ESS101 Modelling and simulation Teacher: Bo Egardt, tel 031-7723721 (visits twice)
January 2020

This exam contains 14 pages (including this cover page) and 5 problems.

You are allowed to use the following material:

- *Modelling And Simulation, Lecture notes for the Chalmers course ESS101*, by S. Gros (with brief annotations)
- *Mathematics Handbook* (Beta)
- *Physics Handbook*
- Chalmers approved calculator
- Formula sheet, appended to the exam

Guidelines:

- 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 nominal grade limits are 22 (3), 28 (4) and 34 (5).

Problem	Points	Score
1	8	
2	8	
3	10	
4	10	
5	4	
Total:	40	

GOOD LUCK !!

1. **Lagrange modelling** Consider the system depicted in Fig. 1, made of two masses m. The upper mass is sliding without friction on the surface depicted in light blue, provided by the scalar equation $C(\mathbf{p}) = 0$. The red link between the masses consists of of a spring-damper combination. The spring has a rest-length L, and its potential energy is given by

$$V_{\text{spring}} = \frac{1}{2} K \left(\|\mathbf{p}_1 - \mathbf{p}_2\| - L \right)^2 , \tag{1}$$

whereas the damper produces a virtual work given by

$$\delta W = -\gamma \left(\dot{\mathbf{p}}_1 - \dot{\mathbf{p}}_2 \right)^\top \left(\delta \mathbf{p}_1 - \delta \mathbf{p}_2 \right) . \tag{2}$$

(a) (4 points) Derive a system of (second-order) differential equations that describes the dynamics of the system.

(b) (4 points) What is the DAE approximating your model for $K \to \infty$?
Hint: In this case, the virtual work of the damper is zero.

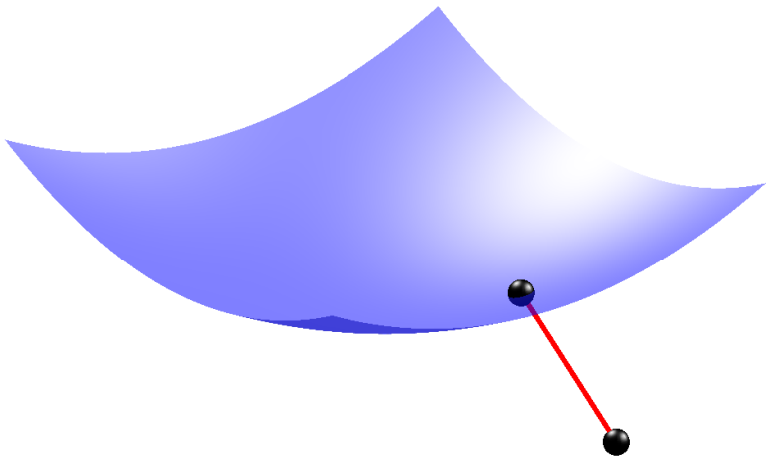


Figure 1: Illustration of the system of question 1.

Solution:

(a) We form the Lagrange function

$$\mathcal{L} = \frac{1}{2} m \sum_{k=1}^{2} \left(\dot{\mathbf{p}}_k^\top \dot{\mathbf{p}}_k - g\, \mathbf{e}^\top \mathbf{p}_k \right) - V_{\text{spring}} - z C(\mathbf{p}_1) \tag{3}$$

where $\mathbf{e}^\top = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$. We will use:

$$\mathbf{q} = \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \end{bmatrix} \tag{4}$$

The dynamics can be constructed using:

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}^\top = m \begin{bmatrix} \ddot{\mathbf{q}}_1 \\ \ddot{\mathbf{q}}_2 \end{bmatrix} \tag{5}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}}^\top = -mg \begin{bmatrix} \mathbf{e} \\ \mathbf{e} \end{bmatrix} - K \left(1 - \frac{L}{\|\mathbf{p}_1 - \mathbf{p}_2\|} \right) \begin{bmatrix} \mathbf{p}_1 - \mathbf{p}_2 \\ \mathbf{p}_2 - \mathbf{p}_1 \end{bmatrix} - \frac{\partial C}{\partial \mathbf{q}}^\top z \tag{6}$$

The generalized (non-conservative) forces caused by the damper are given by (2) (as the

coefficients of the virtual displacements):

$$\mathbf{Q}_{\mathbf{p}_1} = -\gamma\left(\dot{\mathbf{p}}_1 - \dot{\mathbf{p}}_2\right) \tag{7a}$$

$$\mathbf{Q}_{\mathbf{p}_2} = \quad \gamma\left(\dot{\mathbf{p}}_1 - \dot{\mathbf{p}}_2\right) \tag{7b}$$

We can then assemble the model equations

$$m\begin{bmatrix} \ddot{\mathbf{p}}_1 \\ \ddot{\mathbf{p}}_2 \end{bmatrix} = -mg\begin{bmatrix} \mathbf{e} \\ \mathbf{e} \end{bmatrix} - K\left(1 - \frac{L}{\|\mathbf{p}_1 - \mathbf{p}_2\|}\right)\begin{bmatrix} \mathbf{p}_1 - \mathbf{p}_2 \\ \mathbf{p}_2 - \mathbf{p}_1 \end{bmatrix} - \gamma\begin{bmatrix} \dot{\mathbf{p}}_1 - \dot{\mathbf{p}}_2 \\ \dot{\mathbf{p}}_2 - \dot{\mathbf{p}}_1 \end{bmatrix} - \frac{\partial C}{\partial \mathbf{q}}^{\top} z \tag{8}$$

(b) In order to form the DAE approximation of this system, we replace the spring by a rigid link, i.e. we can e.g. introduce the new scalar constraint:

$$\mathcal{C} = \frac{1}{2}\|\mathbf{p}_1 - \mathbf{p}_2\|^2 - L^2 = 0 \tag{9}$$

We observe that

$$\frac{\partial \mathcal{C}}{\partial \mathbf{q}}^{\top} = \begin{bmatrix} \mathbf{p}_1 - \mathbf{p}_2 \\ \mathbf{p}_2 - \mathbf{p}_1 \end{bmatrix} \tag{10}$$

It can be verified that the Lagrange equations then yield similar expressions as (8), but where the potential energy stemming from the elastic link is replaced by:

$$-\frac{\partial \mathcal{C}}{\partial \mathbf{q}}^{\top} z_2 = -z_2\begin{bmatrix} \mathbf{p}_1 - \mathbf{p}_2 \\ \mathbf{p}_2 - \mathbf{p}_1 \end{bmatrix} \tag{11}$$

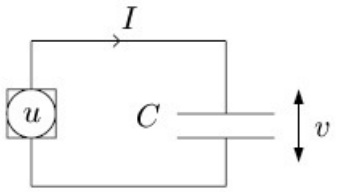
i.e. we obtain (also omitting the friction term as indicated by the hint)

$$m\begin{bmatrix} \ddot{\mathbf{p}}_1 \\ \ddot{\mathbf{p}}_2 \end{bmatrix} = -mg\begin{bmatrix} \mathbf{e} \\ \mathbf{e} \end{bmatrix} - z_2\begin{bmatrix} \mathbf{p}_1 - \mathbf{p}_2 \\ \mathbf{p}_2 - \mathbf{p}_1 \end{bmatrix} - \frac{\partial C}{\partial \mathbf{q}}^{\top} z_1, \tag{12}$$

where z_1 is the Lagrange multiplier introduced in (a). An interpretation of the constrained formulation of the problem is that the constraint $\mathcal{C}$ yields a "spring-like" force term whose magnitude is adjusted (via z_2) to force the link to have a constant length L.

2. **Differential-Algebraic and Implicit Differential Equations**

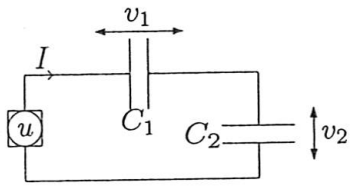
(a) (4 points) Consider the following electrical circuit, where an ideal voltage source is supplying a capacitor.



Let u be input signal and let $z = [v\ I]^T$ be generalized state vector.

1. Determine a DAE on the form
$$E\dot{z} + Fz = Gu$$
.
2. What is the index for the system?
3. How many states does the system have (i.e. how many independent initial conditions need to be specified in order to simulate the system)?

(b) (4 points) Assume that another capacitor is added to the circuit in (a), as depicted below.



1. Determine a DAE for the circuit, with u as input and using the variables v_1, v_2 and I.
2. What is now the index for the system?
3. How many states does the system have?

Solution:

(a) 1. The capacitor obeys the equation $I = C\frac{dv}{dt}$ and the voltage source gives $v = u$, giving the DAE
$$\begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{I} \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} v \\ I \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

2. The current I cannot be solved for from the algebraic equation, but differentiating gives
$$\dot{v} = \dot{u} \quad \Rightarrow \quad I = c\dot{u}$$
and another differentiation would provide $\dot{I}$. Hence, the original DAE has index 2.

3. Since $v = u$ and $I = c\dot{u}$, the variables are defined using only the input. Put differently, the initial conditions must satisfy $v(0) = u(0)$ and $I(0) = c\dot{u}(0)$. Hence, the number of states is 0.

(b) 1. The circuit is described by the DAE system
$$C_1\dot{v}_1 - I = 0 \tag{13}$$
$$C_2\dot{v}_2 - I = 0 \tag{14}$$
$$v_1 + v_2 = u \tag{15}$$

2. Again, in order to solve for I, a differentiation is needed:
$$\dot{v}_1 + \dot{v}_2 = \dot{u} \quad \Rightarrow \quad (\frac{1}{C_1} + \frac{1}{C_2})I = \dot{u} \tag{16}$$

and similarly to the case in (a), another differentiation would provide $\dot{I}$. The DAE is therefore of index 2.

3. The initial conditions must satisfy $v_1(0) + v_2(0) = u(0)$ and $\left(\frac{1}{C_1} + \frac{1}{C_2}\right)I(0) = \dot{u}(0)$, which leaves one degree of freedom. Hence, there is only one state.

3. **System Identification**

(a) (2 points) Consider the following ARX one-step-ahead predictor

$$\hat{y}(t) = ay(t-1) + bu(t-1). \tag{17}$$

Assume that the following data set is available

$$[y(0),\, y(1)] = [0,\, 1] \tag{18}$$
$$[u(0),\, u(1)] = [1,\, 0] \tag{19}$$

Write the predictor (17) in the linear regression form $\hat{y}(t) = h(t)^T\theta$ and find the least-squares estimate for a, b, given the available data.

(b) (2 points) A DC-motor is described by the state equation

$$\dot{\mathbf{x}} = \begin{bmatrix} -R/L & -K_e/L \\ K_m/J_m & -b/J_m \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1/L \\ 0 \end{bmatrix} u \tag{20a}$$
$$y = \begin{bmatrix} 0 & 1 \end{bmatrix}\mathbf{x} \tag{20b}$$

with $x_1 = i$ (current) and $x_2 = \omega$ (angular speed). The model has 5 parameters R, L, J_m, K_m, and K_e. How many of these parameters can be identified from experimental input-output data u, ω? Motivate your answer!

(c) (6 points) Consider the model structure

$$y(t) + \alpha y(t-1) = \beta u(t-1) + e(t) + \gamma e(t-1), \tag{21}$$

where $e(\cdot)$ is a sequence of i.i.d. random variables with zero mean.

1. Compute the one-step ahead predictor for the model (21).
2. How can the Newton method be used to minimize the quadratic criterion

$$V_N(\theta) = \frac{1}{N}\sum_{t=1}^{N}(y(t) - \hat{y}(t|t-1))^2 \tag{22}$$

with $\theta^T = \begin{bmatrix} \alpha & \beta & \gamma \end{bmatrix}$? (Give a short and concise answer).
3. Give an equation that describes a Newton step for the solution above, assuming a Gauss-Newton approximation is used.

Solution:

(a) The linear regression form is

$$y(t) = \begin{bmatrix} y(t-1) & u(t-1) \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}, \tag{23}$$

and the least squares estimate of the parameters can be found by $\hat{\theta}_{LS} = (H^TH)^{-1}H^T\mathbf{y}$. Using the available data we have that

$$H = \begin{bmatrix} y(0) & u(0) \\ y(1) & u(1) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y(0) \\ y(1) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{24}$$

Hence the least squares estimate is $\hat{\theta}_{LS} = [a \quad b] = [1 \quad 0]$.

(b) The input-output relation can be expressed by the transfer function, which becomes

$$G(s) = \frac{\Omega(s)}{U(s)} = \frac{K_m}{(R+Ls)(b+J_ms) + K_eK_m} = \frac{K_m/LJ_m}{s^2 + (R/L + b/J_m)s + (Rb + K_eK_m)/LJ_m}. \tag{25}$$

Since the transfer function has only 3 parameters, it means that only 3 parameters can be determined from input-output data.

(c) Using the backward shift operator q^{-1} (alternatively, z^{-1} can be used), the model can equivalently be written as

$$(1 + \alpha q^{-1})y(t) = \beta q^{-1}u(t) + (1 + \gamma q^{-1})e(t). \tag{26}$$

1. The derivation is given for the general case in the Lecture notes, but in this particular case it can be done simpler. Rewrite the model as

$$(1 + \gamma q^{-1})y(t) = (\gamma - \alpha)q^{-1}y(t) + \beta q^{-1}u(t) + (1 + \gamma q^{-1})e(t), \tag{27}$$

or, equivalently,

$$y(t) = \frac{1}{1 + \gamma q^{-1}}\big[(\gamma - \alpha)q^{-1}y(t) + \beta q^{-1}u(t)\big] + e(t) = \hat{y}(t|t-1) + e(t), \tag{28}$$

where the latter equality follows from the fact that $e(t)$ is the only part of $y(t)$ that cannot be predicted. Hence, the predictor is given by the difference equation

$$(1 + \gamma q^{-1})\hat{y}(t|t-1) = (\gamma - \alpha)y(t-1) + \beta u(t-1) \tag{29}$$

2. The minimization is approached by searching for a solution to the equation

$$\nabla_\theta V_N(\theta) = 0 \tag{30}$$

using the Newton method (since $V_N(\theta)$ is nonlinear in θ).

3. Using the Gauss-Newton approximation, we obtain

$$\nabla_\theta V_N(\theta) = -\frac{2}{N}\sum_{t=1}^{N} \nabla_\theta \hat{y}(t|t-1,\theta)\big(y(t) - \hat{y}(t|t-1,\theta)\big)$$

$$\nabla^2_\theta V_N(\theta) \approx \frac{2}{N}\sum_{t=1}^{N} \nabla_\theta \hat{y}(t|t-1,\theta)\,\nabla_\theta \hat{y}(t|t-1,\theta)^T$$

The Newton step is obtained by solving the following system of linear equations for $\Delta\theta$:

$$\Big[\frac{1}{N}\sum_{t=1}^{N} \nabla_\theta \hat{y}(t|t-1,\theta)\,\nabla_\theta \hat{y}(t|t-1,\theta)^T\Big]\Delta\theta = \frac{1}{N}\sum_{t=1}^{N} \nabla_\theta \hat{y}(t|t-1,\theta)\big(y(t) - \hat{y}(t|t-1,\theta)\big)$$

4. **Simulation**

(a) (4 points) Write a pseudo-code (algorithm) that would deploy an IRK scheme for an implicit DAE

$$\mathbf{F}\left(\dot{\mathbf{x}},\, \mathbf{x},\, \mathbf{z},\, \mathbf{u}\right) = 0 \tag{31}$$

Be specific enough that someone could code it without knowing what the algorithm is about.

(b) (3 points) Verify that the (non-autonomous) scalar ODE

$$\dot{x}(t) = f(x(t), t), \quad x(0) = x_0 \tag{32}$$

can be transformed into the (autonomous) ODE

$$\dot{\mathbf{z}}(t) = \mathbf{f}(\mathbf{z}(t)) \tag{33}$$

using $\mathbf{z}^T(t) = \begin{bmatrix} x(t) & t \end{bmatrix}$. By comparing the RK equations in the two cases, show that the condition $c_i = \sum_{j=1}^{s} a_{ij}$ arises naturally, if we want the RK method to give the same numerical approximation for the two versions of the same problem.

(c) (3 points) Consider an integration scheme, described by the following Butcher array:

0	0	0
1	1/2	1/2
	1/2	1/2

1. Determine the stability function.
2. Is the scheme A-stable?

Solution:

(a) The pseudo-code will look like

Algorithm: Integration of implicit ODE

Input: $\mathbf{x}_0$, $\mathbf{u}(t_0), \ldots, \mathbf{u}(.)$, α and Δt
Set $\mathbf{K}, \mathbf{z} = 0$ (or any better initial guess)
for $k = 0 : N-1$ **do**
  while $\|\mathbf{r}\left(\mathbf{K}, \mathbf{z}, \mathbf{x}_k, \mathbf{u}(.)\right)\| > \text{tol}$ **do**
    Evaluate:

$$\mathbf{r}\left(\mathbf{K}, \mathbf{x}_k, \mathbf{u}(.)\right) = \begin{bmatrix} \mathbf{F}\left(\mathbf{K}_1, \mathbf{z}_1, \mathbf{x}_k + \Delta t \sum_{i=1}^{s} a_{1i}\mathbf{K}_i,\, \mathbf{u}(t_k + c_1 \Delta t)\right) \\ \vdots \\ \mathbf{F}\left(\mathbf{K}_s, \mathbf{z}_s, \mathbf{x}_k + \Delta t \sum_{i=1}^{s} a_{si}\mathbf{K}_i,\, \mathbf{u}(t_k + c_s \Delta t)\right) \end{bmatrix} = 0$$

    and

$$\frac{\partial \mathbf{r}\left(\mathbf{K}, \mathbf{z}, \mathbf{x}_k, \mathbf{u}(.)\right)}{\partial \mathbf{w}}$$

    where $\mathbf{w}$ gathers $\mathbf{K}_{1,\ldots,s}$ and $\mathbf{z}_{1,\ldots,s}$.
    Take the Newton step

$$\mathbf{w} \leftarrow \mathbf{w} - \alpha \frac{\partial \mathbf{r}\left(\mathbf{K}, \mathbf{z}, \mathbf{x}_k, \mathbf{u}(.)\right)}{\partial \mathbf{w}}^{-1} \mathbf{r} \tag{34}$$

  Take the integrator step:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta t \sum_{i=1}^{s} b_i \mathbf{K}_i \tag{35}$$

return $\mathbf{x}_{0,\ldots N}$

(b) First note that

$$\dot{\mathbf{z}}(t) = \begin{bmatrix} \dot{x}(t) \\ 1 \end{bmatrix} = \begin{bmatrix} f(z_1(t), z_2(t)) \\ 1 \end{bmatrix} = \mathbf{f}(\mathbf{z}(t)) \tag{36}$$

The RK equations for the first problem version are

$$.K_i = f\big(x_k + \Delta t \cdot \sum_{j=1}^{s} a_{ij} K_j, t_k + c_i \Delta t\big), \quad i = 1, \ldots, s \tag{37}$$

where K_i are scalar. For the second version (which is autonomous), the RK equations for the, now vector $\mathbf{K}_i = (K_i^1, K_i^2)$, become

$$\mathbf{K}_i = \mathbf{f}(\mathbf{z}_k + \Delta t \cdot \sum_{j=1}^{s} a_{ij} \mathbf{K}_j) = \begin{bmatrix} f\big(x_k + \Delta t \cdot \sum_{j=1}^{s} a_{ij} K_j^1, t_k + \Delta t \cdot \sum_{j=1}^{s} a_{ij} K_j^2\big) \\ 1 \end{bmatrix} \tag{38}$$

Since trivially $K_i^2 = 1$, the equation can be simplified for $K_i = K_i^1$:

$$K_i = f\big(x_k + \Delta t \cdot \sum_{j=1}^{s} a_{ij} K_j, t_k + \Delta t \cdot \sum_{j=1}^{s} a_{ij}\big), \quad i = 1, \ldots, s \tag{39}$$

Comparing (37) with (39) gives the desired result.

(c) 1. Denoting the Butcher array as

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array}$$

the stability function is given by $R(\mu) = 1 + \mu b^T (I - \mu A)^{-1} \mathbf{1}$, where $\mu = \lambda \Delta t$ and $\mathbf{1}$ is a column vector with all entries equal to 1. Thus:

$$R(\mu) = 1 + \mu \begin{bmatrix} 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\mu/2 & 1 - \mu/2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1 + \mu/2}{1 - \mu/2}$$

2. Since $|1 + \mu/2| \leq |1 - \mu/2|$ for all μ in the left half-plane, $|R(\mu)| \leq 1$ for the same μ, i.e. the scheme is A-stable.

5. **Newton** The Newton method aims at solving a set of equations $\mathbf{r}(\mathbf{x}) = 0$ numerically. To that end, it iterates the recursion:

$$\frac{\partial \mathbf{r}(\mathbf{x})}{\partial \mathbf{x}} \Delta \mathbf{x} + \mathbf{r}(\mathbf{x}) = 0 \tag{40a}$$

$$\mathbf{x} \leftarrow \mathbf{x} + \alpha \Delta \mathbf{x} \tag{40b}$$

where $\alpha \in]0, 1]$ is the step-size.

(a) (3 points) The optimization problem

$$\text{minimize}_{\mathbf{x}}\, \Phi(\mathbf{x}, \mathbf{p}), \tag{41}$$

where $\mathbf{x}$ is the optimization variable and $\mathbf{p}$ is a vector of parameters, can be approached by applying the Newton method.

1. How is the function $\mathbf{r}(\mathbf{x})$ chosen in this case?
2. Give a sufficient condition for (41) to have a unique solution $\mathbf{x}^\star(\mathbf{p})$ for a given $\mathbf{p}$.

3. The sensitivity of the solution $\mathbf{x}^\star(\mathbf{p})$ with respect to $\mathbf{p}$ is given by the Jacobian:

$$\frac{\partial \mathbf{x}^\star(\mathbf{p})}{\partial \mathbf{p}} \tag{42}$$

Give an expression for this Jacobian, and state what is required for it to exist.

(b) (1 point) Optimization problems of the type (41) are common in system identification, often with a least-squares criterion. A modification of the Newton method (40) is then often used. Which is the modification and why is it used?

Solution:

(a) 1. The Newton method is applied to the (necessary) condition for a solution:

$$\mathbf{r}(\mathbf{x}) = \nabla_{\mathbf{x}} \Phi(\mathbf{x}, \mathbf{p}) = 0 \tag{43}$$

2. Problem (41) is guaranteed to have a unique solution if it is convex, i.e. if the Hessian $\nabla_{\mathbf{xx}} \Phi(\mathbf{x}, \mathbf{p})$ is positive definite.
3. The solution $\mathbf{x}^\star(\mathbf{p})$ satisfies the implicit equation (43), i.e.

$$\nabla_{\mathbf{x}} \Phi(\mathbf{x}^\star(\mathbf{p}), \mathbf{p}) = 0 \tag{44}$$

We can then use the Implicit Function Theorem and observe that:

$$\left(\frac{\partial}{\partial \mathbf{x}} \nabla_{\mathbf{x}} \Phi(\mathbf{x}, \mathbf{p}) \right) \Big|_{\mathbf{x}=\mathbf{x}^\star(\mathbf{p})} \frac{\partial \mathbf{x}^\star(\mathbf{p})}{\partial \mathbf{p}} + \frac{\partial}{\partial \mathbf{p}} \nabla_{\mathbf{x}} \Phi(\mathbf{x}, \mathbf{p}) \Big|_{\mathbf{x}=\mathbf{x}^\star(\mathbf{p})} = 0 \tag{45}$$

or more simply:

$$\frac{\partial \mathbf{x}^\star(\mathbf{p})}{\partial \mathbf{p}} = -\nabla_{\mathbf{xx}} \Phi(\mathbf{x}, \mathbf{p})^{-1} \nabla_{\mathbf{xp}} \Phi(\mathbf{x}, \mathbf{p}) \Big|_{\mathbf{x}=\mathbf{x}^\star(\mathbf{p})} \tag{46}$$

This calculation is only possible if the square matrix $\nabla_{\mathbf{xx}} \Phi(\mathbf{x}, \mathbf{p})$ is full rank (which is the case if the problem is convex).

(b) For least-squares problems, the derivative of $\mathbf{r}(\mathbf{x})$, i.e. the Hessian of Φ, is often simplified and replaced by a Gauss-Newton approximation, which is simpler to compute and which automatically ensures positive semidefiniteness of (the approximation of) the Hessian.

THE END

Appendix: some possibly useful formula

- Lagrange mechanics is built on the equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} - \frac{\partial \mathcal{L}}{\partial \mathbf{q}} = \mathbf{Q}, \qquad \mathcal{L}\left(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{z}\right) = \mathcal{T} - \mathcal{V} - \mathbf{z}^{\top}\mathbf{C}, \qquad \mathbf{C} = 0, \qquad \langle \delta\mathbf{q},\, \mathbf{Q}\rangle = \delta W,\, \forall\, \delta\mathbf{q} \tag{47}$$

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

$$\mathcal{T} = \frac{1}{2} m \dot{\mathbf{p}}^{\top}\dot{\mathbf{p}}, \qquad \mathcal{V} = mg\mathbf{p}_3 \tag{48}$$

  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:

$$\dot{\mathbf{p}} = \mathbf{v} \tag{49a}$$

$$W\dot{\mathbf{v}} + \frac{\partial \mathbf{C}}{\partial \mathbf{q}}^{\top}\mathbf{z} = \mathbf{Q} - \frac{\partial \mathcal{V}}{\partial \mathbf{q}}^{\top} \tag{49b}$$

$$0 = \mathbf{C}\left(\mathbf{q}\right) \tag{49c}$$

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

$$\mathbf{r}\left(\mathbf{y}, \mathbf{z}\right) = 0 \tag{50}$$

  "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

$$\mathbf{r}\left(\mathbf{y}, \phi\left(\mathbf{y}\right)\right) = 0 \tag{51}$$

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

$$\frac{\partial \mathbf{z}}{\partial \mathbf{y}} = -\frac{\partial \mathbf{r}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{r}}{\partial \mathbf{y}} \tag{52}$$

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

$$\mathbf{x} \leftarrow \mathbf{x} - \alpha\frac{\partial \mathbf{r}}{\partial \mathbf{x}}^{-1}\mathbf{r} \tag{53}$$

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

- Runge-Kutta methods are described by:

$$\begin{array}{c|ccc} c_1 & a_{11} & \dots & a_{1s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s1} & \dots & a_{ss} \\ \hline & b_1 & \dots & b_s \end{array} \qquad \mathbf{K}_j = \mathbf{f}\left(\mathbf{x}_k + \Delta t \sum_{i=1}^{s} a_{ji}\mathbf{K}_i,\, \mathbf{u}\left(t_k + c_j\Delta t\right)\right), \quad j = 1, \dots, s \tag{54a}$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta t \sum_{i=1}^{s} b_i \mathbf{K}_i \tag{54b}$$

- The stability function for RK methods is given by

$$R(\mu) = 1 + \mu b^T (I - \mu A)^{-1}\mathbf{1} = \frac{\det(I - \mu A + \mu \mathbf{1} b^T)}{\det(I - \mu A)}, \quad \mu = \lambda \Delta t \tag{55}$$

- 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:

$$\dot{\mathbf{x}}(t_k + \Delta t \cdot \tau) \approx \dot{\hat{\mathbf{x}}}(t_k + \Delta t \cdot \tau) = \sum_{i=1}^{s} \mathbf{K}_i \ell_i(\tau), \quad \tau \in [0, 1] \tag{56}$$

$$\mathbf{x}(t_k + \Delta t \cdot \tau) \approx \hat{\mathbf{x}}(t_k + \Delta t \cdot \tau) = \mathbf{x}_k + \Delta t \sum_{i=1}^{s} \mathbf{K}_i L_i(\tau) \tag{57}$$

where the Lagrange polynomials are given by:

$$\ell_i(\tau) = \prod_{j=1, j\neq i}^{s} \frac{\tau - \tau_j}{\tau_i - \tau_j}, \quad \text{and} \quad L_i(\tau) = \int_0^{\tau} \ell_i(\xi) \mathrm{d}\xi \tag{58}$$

The Lagrange polynomials satisfy the condition of

$$\text{Punctuality:} \quad \ell_i(\tau_j) = \begin{cases} 1 & \text{if} \quad j = i \\ 0 & \text{if} \quad j \neq i \end{cases} \tag{59}$$

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

$$\dot{\hat{\mathbf{x}}}(t_k + \Delta t \cdot \tau_j) = \mathbf{f}\left(\hat{\mathbf{x}}(t_k + \Delta t \cdot \tau_j),\, \mathbf{u}\left(t_k + \Delta t \cdot \tau_j\right)\right), \qquad \text{in the explicit ODE case} \tag{60a}$$

$$\mathbf{F}\left(\dot{\hat{\mathbf{x}}}(t_k + \Delta t \cdot \tau_j), \hat{\mathbf{x}}(t_k + \Delta t \cdot \tau_j),\, \mathbf{u}\left(t_k + \Delta t \cdot \tau_j\right)\right) = 0, \qquad \text{in the implicit ODE case} \tag{60b}$$

$$\mathbf{F}\left(\dot{\hat{\mathbf{x}}}(t_k + \Delta t \cdot \tau_j), \hat{\mathbf{z}}_j, \hat{\mathbf{x}}(t_k + \Delta t \cdot \tau_j),\, \mathbf{u}\left(t_k + \Delta t \cdot \tau_j\right)\right) = 0, \qquad \text{in the fully-implicit DAE case} \tag{60c}$$

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

$$P_s(\tau) = \frac{1}{s!} \frac{\mathrm{d}^s}{\mathrm{d}\tau^s} \left[\left(\tau^2 - \tau\right)^s\right], \tag{61}$$

implying that the Lagrange polynomials also satisfy the condition of

$$\text{Orthogonality:} \quad \int_0^1 \ell_i(\tau)\ell_j(\tau)\, \mathrm{d}\tau = 0 \quad \text{for} \quad i \neq j \tag{62}$$

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

$$\max_{\boldsymbol{\theta}} \quad \mathbb{P}\left[e_k = y_k - \hat{y}_k \quad \text{for} \quad k = 1, \ldots, N \mid \boldsymbol{\theta}\right] \tag{63}$$

If the noise sequence is uncorrelated, then

$$\mathbb{P}\left[e_k = y_k - \hat{y}_k \quad \text{for} \quad k = 0, \ldots, N \mid \boldsymbol{\theta}\right] = \prod_{k=1}^{N} \mathbb{P}\left[e_k = y_k - \hat{y}_k \mid \boldsymbol{\theta}\right] \tag{64}$$

- The solution of a linear least-squares problem

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{2} \left\| A\boldsymbol{\theta} - \mathbf{y} \right\|_{\Sigma_e^{-1}}^2 \tag{65}$$

reads as:

$$\hat{\boldsymbol{\theta}} = \left(A^\top \Sigma_e^{-1} A\right)^{-1} A^\top \Sigma_e^{-1} \mathbf{y} \tag{66}$$

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

$$\Sigma_{\hat{\boldsymbol{\theta}}} = \left(A^\top \Sigma_e^{-1} A\right)^{-1} \tag{67}$$

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

$$H(z)\hat{y}(t|t-1) = G(z)u(t) + (H(z) - 1)y(t) \tag{68}$$

- The Gauss-Newton approximation in an optimization problem

$$\min_{\mathbf{x}} \quad J\left(\mathbf{x}\right) = \frac{1}{2} \left\| \mathbf{R}\left(\mathbf{x}\right) \right\|^2 \tag{69}$$

uses the approximation:

$$\frac{\partial^2 J}{\partial \mathbf{x}^2} \approx \frac{\partial R}{\partial \mathbf{x}}^\top \frac{\partial R}{\partial \mathbf{x}} \tag{70}$$

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

$$\mathbf{x}(t) = e^{At}\mathbf{x}(0) + \int_0^t e^{A(t-\tau)} B\mathbf{u}(\tau) \mathrm{d}\tau \tag{71}$$

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

$$G(s) = C\left(sI - A\right)^{-1} B + D \tag{72}$$