Modelling and Simulation ESS101
1 November 2024, Final Exam

This exam contains 12 pages (including this cover page) and 4 problems.

You are allowed to use the following material:

- *Modelling And Simulation, Lecture notes for the Chalmers course ESS101*, by S. Gros (annotations are not allowed)
- *Mathematics Handbook* (Beta)
- *Physics Handbook*
- Chalmers approved calculator
- Formula sheet, appended to the exam.

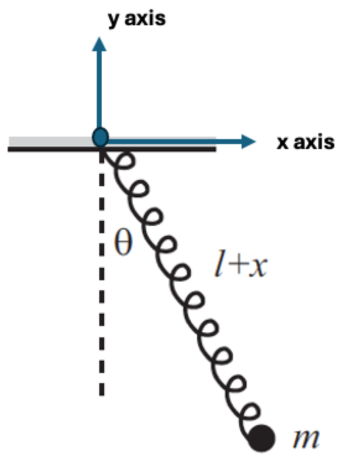
- Organize your work in a reasonably neat and coherent way. Work scattered all over the page without a clear ordering may receive less credit.
- Mysterious or unsupported answers will not receive credit, but an incorrect answer supported by substantially correct calculations and explanations will receive partial credit.
- None of the proposed questions require extremely long computations. If you get caught in endless algebra, you have probably missed the simple way of doing it.
- The passing grade will be given at 20 points, grade 4 at 27 and the top grade at 34 points.

Problem	Points	Score
1	12	
2	10	
3	7	
4	11	
Total:	40	

Best of luck to all !!

Examiner: Yasemin Bekiroglu, +46 70 148 72 71

1. Consider a pendulum made of a spring with a mass m on the end (see the figure below). The spring is arranged to lie in a straight line (which we can arrange by wrapping the spring around a rigid massless rod). The equilibrium length of the spring is l. Let the spring have length $l + x(t)$, and let its angle with the vertical be $\theta(t)$. Assuming that the motion takes place in a vertical plane, you will find the equations of motion for x and θ. The kinetic energy may be broken up into the radial and tangential parts, so we have an additional term $\frac{1}{2}m(l+x)^2\dot{\theta}^2$, besides the usual term with the velocity. The potential energy comes from both gravity and the spring constant k, so we have an additional term stemming from the spring $\frac{1}{2}kx^2$, besides the usual term including gravity.



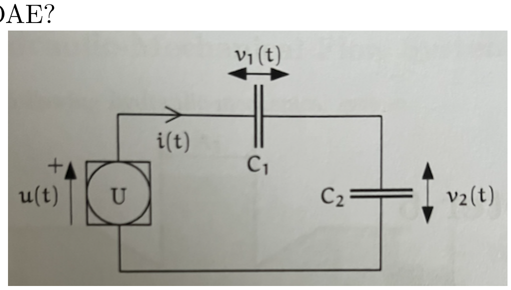
(a) (4 points) First determine the Lagrange function for the system.

(b) (5 points) Then write Euler Lagrange equations that describe the dynamics of this system.

(c) (3 points) Consider the following electric circuit where an ideal voltage source feeds two capacitors. Let $\mathbf{u}(t)$ be the input. The system is described by the equations

$$\begin{aligned} C_1\dot{\mathbf{v}}_1(t) - \mathbf{i}(t) &= 0 \\ C_2\dot{\mathbf{v}}_2(t) - \mathbf{i}(t) &= 0 \\ \mathbf{v}_1(t) + \mathbf{v}_2(t) &= \mathbf{u}(t) \end{aligned}$$

What is the index of this DAE?



Solution:

(a)

$$q = \begin{bmatrix} x \\ \theta \end{bmatrix}$$

$$T = \frac{1}{2} m \left(\dot{x}^2 + (\ell + x)^2 \dot{\theta}^2 \right) = \frac{1}{2} \dot{q}^T \underbrace{\begin{bmatrix} m & 0 \\ 0 & m(\ell + x)^2 \end{bmatrix}}_{W(q)} \dot{q}$$

$$V = -mg(\ell + x)\cos\theta + \frac{1}{2} k x^2$$

a) $$L = T - V = \frac{1}{2} m \left(\dot{x}^2 + (\ell + x)^2 \dot{\theta}^2 \right) + mg(\ell + x)\cos\theta - \frac{1}{2} k x^2$$

b) $$\nabla_{\dot{q}} L = \nabla_{\dot{q}} T - \underbrace{\nabla_{\dot{q}} V}_{0}$$

$$= W(q)\dot{q}$$

$$\frac{d}{dt} \nabla_{\dot{q}} L = \frac{\partial}{\partial q} \left(W(q)\dot{q} \right) \dot{q} + W(q)\ddot{q}$$

$$= \frac{\partial}{\partial q} \left(\begin{bmatrix} \dot{x} m \\ m(\ell + x)^2 \dot{\theta} \end{bmatrix} \right) \dot{q} + W(q)\ddot{q}$$

$$= \begin{bmatrix} 0 & 0 \\ 2m(\ell + x)\dot{\theta} & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} m\ddot{x} \\ m(\ell + x)^2 \ddot{\theta} \end{bmatrix}$$

$$\frac{d}{dt} \nabla_{\dot{q}} L = \begin{bmatrix} 0 \\ 2m(\ell + x)\dot{x}\dot{\theta} \end{bmatrix} + \begin{bmatrix} m\ddot{x} \\ m(\ell + x)^2 \ddot{\theta} \end{bmatrix}$$

$$\nabla_q L = \nabla_q T - \nabla_q V$$

$$\nabla_q L = \begin{bmatrix} m(\ell+x)\dot{\theta}^2 \\ 0 \end{bmatrix} - \begin{pmatrix} -mg\cos\theta + kx \\ mg\sin\theta\,(\ell+x) \end{pmatrix}$$

$$\nabla_q L = \begin{pmatrix} m(\ell+x)\dot{\theta}^2 + mg\cos\theta - kx \\ -mg\sin\theta\,(\ell+x) \end{pmatrix}$$

E-L: $$0 = \frac{d}{dt}\nabla_{\dot{q}} L - \nabla_q L$$

$$0 = \begin{bmatrix} m\ddot{x} \\ 2m(\ell+x)\dot{x}\dot{\theta} + m(\ell+x)^2\ddot{\theta} \end{bmatrix} - \begin{pmatrix} m(\ell+x)\dot{\theta}^2 + mg\cos\theta - kx \\ -mg\sin\theta\,(\ell+x) \end{pmatrix}$$

$$0 = \begin{pmatrix} m\ddot{x} - m(\ell+x)\dot{\theta}^2 - mg\cos\theta + kx \\ 2m(\ell+x)\dot{x}\dot{\theta} + m(\ell+x)^2\ddot{\theta} + mg\sin\theta(\ell+x) \end{pmatrix}$$

The resulting two equations:

$$\begin{cases} 0 = m\ddot{x} - m(\ell+x)\dot{\theta}^2 - mg\cos\theta + kx \\ 0 = 2m(\ell+x)\dot{x}\dot{\theta} + m(\ell+x)^2\ddot{\theta} + mg\sin\theta(\ell+x) \end{cases}$$

$$\begin{cases} m\ddot{x} = m(\ell+x)\dot{\theta}^2 + mg\cos\theta - kx \\ -mg\sin\theta = m2\dot{x}\dot{\theta} + m(\ell+x)\ddot{\theta} \end{cases}$$

(b) Check how many derivations are needed to be able to solve for the derivatives of all variables. Since $\dot{\mathbf{v}}_1(t)$ and $\dot{\mathbf{v}}_2(t)$ can be resolved from the first two equations directly, only an equation for $\frac{d}{dt}i(t)$ is needed. Differentiation of the third equation yields:

$\dot{\mathbf{v}}_1(t) + \dot{\mathbf{v}}_2(t) = \dot{\mathbf{u}}(t)$.

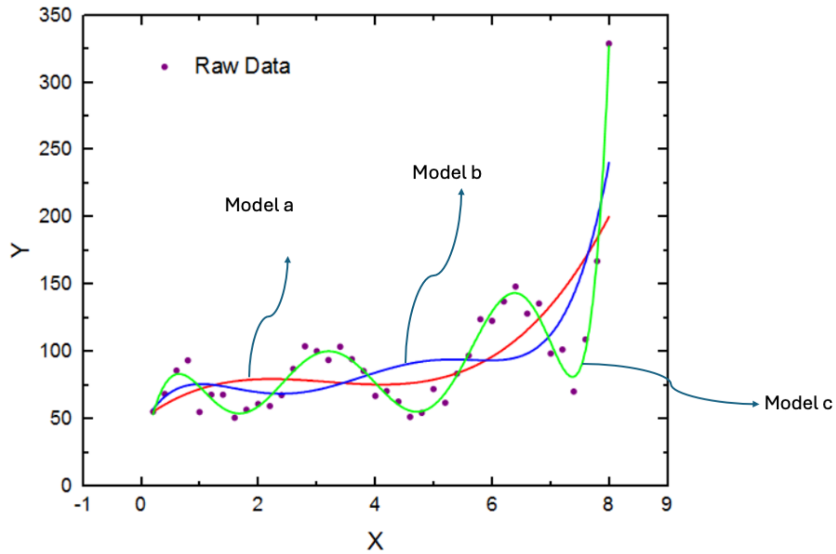
Insertion of the expressions for $\dot{\mathbf{v}}_1(t)$ and $\dot{\mathbf{v}}_2(t)$ results in:

$(\frac{1}{C_1} + \frac{1}{C_2})i(t) = \dot{\mathbf{u}}(t)$.

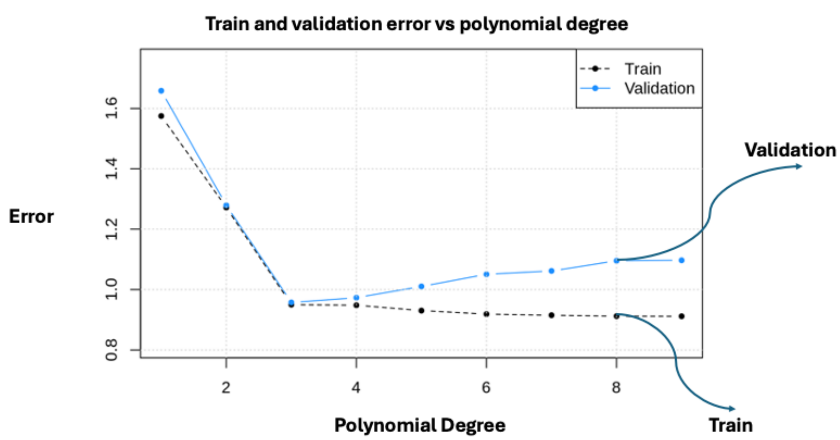
A second differentiation of this expression gives $\frac{d}{dt}i(t)$, and hence the index of the system is 2.

2. (a) (3 points) Given the data pairs $\{[x(i), y(i)]\} = \{[0,2],[1,3],[-1,3]\}$. Using a least squares approach capturing the relationship between x and y, we write $\hat{\mathbf{y}} = \boldsymbol{\theta}^\top \boldsymbol{\varphi}$ (where $\boldsymbol{\varphi}$ is the *regression vector* holding the *regressors* $(1, x, ...)$). We need to construct f_N and R_N matrices to find the values in the parameter vector θ that would fit a second order polynomial to the data using the Linear Least Squares approach. Using the given three data points, construct the two matrices f_N and R_N. You do not need to calculate the final parameter vector values of θ.

(b) (1 point) Match the following models 1. Degree 3 polynomial, 2. Degree 5 polynomial, 3. Degree 9 polynomial, with the model fittings in the figures below, i.e. which model is used to plot the curve fittings. Explain your reasoning.



(c) (1 point) Given the following curve that shows how the model fit error changes with respect to increasing model order for a trainset and a validationset. Which model order (polynomial degree) would you choose to avoid overfitting, given this specific setup? Explain your reasoning.



(d) (2 points) Consider the following model a: $y(t) + ay(t-1) = bu(t-1) + e(t)$
model b: $y(t) = b_1 u(t-1) + b_2 u(t-2) + e(t)$,
where $u(t)$ and $y(t)$ denote the input ands output signal respectively, while $e(t)$ is unknown disturbance. What type of models are these?

(e) (3 points) For the model $y(t) + ay(t-1) = bu(t-1)$ five intermediate sums were calculated:

$\sum_{t=1}^{101} y^2(t) = 5.0$

$\sum_{t=1}^{101} y(t)u(t) = 1.0$
$\sum_{t=1}^{101} u^2(t) = 1.0$
$\sum_{t=2}^{102} y(t-1)y(t) = 4.5$
$\sum_{t=2}^{102} u(t-1)y(t) = 1.0$

Find the values of the parameter vector $\theta^T = [a \quad b]$ that minimizes the error between the predictions and true values.

Solution:

(a) Using the data vectors $x = [0, 1, -1]$ and $y = [2, 3, 3]$ in the least-squares estimate where $k = 2$ for fitting a second order polynomial, our f_N and R_N would be as follows:

$$\hat{\boldsymbol{\theta}}_N = R_N^{-1} f_N = \Big(\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\varphi}(i)\boldsymbol{\varphi}^\top(i)\Big)^{-1} \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\varphi}(i)y(i)$$

$$\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\varphi}(i)\boldsymbol{\varphi}^\top(i) = \frac{1}{N}\begin{bmatrix} N & \sum_{i=1}^{N} x(i) & \dots & \sum_{i=1}^{N} x^k(i) \\ \sum_{i=1}^{N} x(i) & \sum_{i=1}^{N} x^2(i) & \dots & \sum_{i=1}^{N} x^{k+1}(i) \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ \sum_{i=1}^{N} x^k(i) & \sum_{i=1}^{N} x^{k+1}(i) & \dots & \sum_{i=1}^{N} x^{2*k}(i) \end{bmatrix} \tag{1}$$

$$\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\varphi}(i)y(i) = \frac{1}{N}\begin{bmatrix} \sum_{i=1}^{N} y(i) \\ \sum_{i=1}^{N} x(i)y(i) \\ . \\ . \\ . \\ \sum_{i=1}^{N} x^k(i)y(i) \end{bmatrix} \tag{2}$$

$$R_N = \frac{1}{3}.\begin{bmatrix} 3 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix} \tag{3}$$

$$f_N = \frac{1}{3}.\begin{bmatrix} 8 \\ 0 \\ 6 \end{bmatrix} \tag{4}$$

Although it is not asked, the resulting parameter estimate then would be: $\hat{\boldsymbol{\theta}}_N = [a, b, c]^T$, $a = 2$, $b = 0$, $c = 1$ in $\hat{\mathbf{y}} = a + bx + cx^2 = \boldsymbol{\theta}^\top\boldsymbol{\varphi}$.

(b) Model a = degree 3, model b = degree 5, model c = degree 9.

(c) Around degree 3 or 4 should be good enough as increasing the order after that does not make a big change.

(d) ARX, FIR

(e) Let $\theta^T = [a \quad b]$ and $\varphi(t) = [-y(t-1) \quad u(t-1)]^T$ so that $y(t) = \theta^T\varphi(t)$. The least square estimate is (N=102):

$$\hat{\theta} = R_N^{-1} f_N = \left[\ \frac{1}{N}\sum_{t=2}^{N} \varphi(t)\varphi^T(t) \ \right]^{-1} \left[\ \frac{1}{N}\sum_{t=2}^{N} \varphi(t)y(t) \ \right]$$

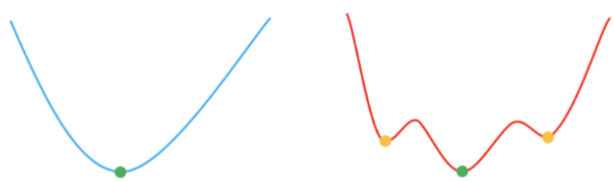
$$\begin{bmatrix} \sum_{t=1}^{101} y^2(t) & -\sum_{t=1}^{101} y(t)u(t) \\ -\sum_{t=1}^{101} y(t)u(t) & \sum_{t=1}^{101} u^2(t) \end{bmatrix}^{-1} \cdot \begin{bmatrix} -\sum_{t=2}^{102} y(t-1)y(t) \\ \sum_{t=2}^{102} u(t-1)y(t) \end{bmatrix} =$$

$$\begin{bmatrix} 5.0 & -1.0 \\ -1.0 & 1.0 \end{bmatrix}^{-1} \begin{bmatrix} -4.5 \\ 1.0 \end{bmatrix} = \begin{bmatrix} -0.875 \\ 0.125 \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{a}_N} \\ \hat{\mathbf{b}_N} \end{bmatrix}$$

3. (a) (2 points) Apply the Newton method to find the cube root of 3, i.e. solve $x^3 = 3$. Start with $x_0 = 0.5$ and iterate only once.

   (b) (3 points) Consider the function $\mathbf{f} : \mathbb{R}^2 \mapsto \mathbb{R}^2$, $f(x, y) = \begin{bmatrix} x - y + 2x^2 + y^2 \\ x^3 - 3xy + y^3 \end{bmatrix}$, for which we construct approximate solutions to the equation $f(x, y) = [0, 0]^T$ using (full step) Newton method and the initial guess $\mathbf{x} = [x_0, y_0] = [0, 1]$. Calculate the resulting solution from applying Newton iteration only once, i.e. $[x_1, y_1]$.

   (c) (1 point) When using the Newton method for optimization of the following two functions (in Figure (a) and (b) below) to find their extreme points, which one would be easier to optimize with Newton and what could be potential issues with the more challenging case?



(a) (b)

   (d) (1 point) Comment on the effect of the initial solution on the performance of the Newton Method. When could the choice of the initial solution lead to potential problems?

Solution:

(a) We need to solve for x in $x^3 = 3$ using the Newton method, so we organize it into $x^3 - 3 = 0$ and apply the Newton update rule once:

$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$, substituting for $x_0 = 0.5$ and $f'(x_0) = 3{x_0}^2 = 0.75$ yields $x_1 = 0.5 - \frac{0.5^3 - 3}{0.75} = 4.3333$

(b) Plugging in the following in the update formula: $\partial f(x, y) = \begin{bmatrix} 1 + 4x & 2y - 1 \\ 3x^2 - 3y & -3x + 3y^2 \end{bmatrix}$,

$det \partial f(x, y) = (1+4x)(3y^2-3x)-(3x^2-3y)(2y-1)$, $\partial f(x, y)^{-1} = \frac{1}{det \partial f(x,y)} \begin{bmatrix} -3x + 3y^2 & 1 - 2y \\ 3y - 3x^2 & 1 + 4x \end{bmatrix}$

Thus, $\begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} x_{k-1} \\ y_{k-1} \end{bmatrix} - \partial f(x_{k-1}, y_{k-1})^{-1}.f(x_{k-1}, y_{k-1})$

Beginning with $x_0 = 0, y_0 = 1$, $x_1 = \frac{1}{6}$, $y_1 = \frac{5}{6}$

(c) The first case (a) is a convex function therefore Newton can find a solution. In the second case (b) Newton might get stuck in a local minima.

(d) The initial solution should be chosen well so it is not too far from the actual solution, the step size also needs to be adapted (reduced step size might help to deal with divergence issues). When the Jacobian is singular, it fails, so if the initial solution is close to a local minima, Newton might get stuck there.

4. (a) (1 point) Consider a Runge-Kutta scheme defined by the following Butcher array:

$$\begin{array}{c|cc} 1/3 & 5/12 & -1/12 \\ 1 & 3/4 & 1/4 \\ \hline & 3/4 & 1/4 \end{array}$$

Is the RK scheme above explicit or implicit? How many stages are there?

(b) (1 point) For the RK scheme above, write the equations describing an update of the solution sequence $\{x_k\}$.

(c) (3 points) We will use Euler's standard method to numerically solve the differential equation:

$$\dot{\mathbf{x}}(t) = \mathbf{x}^2(t)$$

where $x(0) = 1$ (i.e. the true value in the beginning at time $t = 0$ is 1.). The exact solution is:

$$x(t) = \frac{1}{1-t}$$

(i) What approximate value for $x(0.2)$ is obtained (for time $t = 0.2$) if the step length is $\Delta t = 0.1$? Calculate your approximation $x_k \approx x(0.2)$ for time $t = 0.2$ starting from $x_{k=0} = x(0)$, using standard Euler RK scheme, i.e. $\mathbf{x_{k+1}} = \mathbf{x_k} + \Delta t \mathbf{f(x_k)}$?

(d) (2 points) (ii) What approximate value for x(0.2) is obtained (for time $t = 0.2$) if the step length is $\Delta t = 0.2$, again using standard Euler RK scheme as above in (i)?

(e) (4 points) What error estimates can be achieved by comparing the result from above (i) and (ii) cases? How well does this align with true errors? (Hint: Euler's method has a global error proportional to the step length, i.e. it is in the order of $\mathcal{O}(\Delta t)$. You need to find an error estimate for each approximation that satisfies this criteria.)

Solution:

(a) The RK scheme is implicit and has 2 stages.

(b)

$$\begin{aligned}
\mathbf{K}_1 &= \mathbf{f}\left(\mathbf{x}_k + \frac{5\Delta t}{12} \cdot \mathbf{K}_1 - \frac{\Delta t}{12} \cdot \mathbf{K}_2, \mathbf{u}(t_k + \frac{\Delta t}{3})\right) \\
\mathbf{K}_2 &= \mathbf{f}\left(\mathbf{x}_k + \frac{3\Delta t}{4} \cdot \mathbf{K}_1 + \frac{\Delta t}{4} \cdot \mathbf{K}_2, \mathbf{u}(t_k + \Delta t)\right) \\
\mathbf{x}_{k+1} &= \mathbf{x}_k + \frac{3\Delta t}{4}\mathbf{K}_1 + \frac{\Delta t}{4}\mathbf{K}_2
\end{aligned}$$

(c) With the standard Euler method $x(t+\Delta t) \approx x_{k+1} = x_k + \Delta t f(x_k)$ the solution to the differential equation $\dot{\mathbf{x}} = \mathbf{x}^2 = f(x(t))$ is given by $x_{k+1} = x_k + \Delta t x(t)^2$.

For $\Delta t = 0.1$, and $x_0 = x(0) = 1$ we get the sequence

$$x_0 = x(t=0) = 1$$
$$x(t=0.1) \approx x_1 = x_0 + 0.1x_0^2 = 1 + 0.1 = 1.1$$
$$x(t=0.2) \approx x_2 = x_1 + 0.1x_1^2 = 1.1 + 0.1.1.1^2 = 1.221$$

(d) For $\Delta t = 0.2$, and $x_0 = x(0) = 1$ we get the sequence

$$x_0 = x(t=0) = 1$$
$$x(t=0.2) \approx x_1 = x_0 + 0.2x_0^2 = 1 + 0.2 = 1.2$$

(e) Euler's method has a global error proportional to the step length Δt, ie. it is in the order of $\mathcal{O}(\Delta t)$. Assume that the global error for the step length $\Delta t' = 0.1$ is E. The global error for the step length $\Delta t'' = 0.2$ should then be $2E$. Now let $x(t = 0.2)$ denote the exact (but in the error estimation, unknown) solution. Let us further denote our approximations for each step length at time 0.2 as:
$x_{\Delta t'} = 1.221$ and $x_{\Delta t''} = 1.2$

Calculating the error between the true value $x(0.2)$ and our two approximations:

$$E = x(0.2) - x_{\Delta t'}$$
$$2E = x(0.2) - x_{\Delta t''}$$

$$E = x(0.2) - 1.221$$
$$2E = x(0.2) - 1.2$$

from above two equations we can get E as $E = 0.021$ and according to this error: $x(0.2) = 1.242$. So the error estimate is $E = 0.021$ when the step length Δt is 0.1 and 2E when the step length Δt is 0.2. We know that the exact solution to the differential equation at $t = 0.2$ is in fact:

$$x(0.2) = \frac{1}{1 - 0.2} = 1.25$$

The true global errors are then $1.25 - 1.221 = 0.029$ for the case $\Delta t = 0.1$ (where we just calculated the error estimate E as 0.021 which is very close to the true error) and $1.25 - 1.2 = 0.05$ for $\Delta t = 0.2$ (where the error estimate is 0.042). These values are well in line with the errors estimated above.

Appendix: some possibly useful formula

- Lagrange mechanics is built on the equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} - \frac{\partial \mathcal{L}}{\partial \mathbf{q}} = \mathbf{Q}, \qquad \mathcal{L}\left(\mathbf{q},\dot{\mathbf{q}},\mathbf{z}\right) = \mathcal{T} - \mathcal{V} - \mathbf{z}^\top \mathbf{C}, \qquad \mathbf{C} = 0, \qquad \langle \delta\mathbf{q},\, \mathbf{Q}\rangle = \delta W,\, \forall\, \delta\mathbf{q} \tag{5}$$

  The kinetic and potential energy of a point mass are given by:

$$\mathcal{T} = \frac{1}{2} m \dot{\mathbf{p}}^\top \dot{\mathbf{p}}, \qquad \mathcal{V} = mg\mathbf{p}_3 \tag{6}$$

  respectively, where $\mathbf{p} \in \mathbb{R}^3$ is the position of the mass in a cartesian reference frame having the third coordinate as the vertical axis pointing up. The generalized forces are identical to the external forces applied to a point mass if the position of that point is expressed in cartesian coordinates in the generalized coordinates $\mathbf{q}$.

- In the case $\mathcal{T} = \frac{1}{2}m\dot{\mathbf{q}}^\top W \dot{\mathbf{q}}$ with W constant $\mathcal{V} = \mathcal{V}\left(\mathbf{q}\right)$ and $\mathbf{C} = \mathbf{C}\left(\mathbf{q}\right)$, the Lagrange equations simplify to the dynamics in the semi-explicit index-3 DAE form:

$$\dot{\mathbf{p}} = \mathbf{v} \tag{7a}$$

$$W\dot{\mathbf{v}} + \frac{\partial \mathbf{C}}{\partial \mathbf{q}}^\top \mathbf{z} = \mathbf{Q} - \frac{\partial \mathcal{V}}{\partial \mathbf{q}}^\top \tag{7b}$$

$$0 = \mathbf{C}\left(\mathbf{q}\right) \tag{7c}$$

- The Implicit Function Theorem (IFT) guarantees that a nonlinear set of equations

$$\mathbf{r}\left(\mathbf{y},\mathbf{z}\right) = 0 \tag{8}$$

  "can be solved" in terms of $\mathbf{z}$ for a given $\mathbf{y}$ iff the Jacobian $\frac{\partial \mathbf{r}(\mathbf{y},\mathbf{z})}{\partial \mathbf{z}}$ is full rank at the solution. More specifically, it guarantees that there is a function $\phi\left(\mathbf{y}\right)$ such that

$$\mathbf{r}\left(\mathbf{y},\phi\left(\mathbf{y}\right)\right) = 0 \tag{9}$$

  holds in the neighborhood of the point $\mathbf{y}$ where the Jacobian is evaluated. Furthermore, the IFT specifies that:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{y}} = -\frac{\partial \mathbf{r}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{r}}{\partial \mathbf{y}} \tag{10}$$

- For solving a problem $\mathbf{r}\left(\mathbf{x}\right) = 0$, Newton iterates:

$$\mathbf{x} \leftarrow \mathbf{x} - \alpha \frac{\partial \mathbf{r}}{\partial \mathbf{x}}^{-1}\mathbf{r} \tag{11}$$

  until $\mathbf{r}\left(\mathbf{x}\right) \approx 0$ where $\alpha \in [0, 1]$

- Runge-Kutta methods are described by:

$$\begin{array}{c|ccc} c_1 & a_{11} & \dots & a_{1s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s1} & \dots & a_{ss} \\ \hline & b_1 & \dots & b_s \end{array}$$

$$\mathbf{K}_j = \mathbf{f}\left(\mathbf{x}_k + \Delta t \sum_{i=1}^{s} a_{ji}\mathbf{K}_i,\, \mathbf{u}\left(t_k + c_j \Delta t\right)\right), \quad j = 1, \dots, s \tag{12a}$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta t \sum_{i=1}^{s} b_i \mathbf{K}_i \tag{12b}$$

- For ERK methods, the relationship between the (minimum) number of stages s to the order o is given by:

s	1	2	3	4	6	7	9	11	...
o	1	2	3	4	5	6	7	8	...

Table 1: Stage to order of ERK methods

- Collocation methods use:

$$\dot{\mathbf{x}}(t_k + \Delta t \cdot \tau) \approx \dot{\hat{\mathbf{x}}}(t_k + \Delta t \cdot \tau) = \sum_{i=1}^{s} \mathbf{K}_i \ell_i(\tau), \quad \tau \in [0,\, 1] \tag{13}$$

$$\mathbf{x}(t_k + \Delta t \cdot \tau) \approx \hat{\mathbf{x}}(t_k + \Delta t \cdot \tau) = \mathbf{x}_k + \Delta t \sum_{i=1}^{s} \mathbf{K}_i L_i(\tau) \tag{14}$$

where the Lagrange polynomials are given by:

$$\ell_i(\tau) = \prod_{j=1, j\neq i}^{s} \frac{\tau - \tau_j}{\tau_i - \tau_j}, \quad \text{and} \quad L_i(\tau) = \int_0^{\tau} \ell_i(\xi)\mathrm{d}\xi \tag{15}$$

The Lagrange polynomials satisfy the conditions of

$$\text{Orthogonality:} \quad \int_0^1 \ell_i(\tau)\ell_j(\tau)\,\mathrm{d}\tau = 0 \quad \text{for} \quad i \neq j \tag{16a}$$

$$\text{Punctuality:} \quad \ell_i(\tau_j) = \begin{cases} 1 & \text{if} \quad j = i \\ 0 & \text{if} \quad j \neq i \end{cases} \tag{16b}$$

and enforce the collocation equations (for $j = 1, \dots, s$):

$$\dot{\hat{\mathbf{x}}}(t_k + \Delta t \cdot \tau_j) = \mathbf{f}\left(\hat{\mathbf{x}}(t_k + \Delta t \cdot \tau_j),\, \mathbf{u}\left(t_k + \Delta t \cdot \tau_j\right)\right), \qquad \text{in the explicit ODE case} \tag{17a}$$

$$\mathbf{F}\left(\dot{\hat{\mathbf{x}}}(t_k + \Delta t \cdot \tau_j), \hat{\mathbf{x}}(t_k + \Delta t \cdot \tau_j),\, \mathbf{u}\left(t_k + \Delta t \cdot \tau_j\right)\right) = 0, \qquad \text{in the implicit ODE case} \tag{17b}$$

$$\mathbf{F}\left(\dot{\hat{\mathbf{x}}}(t_k + \Delta t \cdot \tau_j), \hat{\mathbf{z}}_j, \hat{\mathbf{x}}(t_k + \Delta t \cdot \tau_j),\, \mathbf{u}\left(t_k + \Delta t \cdot \tau_j\right)\right) = 0, \qquad \text{in the fully-implicit DAE case} \tag{17c}$$

- Gauss-Legendre collocation methods select the set of points $\tau_{1,\dots,s}$ as the zeros of the (shifted) Legrendre polynomial:

$$P_s\left(\tau\right) = \frac{1}{s!}\frac{\mathrm{d}^s}{\mathrm{d}\tau^s}\left[\left(\tau^2 - \tau\right)^s\right] \tag{18}$$

They achieve the order $\|\mathbf{x}_N - \mathbf{x}\left(t_{\mathrm{f}}\right)\| = \mathcal{O}\left(\Delta t^{2s}\right)$.

- Maximum-likelihood estimation is based on

$$\max_{\boldsymbol{\theta}} \quad \mathbb{P}\left[e_k = y_k - \hat{y}_k \quad \text{for} \quad k = 1, \dots, N \mid \boldsymbol{\theta}\right] \tag{19}$$

If the noise sequence is uncorrelated, then

$$\mathbb{P}\left[e_k = y_k - \hat{y}_k \quad \text{for} \quad k = 0, \dots, N \mid \boldsymbol{\theta}\right] = \prod_{k=1}^{N} \mathbb{P}\left[e_k = y_k - \hat{y}_k \mid \boldsymbol{\theta}\,\right] \tag{20}$$

- The solution of a linear least-squares problem

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{2}\left\|A\boldsymbol{\theta} - \mathbf{y}\right\|^2_{\Sigma_e^{-1}} \tag{21}$$

reads as:

$$\hat{\boldsymbol{\theta}} = \left(A^\top \Sigma_e^{-1} A\right)^{-1} A^\top \Sigma_e^{-1} \mathbf{y} \tag{22}$$

and the covariance of the parameter estimation based is given by the formula:

$$\Sigma_{\hat{\boldsymbol{\theta}}} = \left(A^\top \Sigma_e^{-1} A\right)^{-1} \tag{23}$$

- In system identification, given the a plant $G(z)$ and a noise $H(z)$ model description, the one-step-ahead predictor $\hat{y}(k|k-1)$ can be retrieved with

$$H(z)\hat{y}(z) = G(z)u(z) + (H(z) - 1)y(z) \tag{24}$$

- The Gauss-Newton approximation in an optimization problem

$$\min_{\mathbf{x}} \quad J(\mathbf{x}) = \frac{1}{2} \|\mathbf{R}(\mathbf{x})\|^2 \tag{25}$$

uses the approximation:

$$\frac{\partial^2 J}{\partial \mathbf{x}^2} \approx \frac{\partial R}{\partial \mathbf{x}}^\top \frac{\partial R}{\partial \mathbf{x}} \tag{26}$$

- The solution to an LTI system $\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}$ is given by:

$$\mathbf{x}(t) = e^{At}\mathbf{x}(0) + \int_0^t e^{A(t-\tau)} B\mathbf{u}(\tau)\mathrm{d}\tau \tag{27}$$

and the transformation state-space to transfer function is given by:

$$G(s) = C\left(sI - A\right)^{-1} B + D \tag{28}$$

- $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $det(A) = ad - bc$
- $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$, $det(A) = a.det(\begin{bmatrix} e & f \\ h & i \end{bmatrix}) - b.det(\begin{bmatrix} d & f \\ g & i \end{bmatrix}) + c.det(\begin{bmatrix} d & e \\ g & h \end{bmatrix})$
- $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $det(A) = ad - bc$, $A^{-1} = \frac{1}{det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$
- $\alpha = \mathbf{x}^T A\mathbf{x}$, where A is a symmetric matrix and $\mathbf{x}$ is $n \times 1$, A is $n \times n$, and A does not depend on $\mathbf{x}$, then, $\frac{\partial \alpha}{\partial \mathbf{x}} = 2\mathbf{x}^T A$.