Modelling and Simulation ESS101
5 January 2022 (Re-exam)

This exam contains 11 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, but cannot contain solutions to the exercises or previous exams)
- *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	10	
2	11	
3	9	
4	6	
5	4	
Total:	40	

Best of luck to all !!

Examiner: Yasemin Bekiroglu, +46 70 148 72 71
Teaching Assistant: Ahmet Ercan Tekden, +46 72 371 03 10
Lecturer: Hadi Hajieghrary, +46 72 177 90 46

1. (a) (10 points) We will consider the ring problem illustrated in Fig. 1. A massless ring-shaped rail is rotating around its vertical axis, subject to a torque T. A mass m slides along the ring without friction. Choose a set of coordinates to describe the system, and write the model equations of the system (using the Lagrange function and assembling the Lagrange equations).
Hint: let us consider a point $\mathbf{p} \in \mathbb{R}^3$ *on the ring. The work* δW *produced by* T *for a displacement* $\delta\mathbf{p}$ *is then given by:*

$$\delta W = \frac{1}{R^2}\det\left(\begin{bmatrix} \mathbf{T} & \mathbf{p} & \delta\mathbf{p} \end{bmatrix}\right) \tag{1}$$

where $\mathbf{T} = \begin{bmatrix} 0 & 0 & T \end{bmatrix}^\top$ *(also det denotes determinant). You may want to use this observation or not, depending on how you decide to approach the problem.*

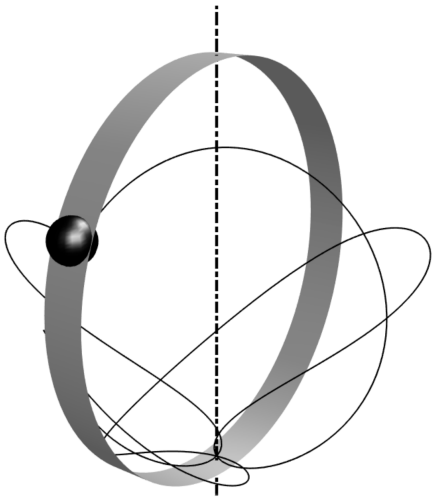


Figure 1: Illustration of the ring problem. The ring has a negligible inertia, a radius R, (the ball to be at distance R from the origin), and is subject to a torque T on its vertical rotation axis (dashed-dotted line). The mass (black ball) slides without friction on the ring.

Solution:

(a) Let us pick

$$\mathbf{q} = \mathbf{p} \tag{2}$$

as our generalized coordinates, where $\mathbf{p} \in \mathbb{R}^3$ is the position of the mass in the cartesian reference frame attached to the axis of the ring (vertical axis aligned with the vertical ring axis, origin at the center of the ring). The kinetic and potential energies read as:

$$T = \frac{1}{2}m\dot{\mathbf{p}}^\top\dot{\mathbf{p}}, \qquad V = mg\mathbf{p}_3 \tag{3}$$

and we use the constraints:

$$\mathbf{c} = \frac{1}{2}\left(\mathbf{p}^\top\mathbf{p} - R^2\right) = 0 \tag{4}$$

that imposes the ball to be at distance R from the origin. We then write the Lagrange function:

$$\mathcal{L} = \frac{1}{2} m \dot{\mathbf{p}}^\top \dot{\mathbf{p}} - mg\mathbf{p}_3 - z\frac{1}{2}\left(\mathbf{p}^\top \mathbf{p} - R^2\right) \tag{5}$$

where $z \in \mathbb{R}$. We can then trivially compute:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}}^\top = -\begin{bmatrix} 0 \\ 0 \\ mg \end{bmatrix} - z\mathbf{p}, \qquad \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}^\top = m\dot{\mathbf{p}}, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}^\top = m\ddot{\mathbf{p}} \tag{6}$$

and assemble the Lagrange equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}^\top - \frac{\partial \mathcal{L}}{\partial \mathbf{q}}^\top = m\ddot{\mathbf{p}} + \begin{bmatrix} 0 \\ 0 \\ mg \end{bmatrix} + z\mathbf{p} = \mathbf{Q} \tag{7a}$$

$$\frac{1}{2}\left(\mathbf{p}^\top \mathbf{p} - R^2\right) = 0 \tag{7b}$$

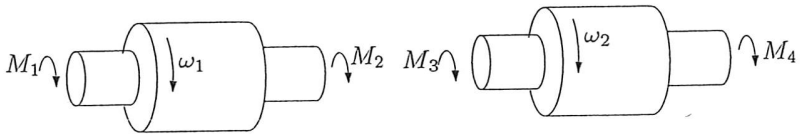
We finally compute the generalized forces $\mathbf{Q}$. Using (1), we have:

$$\delta W = \frac{T}{R^2}\left(\delta \mathbf{p}_2 \mathbf{p}_1 - \delta \mathbf{p}_1 \mathbf{p}_2\right) \tag{8}$$

such that

$$\mathbf{Q} = \frac{T}{R^2}\begin{bmatrix} -\mathbf{p}_2 \\ \mathbf{p}_1 \\ 0 \end{bmatrix} \tag{9}$$

2. (a) (5 points) Consider the following rotating parts with moments of inertia J_1 and J_2.



The two parts are joined, resulting in the equations

$$J_1\dot{\omega}_1 = M_1 + M_2 \tag{10a}$$
$$J_2\dot{\omega}_2 = M_3 + M_4 \tag{10b}$$
$$M_2 = -M_3 \tag{10c}$$
$$\omega_1 = \omega_2, \tag{10d}$$

where M_1 and M_4 are considered to be inputs.
What is the index of this system?

(b) (6 points) Consider the following DAE, where u is the input:

$$\dot{x}_1 = x_2 \tag{11a}$$
$$\dot{x}_2 = -z_1 + u \tag{11b}$$
$$x_1 = z_1 + z_2 \tag{11c}$$
$$z_1 = z_2 \cdot |z_2| \tag{11d}$$

1. What is the differential index of (11)?

2. Transform (11) into an ODE (ordinary differential equation).

Solution:

(a) From the two first equations of (10), $\dot{\omega}_1$ and $\dot{\omega}_2$ can be solved for. Differentiating the third and fourth equations give

$$\dot{M}_2 = -\dot{M}_3 \tag{12}$$
$$\dot{\omega}_1 = \dot{\omega}_2 \tag{13}$$

and inserting the expressions for $\dot{\omega}_1$ and $\dot{\omega}_2$ results in

$$\dot{M}_2 = -\dot{M}_3 \tag{14}$$
$$(M_1 + M_2)/J_1 = (M_3 + M_4)/J_2 \tag{15}$$

Differentiation of the latter of these two equations then gives

$$\dot{M}_2 = -\dot{M}_3 \tag{16}$$
$$(\dot{M}_1 + \dot{M}_2)/J_1 = (\dot{M}_3 + \dot{M}_4)/J_2, \tag{17}$$

from which we can solve for $\dot{M}_1$ and $\dot{M}_2$ (remember that M_1 and M_4 are inputs). Hence, the system has index 2.

(b) 1. We observe that (11) is a semi-explicit DAE with

$$\mathbf{g}(\mathbf{x}, \mathbf{z}, u) = \begin{bmatrix} x_1 - (z_1 + z_2) \\ z_1 - z_2 \cdot |z_2| \end{bmatrix}. \tag{18}$$

By noting that the derivative of $z_2 \cdot |z_2|$ can be computed by considering the function for positive and negative z_2 separately, we have

$$\frac{\partial \mathbf{g}}{\partial \mathbf{z}} = \begin{bmatrix} -1 & -1 \\ 1 & -2|z_2| \end{bmatrix}. \tag{19}$$

Since the determinant is nonzero for all z_2, the differential index is equal to 1.

2. Since the index is 1, it should be sufficient to differentiate the algebraic equations once to obtain an ODE. Indeed, we get

$$\dot{x}_1 = \dot{z}_1 + \dot{z}_2 \tag{20}$$
$$\dot{z}_1 = 2|z_2|\dot{z}_2, \tag{21}$$

which can be solved for $\dot{z}_1$ and $\dot{z}_2$:

$$\dot{z}_1 = \frac{2|z_2|x_2}{1 + 2|z_2|} \tag{22}$$
$$\dot{z}_2 = \frac{x_2}{1 + 2|z_2|} \tag{23}$$

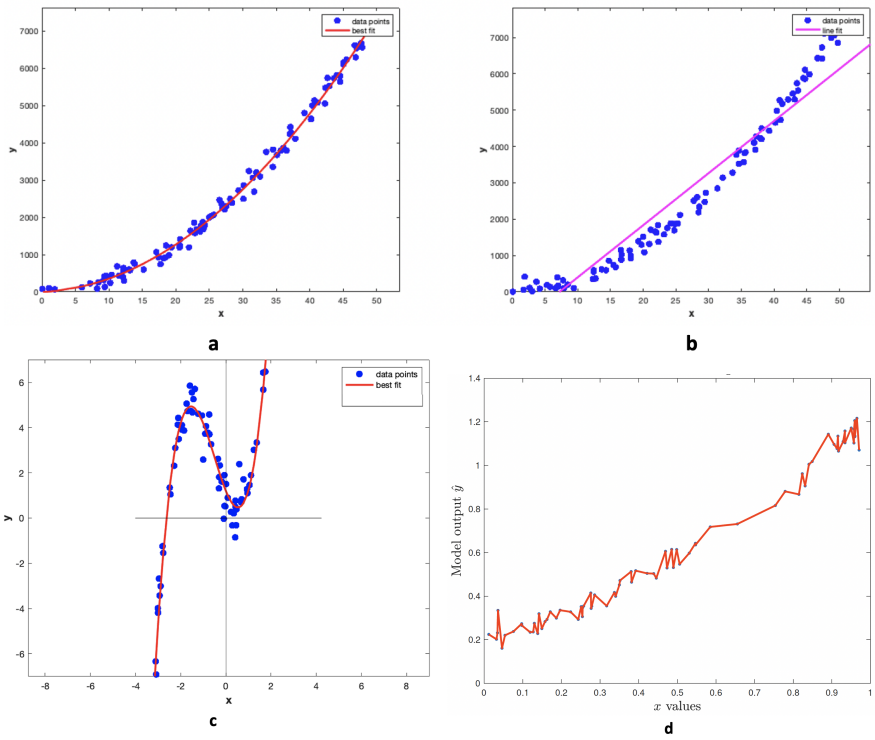


Figure 2: Linear regression examples based on least squares.

The complete ODE is thus

$$\dot{x}_1 = x_2 \tag{24}$$
$$\dot{x}_2 = -z_1 + u \tag{25}$$
$$\dot{z}_1 = \frac{2|z_2|x_2}{1+2|z_2|} \tag{26}$$
$$\dot{z}_2 = \frac{x_2}{1+2|z_2|} \tag{27}$$

3. (a) (4 points) Consider the curve fitting results in Figure 2a-d. Comment on which curve fitting is better when Figure 2a and Figure 2b is compared. How is better fitting determined? What is different in the curve fitting approach in Figure 2c, what mainly changes in the fitted curve models from Figure 2a to Figure 2c? Comment on the quality of the model fitting in Figure 2d. Is it an acceptable fitting result? What could be potential problems and how they can be resolved?

   (b) (2 points) Consider the ARX model:

$$y_k + a_1 y_{k-1} + a_2 y_{k-2} = b_0 u_k + e_k \tag{28}$$

   where e_k is the additive noise and the associated data $y_{0,\ldots,N}$ and $u_{0,\ldots,N}$ obtained from applying the input sequence $u_{0,\ldots,N}$ to the real system, started with $y_{k<0} = 0$. Write out the one-step ahead predictor $\hat{y}_k$ in the form of $\hat{y}_k = \boldsymbol{\theta}^T \boldsymbol{\varphi}$, where $\boldsymbol{\theta}$ is the parameter vector and $\boldsymbol{\varphi}$ is the regression vector and define the entries in $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$.

   (c) (3 points) Consider a two-dimensional dataset $\{x(i), y(i)\}, i = 1, \ldots N$. Using a least squares ap-

proach capturing the relationship between x and y,

$$\hat{\mathbf{y}} = a + bx + cx^2 + dx^3 + fx^4 = \boldsymbol{\theta}^\top \boldsymbol{\varphi},$$

where $\boldsymbol{\theta}$ is the parameter vector and $\boldsymbol{\varphi}$ is the *regression vector* (holding the *regressors* $1, x, ...$), write out all the components in the R_N and f_N matrices in the least squares estimate defined as

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

Solution:

(a) Model fit in Figure 2a is better, in comparison to Figure 2b, as the fitting error is less. Model fit quality can be determined by a loss function based on the differences between the predictions and the ground truth values. Model in Figure 2c has a higher order, it is a more complex model. Figure 2d shows overfitting which is problematic as it leads to bad generalization performance, i.e. model predictions on unseen data would not be of good quality. Cross-validation can help to find a good model order to avoid overfitting.

(b) The one-step ahead predictor reads as:

$$\hat{y}_k = -a_1 y_{k-1} - a_2 y_{k-2} + b_0 u_k \tag{29}$$

and the mismatch between the data and the predictor is given by:

$$e_k = \hat{y}_k - y_k = -y_k - a_1 y_{k-1} - a_2 y_{k-2} + b_0 u_k = \boldsymbol{\theta}^T \boldsymbol{\varphi} - y_k \tag{30}$$

where $\boldsymbol{\varphi} = \begin{bmatrix} -y_{k-1} \\ -y_{k-2} \\ u_k \end{bmatrix}$, $\boldsymbol{\theta} = \begin{bmatrix} a_1 \\ a_2 \\ b_0 \end{bmatrix}$.

(c) The matrices are written as below where $k = 4$

$$\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\varphi}(i)\boldsymbol{\varphi}^\top(i) = \frac{1}{N}\begin{bmatrix} N & \sum_{i=1}^{N} x(i) & ... & \sum_{i=1}^{N} x^k(i) \\ \sum_{i=1}^{N} x(i) & \sum_{i=1}^{N} x^2(i) & ... & \sum_{i=1}^{N} x^{k+1}(i) \\ . & . & . & \\ . & . & . & \\ . & . & . & \\ \sum_{i=1}^{N} x^k(i) & \sum_{i=1}^{N} x^{k+1}(i) & ... & \sum_{i=1}^{N} x^{2*k}(i) \end{bmatrix} \tag{31}$$

$$\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\varphi}(i)y(i) = \frac{1}{N}\begin{bmatrix} \sum_{i=1}^{N} y(i) \\ \sum_{i=1}^{N} x(i)y(i) \\ . \\ . \\ . \\ \sum_{i=1}^{N} x^k(i)y(i) \end{bmatrix} \tag{32}$$

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

$$M\Delta\mathbf{x} + \mathbf{r}(\mathbf{x}) = 0 \tag{33a}$$

$$\mathbf{x} \leftarrow \mathbf{x} + \alpha\Delta\mathbf{x} \tag{33b}$$

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

(a) (1 point) How should matrix M be ideally chosen?

(b) (1 point) Explain in words what condition(s) is (are) required for Newton to converge with $\alpha = 1$.

(c) (2 points) Why do we need α and how should it be chosen?

(d) (2 points) The local convergence rate of an exact, full-step Newton method can be summarized as:

$$\|\mathbf{x}_+ - \mathbf{x}_\star\| \leq c \|\mathbf{x} - \mathbf{x}_\star\|^2 \tag{34}$$

where $\mathbf{x}_\star$ is a solution of $\mathbf{r}(\mathbf{x}_\star)$. What is the meaning of this formula? When does it (doesn't it) occur?

Solution:

(a) Ideally, we ought to chose M as the Jacobian matrix $\frac{\partial \mathbf{r}}{\partial \mathbf{x}}$. In practice, approximations are often used.

(b) Full Newton steps are guaranteed to converge in a neighborhood of a solution only. The "size" of that neighborhood depends on how nonlinear $\mathbf{r}(\mathbf{x})$ is, and the Jacobian $\frac{\partial \mathbf{r}(\mathbf{x})}{\partial \mathbf{x}}$ must be full rank throughout this neighborhood.

(c) The step-size α allows the Newton iteration to converge even if $\mathbf{x}$ is not close to the solution $\mathbf{x}_\star$. It is typically chosen so as to ensure that

$$\|\mathbf{r}(\mathbf{x} + \alpha \Delta \mathbf{x})\| < \|\mathbf{r}(\mathbf{x})\| \tag{35}$$

Finding α is the role of a (small) computer code usually labelled "line-search".

(d) This formula states that the exact, full-step Newton iteration converges quadratically to a solution. That is, the number of accurate digits in the $\mathbf{x}$ is doubled at every iteration. Achieving the quadratic contraction rate requires basically what is stated in the question, namely:

- Exact Newton steps, i.e. an exact Jacobian $M = \frac{\partial \mathbf{r}(\mathbf{x})}{\partial \mathbf{x}}$ is used and system (33a) is solved to machine precision.
- Full steps are taken, i.e. $\alpha = 1$ throughout the iterations.
- The quadratic convergence rate is local, i.e. it occurs in a neighborhood of the solution $\mathbf{x}_\star$.

5. (a) (2 points) What is the maximum order (for a given number of stages s) that an IRK method can achieve? What one needs to do to achieve that order?

Consider the following Runge-Kutta equations for integration of an ODE $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u})$:

$$\begin{aligned} \mathbf{K}_1 &= \mathbf{f}\left(\mathbf{x}_k, \mathbf{u}(t_k)\right) \\ \mathbf{K}_2 &= \mathbf{f}\big(\mathbf{x}_k + \Delta t \cdot \mathbf{K}_1, \mathbf{u}(t_k + \Delta t)\big) \\ \mathbf{x}_{k+1} &= \mathbf{x}_k + \frac{\Delta t}{2}(\mathbf{K}_1 + \mathbf{K}_2) \end{aligned}$$

(b) (1 point) Is the RK scheme explicit or implicit? How many stages are there?

(c) (1 point) What is the Butcher array describing the scheme?

Solution:

(a) The family of IRK methods includes the Gauss-Legendre collocation methods (this is easy to verify from the equations provided in the appendix), which achieve an order up to $2s$. That is the maximum order that IRK methods can achieve for a given number of stages s. Gauss-Legendre collocation schemes yield a very specific Butcher tableau (a, b, c) to be used in the IRK scheme. The order $2s$ is achieved only if this specific Butcher tableau is used.

(b) The RK scheme is explicit and has 2 stages.

(c)

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

Appendix: some possibly useful formula

- Lagrange mechanics is built on the equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} - \frac{\partial \mathcal{L}}{\partial \mathbf{q}} = \mathbf{Q}, \qquad \mathcal{L}\left(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{z}\right) = \mathcal{T} - \mathcal{V} - \mathbf{z}^\top \mathbf{C}, \qquad \mathbf{C} = 0, \qquad \langle \delta\mathbf{q},\, \mathbf{Q}\rangle = \delta W,\, \forall\, \delta\mathbf{q} \tag{36}$$

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

$$\mathcal{T} = \frac{1}{2} m \dot{\mathbf{p}}^\top \dot{\mathbf{p}}, \qquad \mathcal{V} = mg\mathbf{p}_3 \tag{37}$$

  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:

$$\dot{\mathbf{p}} = \mathbf{v} \tag{38a}$$

$$W\dot{\mathbf{v}} + \frac{\partial \mathbf{C}}{\partial \mathbf{q}}^\top \mathbf{z} = \mathbf{Q} - \frac{\partial \mathcal{V}}{\partial \mathbf{q}}^\top \tag{38b}$$

$$0 = \mathbf{C}\left(\mathbf{q}\right) \tag{38c}$$

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

$$\mathbf{r}\left(\mathbf{y}, \mathbf{z}\right) = 0 \tag{39}$$

  "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

$$\mathbf{r}\left(\mathbf{y}, \phi\left(\mathbf{y}\right)\right) = 0 \tag{40}$$

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

$$\frac{\partial \mathbf{z}}{\partial \mathbf{y}} = -\frac{\partial \mathbf{r}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{r}}{\partial \mathbf{y}} \tag{41}$$

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

$$\mathbf{x} \leftarrow \mathbf{x} - \alpha \frac{\partial \mathbf{r}}{\partial \mathbf{x}}^{-1} \mathbf{r} \tag{42}$$

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

- Runge-Kutta methods are described by:

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

$$\mathbf{K}_j = \mathbf{f}\left(\mathbf{x}_k + \Delta t \sum_{i=1}^{s} a_{ji}\mathbf{K}_i,\, \mathbf{u}\left(t_k + c_j \Delta t\right)\right), \quad j = 1, \dots, s \tag{43a}$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta t \sum_{i=1}^{s} b_i \mathbf{K}_i \tag{43b}$$

- 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:

$$\dot{\mathbf{x}}(t_k + \Delta t \cdot \tau) \approx \dot{\hat{\mathbf{x}}}(t_k + \Delta t \cdot \tau) = \sum_{i=1}^{s} \mathbf{K}_i \ell_i(\tau), \quad \tau \in [0,\, 1] \tag{44}$$

$$\mathbf{x}(t_k + \Delta t \cdot \tau) \approx \hat{\mathbf{x}}(t_k + \Delta t \cdot \tau) = \mathbf{x}_k + \Delta t \sum_{i=1}^{s} \mathbf{K}_i L_i(\tau) \tag{45}$$

where the Lagrange polynomials are given by:

$$\ell_i(\tau) = \prod_{j=1, j\neq i}^{s} \frac{\tau - \tau_j}{\tau_i - \tau_j}, \quad \text{and} \quad L_i(\tau) = \int_0^{\tau} \ell_i(\xi) \mathrm{d}\xi \tag{46}$$

The Lagrange polynomials satisfy the conditions of

$$\text{Orthogonality:} \quad \int_0^1 \ell_i(\tau)\ell_j(\tau)\,\mathrm{d}\tau = 0 \quad \text{for} \quad i \neq j \tag{47a}$$

$$\text{Punctuality:} \quad \ell_i(\tau_j) = \begin{cases} 1 & \text{if} \quad j = i \\ 0 & \text{if} \quad j \neq i \end{cases} \tag{47b}$$

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

$$\dot{\hat{\mathbf{x}}}(t_k + \Delta t \cdot \tau_j) = \mathbf{f}\left(\hat{\mathbf{x}}(t_k + \Delta t \cdot \tau_j),\, \mathbf{u}\left(t_k + \Delta t \cdot \tau_j\right)\right), \qquad \text{in the explicit ODE case} \tag{48a}$$

$$\mathbf{F}\left(\dot{\hat{\mathbf{x}}}(t_k + \Delta t \cdot \tau_j), \hat{\mathbf{x}}(t_k + \Delta t \cdot \tau_j),\, \mathbf{u}\left(t_k + \Delta t \cdot \tau_j\right)\right) = 0, \qquad \text{in the implicit ODE case} \tag{48b}$$

$$\mathbf{F}\left(\dot{\hat{\mathbf{x}}}(t_k + \Delta t \cdot \tau_j), \hat{\mathbf{z}}_j, \hat{\mathbf{x}}(t_k + \Delta t \cdot \tau_j),\, \mathbf{u}\left(t_k + \Delta t \cdot \tau_j\right)\right) = 0, \qquad \text{in the fully-implicit DAE case} \tag{48c}$$

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

$$P_s\left(\tau\right) = \frac{1}{s!}\frac{\mathrm{d}^s}{\mathrm{d}\tau^s}\left[\left(\tau^2 - \tau\right)^s\right] \tag{49}$$

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

$$\max_{\boldsymbol{\theta}} \quad \mathbb{P}\left[e_k = y_k - \hat{y}_k \quad \text{for} \quad k = 1, \ldots, N \mid \boldsymbol{\theta}\right] \tag{50}$$

If the noise sequence is uncorrelated, then

$$\mathbb{P}\left[e_k = y_k - \hat{y}_k \quad \text{for} \quad k = 0, \ldots, N \mid \boldsymbol{\theta}\right] = \prod_{k=1}^{N} \mathbb{P}\left[e_k = y_k - \hat{y}_k \,|\, \boldsymbol{\theta}\,\right] \tag{51}$$

- The solution of a linear least-squares problem

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{2}\left\|A\boldsymbol{\theta} - \mathbf{y}\right\|^2_{\Sigma_e^{-1}} \tag{52}$$

reads as:

$$\hat{\boldsymbol{\theta}} = \left(A^\top \Sigma_e^{-1} A\right)^{-1} A^\top \Sigma_e^{-1} \mathbf{y} \tag{53}$$

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

$$\Sigma_{\hat{\boldsymbol{\theta}}} = \left(A^\top \Sigma_e^{-1} A\right)^{-1} \tag{54}$$

- 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

$$H(z)\hat{y}(z) = G(z)u(z) + (H(z) - 1)y(z) \tag{55}$$

- The Gauss-Newton approximation in an optimization problem

$$\min_{\mathbf{x}} \quad J(\mathbf{x}) = \frac{1}{2} \|\mathbf{R}(\mathbf{x})\|^2 \tag{56}$$

uses the approximation:

$$\frac{\partial^2 J}{\partial \mathbf{x}^2} \approx \frac{\partial R}{\partial \mathbf{x}}^\top \frac{\partial R}{\partial \mathbf{x}} \tag{57}$$

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

$$\mathbf{x}(t) = e^{At}\mathbf{x}(0) + \int_0^t e^{A(t-\tau)} B\mathbf{u}(\tau) \mathrm{d}\tau \tag{58}$$

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

$$G(s) = C\left(sI - A\right)^{-1} B + D \tag{59}$$

- $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $det(A) = ad - bc$
- $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$, $det(A) = a.det(\begin{bmatrix} e & f \\ h & i \end{bmatrix}) - b.det(\begin{bmatrix} d & f \\ g & i \end{bmatrix}) + c.det(\begin{bmatrix} d & e \\ g & h \end{bmatrix})$
- $\alpha = \mathbf{x}^T A \mathbf{x}$, 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, $\frac{\partial \alpha}{\partial \mathbf{x}} = 2\mathbf{x}^T A$.