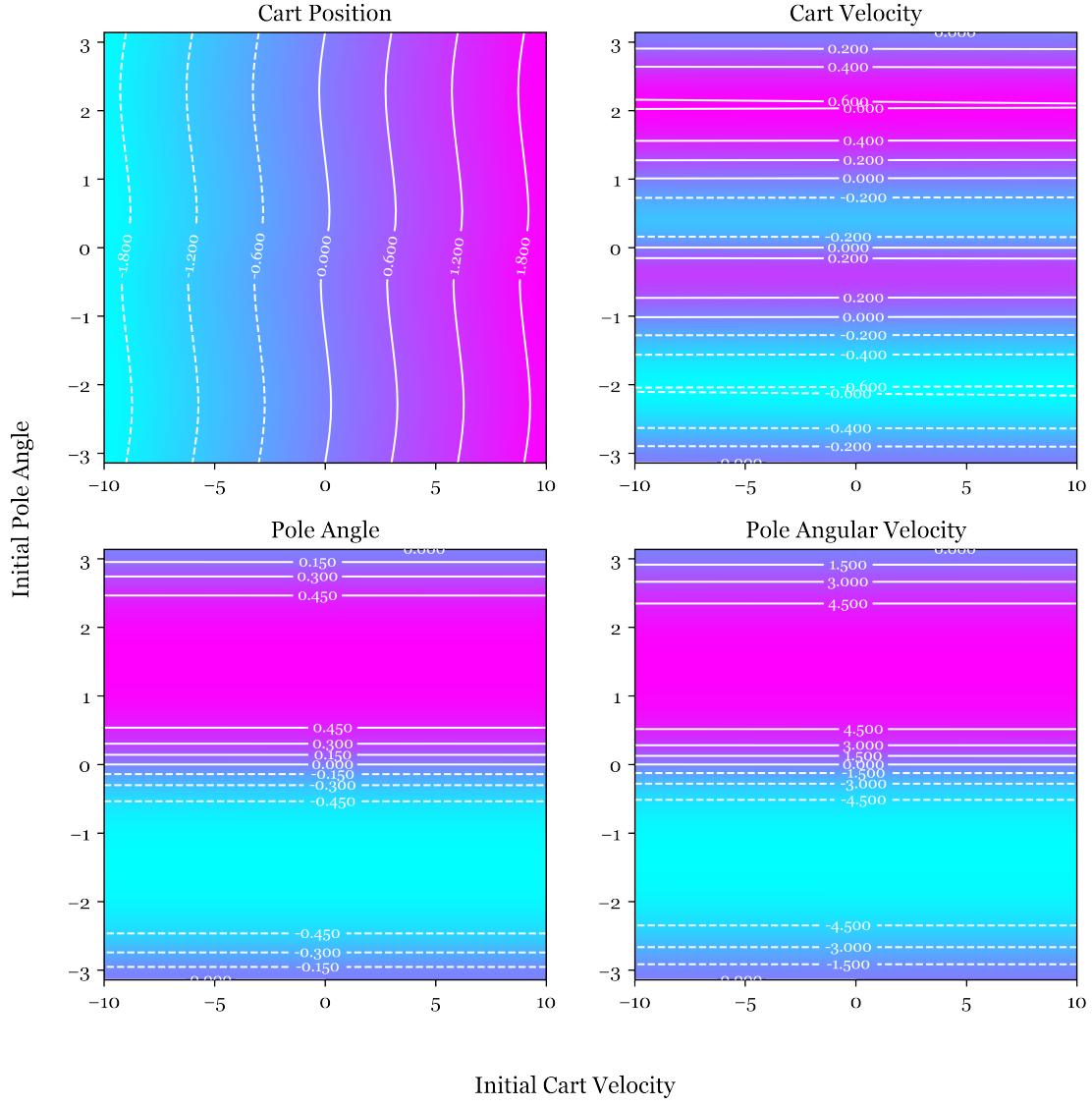


In the plot above, I am varying initial pole angle along the lines in the x direction, and varying initial cart velocity between lines, from -10 in cyan to +10 in magenta. The y axis is the change in cart position. We can see that the lines are equally spaced, so the change in cart position seems to be a linear function of cart velocity which makes sense. However, the change in cart position is a (weakly) nonlinear function of the pole angle since there is a slight wobble in the lines.

Plotting the changes in other state variables in this way yields coincident lines which are better represented in a contour plot:

```
Text(0.06, 0.5, 'Initial Pole Angle')
```

### Changes in State as a Function of Initial State

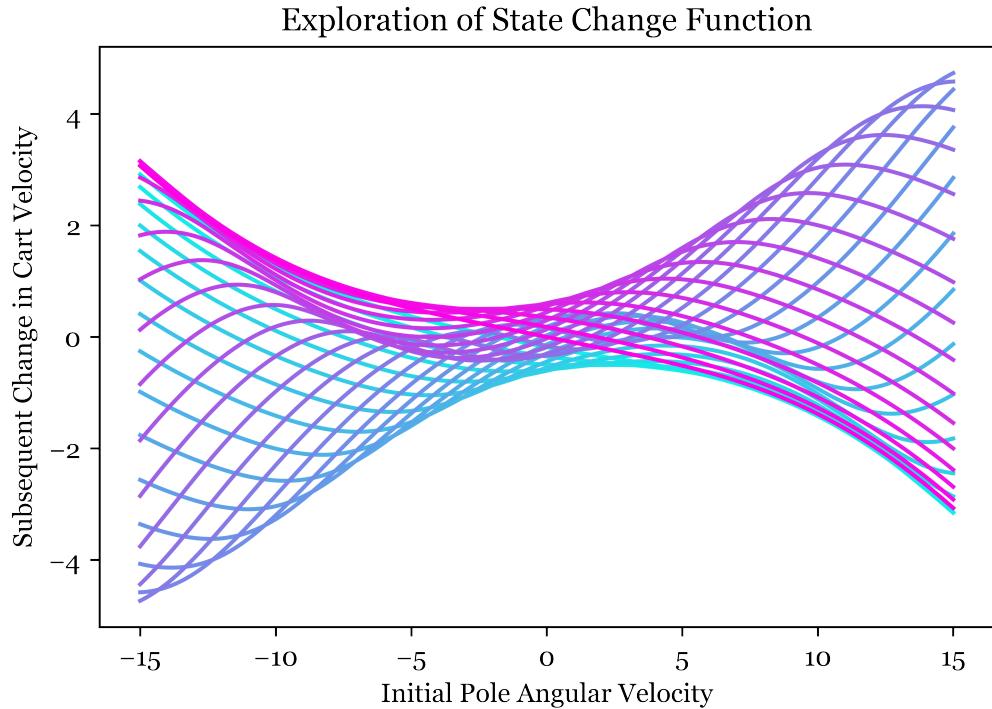


In the plot above, the initial cart velocity varies in the x direction and the intial pole angle in the y direction. The subsequent changes in cart position, cart velocity and pole angle are shown as contour plots. The changes in cart velocity, pole angle and pole angular velocity are not functions of the initial cart velocity. This must be the case because any initial cart velocity can be achieved by assuming a different intertial reference frame, leaving the rest of the dynamics unchanged.

The same reasoning leads to the contour plots for varying cart velocity and pole angular velocity having straight lines so I will move to the more interesting combination of varying pole angle and angular velocity.

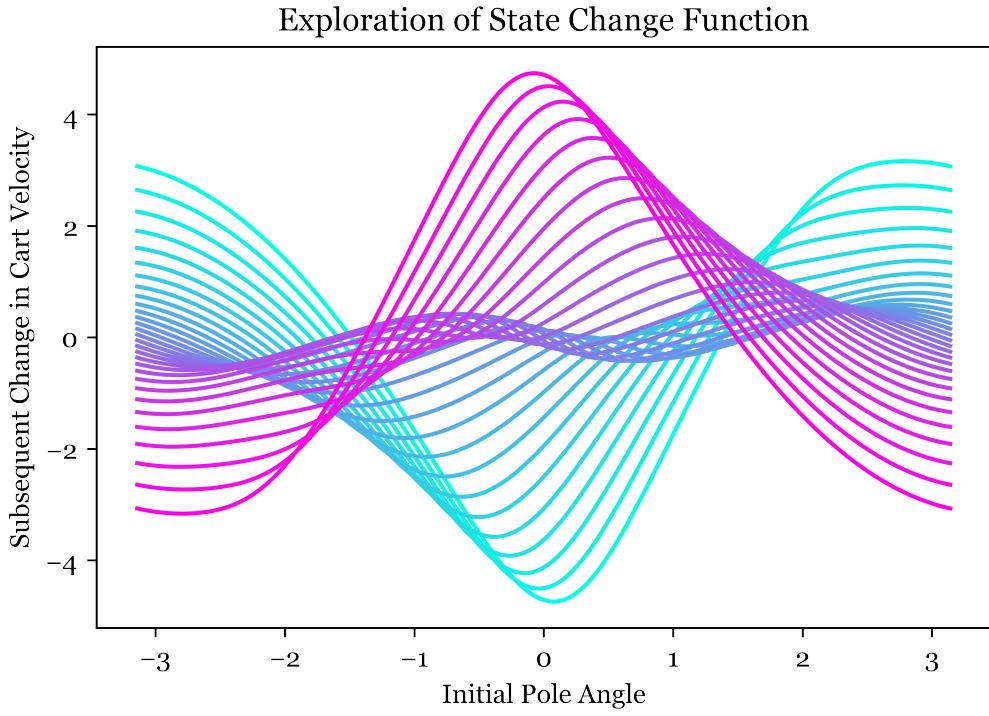
## 2.2 Effect of Varying Initial Pole Angle and Angular Velocity

Text(0, 0.5, 'Subsequent Change in Cart Velocity')



The plot above shows how the change in cart velocity at the next time step is a highly nonlinear function of the two variables that we are changing in the above plot - the intial pole angle and angular velocity. The more cyan curves are an intial angle around  $-\pi$  and the more magenta curves are an initial angle around  $\pi$ .

Text(0, 0.5, 'Subsequent Change in Cart Velocity')



This is the same as the plot above but this time plotting change in cart velocity as a function of initial pole angle, varying the initial angular velocity from -15 in cyan to +15 in magenta. It looks like a linear estimator of  $Y$  from  $X$  will perform poorly, but looking at the central, dense part of the graph above (which represents a small initial angular velocity) we can see the curvature of the lines are much smaller. Keeping in mind the truncated Taylor series approximation to a function for small deviations, there is some scope to use linear estimation especially for small deviations about an equilibrium.

`Text(0.06, 0.5, 'Initial Pole Angular Velocity')`

In the plot above I vary initial pole angle in the x direction and initial pole angular velocity in the y direction and show the changes in the state variables in the contour plots. It is quite hard to find linearity in the plots.

### 3 Task 1.3: Linear Model

I assume that the change in state is a linear function of the current state:

$$Y = f(X) = XC$$

where  $C$  is a  $4 \times 4$  matrix. First I generate 500 input/output pairs by randomly generating 500 initial states and then performing 1 step (one call to `perform_action`) to generate 500 subsequent states.

### 3.1 Generating Training Data

### 3.2 Finding the Least-Squares Fit

The Moore-Penrose pseudoinverse can then be used to construct the least-squares matrix that maps  $X$  to  $Y$ :

$$Y = XC$$
$$X^+Y = X^+XC = C$$

### 3.3 Evaluating the Linear Estimator

We can now compare our linear model to the experimental data we plotted in task 1.2.

### 3.4 Plotting Predicted State Change Against Target State Change

The plot above shows one dot per training data element, and I plotted one graph per element in the state change vector. For each dot, the x coordinate is the target change in state as returned from `perform_action` and the y coordinate is the predicted change as returned by the linear estimator. So for perfect predictions, these would be equal and the points would all lie on the line  $y = x$  shown in dotted black.

It is clear that the changes in cart position and pole angle are well predicted by the linear estimator, with the points lying close to the ideal line. The change in cart velocity at least has the central mass of points lying near the ideal line, but with many outliers at the edges, while the change in pole angular velocity is very poorly explained by the linear model with the points avoiding the ideal line.

### 3.5 Linear Model for Small Deflections Only

In the following cells I have repeated the above procedure, but limited the training data to small deflections from equilibrium (initial states between -0.25 and 0.25). Thinking back to the Taylor series approximation for a function I expected this scheme to yield better predictions near the equilibrium.

It is clear that the linear model is working nicely for small deflections, with all the points near to the ideal lines. Perhaps we could build a good model by using many linear models centered around different points, which is the motivation behind moving to a nonlinear model.

### 3.6 Comparing Prediction Contour Plots to Target Contour Plots

I generated 4 contour plots for the predicted state change of each the state vector elements as predicted by the linear model. I am varying the initial pole angle and angular velocity again, so I am aiming for these contour plots to be identical to the final 4 I plotted in task 1.2. Note that we are back to using the linear estimate for large deviations.

The contours from the linear estimator are in white and the target contours are in black. Clearly the nonlinearities in the data have been discarded and the linear estimator has picked up on only the lowest-frequency linear trends. I would expect this model to miss most of the nuance in the motion of the cartpole, especially in the nonlinear rotational motion of the pole.

However, the linear estimator works better when we plot contours while varying the initial cart velocity and pole angular velocity as above. This is the pair of variables I didn't plot contours for in task 1.2, but here we can see how the 2 trends on the left are closer to linear and the white and black contour lines are in close agreement. This is in agreement with the results from the scatter plots where the predictions for change in cart position and pole angle were much more accurate.

The elements that are well predicted by the linear model are the change in cart location and pole angle. The change in cart location will be quite similar to the velocity of the cart times the timestep (neglecting acceleration during the timestep), so it is expected that we could predict this as a linear function of the velocity. The change in pole angle is similar in that it will be quite similar to the angular velocity.

## 4 Task 1.4: Prediction of System Evolution

Now we can use the linear model to predict the system evolution by adding the prediction for the state change to the current state and repeating. This can be used as an alternative to the analytic system dynamics, so I have reproduced the three plots from task 1.1 but now using the linear model rather than the actual dynamics.

It is vital to remap the pole angle to the range  $-\pi$  to  $\pi$  during this process. If this is not done, the angle grows beyond anything seen in the training data and the result is an abnormally large input of energy to the system which causes divergence.

### 4.1 Small Oscillations

The plot above is initialised the same as the small oscillations case from the first plot in this report, and shows the time evolution of the pole angular velocity plotted against cart velocity. The target behaviour is shown in orange, which is the data from the small oscillation plot using the analytic dynamics. The linear estimation time evolution is the curve that starts in magenta at initialisation and turns more cyan as time goes on.

The linear model successfully brings the system into an oscillating state, shown by the loops about the origin, although these oscillations are of much larger amplitude than they should be (the loops are larger than the orange line). After a few oscillations, the system spontaneously shifts into a regime where it is still oscillating but also moving to the right, shown by the loops shifted to the right. It's important to note that this is a particularly lucky result based on favourable random generation of training data - most plots I made with other training data diverged from the oscillating regime in fewer time steps. The random increase in kinetic energy indicate that conservation of energy is a nonlinear detail in the data that the linear model struggles to capture.

The plot above is the same as the previous one, but starting with a larger pole angular velocity so that the oscillations are larger. Again as time progresses the line goes from magenta to cyan, and we can see that both the pole angular velocity and cart velocity are increasing. The outward spiral is indicative of an unstable system; again the linear model can't enforce conservation of energy. However, the spiral structure does indicate that the model has some understanding that a larger cart velocity implies a smaller pole angular velocity.

This final plot has the pole angular velocity plotted against the pole angle as the system evolves. The smooth phase space from task 1.1 has been replaced by incoherent jagged lines which don't reflect the motion of the cartpole. This was to be expected because in the scatter plots that showed

prediction accuracy, the pole angular velocity was the least well-explained by the linear model. So especially at the large angular velocities in this case, where nonlinear effects dominate, the linear model is incapable of explaining the data.