Appendix: some possibly useful formula

- Lagrange mechanics is built on the equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} - \frac{\partial \mathcal{L}}{\partial \mathbf{q}} = \mathbf{Q}, \qquad \mathcal{L}\left(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{z}\right) = \mathcal{T} - \mathcal{V} - \mathbf{z}^\top \mathbf{C}, \qquad \mathbf{C} = 0, \qquad \langle \delta\mathbf{q}, \mathbf{Q}\rangle = \delta W, \,\forall\, \delta\mathbf{q} \tag{1}$$

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

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

  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 $\mathcal{V} = \mathcal{V}(\mathbf{q})$ and $\mathbf{C} = \mathbf{C}(\mathbf{q})$, the Lagrange equations simplify to the dynamics in the semi-explicit index-3 DAE form:

$$\dot{\mathbf{p}} = \mathbf{v} \tag{3a}$$

$$W\dot{\mathbf{v}} + \frac{\partial \mathbf{C}}{\partial \mathbf{q}}^\top \mathbf{z} = \mathbf{Q} - \frac{\partial \mathcal{V}}{\partial \mathbf{q}}^\top \tag{3b}$$

$$0 = \mathbf{C}(\mathbf{q}) \tag{3c}$$

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

$$\mathbf{r}(\mathbf{y}, \mathbf{z}) = 0 \tag{4}$$

  "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 $\phi(\mathbf{y})$ such that

$$\mathbf{r}(\mathbf{y}, \phi(\mathbf{y})) = 0 \tag{5}$$

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

$$\frac{\partial \mathbf{z}}{\partial \mathbf{y}} = -\frac{\partial \mathbf{r}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{r}}{\partial \mathbf{y}} \tag{6}$$

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

$$\mathbf{x} \leftarrow \mathbf{x} - \alpha \frac{\partial \mathbf{r}}{\partial \mathbf{x}}^{-1} \mathbf{r} \tag{7}$$

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

- Runge-Kutta methods are described by:

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

$$\mathbf{K}_j = \mathbf{f}\left(\mathbf{x}_k + \Delta t \sum_{i=1}^{s} a_{ji}\mathbf{K}_i,\, \mathbf{u}\left(t_k + c_j \Delta t\right)\right), \quad j = 1, \dots, s \tag{8a}$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta t \sum_{i=1}^{s} b_i \mathbf{K}_i \tag{8b}$$

- 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:

$$\dot{\mathbf{x}}(t_k+\Delta t\cdot\tau)\approx\dot{\hat{\mathbf{x}}}(t_k+\Delta t\cdot\tau)=\sum_{i=1}^{s}\mathbf{K}_i\ell_i(\tau),\quad \tau\in[0,\,1] \tag{9}$$

$$\mathbf{x}(t_k+\Delta t\cdot\tau)\approx\hat{\mathbf{x}}(t_k+\Delta t\cdot\tau)=\mathbf{x}_k+\Delta t\sum_{i=1}^{s}\mathbf{K}_iL_i(\tau) \tag{10}$$

where the Lagrange polynomials are given by:

$$\ell_i(\tau)=\prod_{j=1,j\neq i}^{s}\frac{\tau-\tau_j}{\tau_i-\tau_j},\quad\text{and}\quad L_i(\tau)=\int_0^{\tau}\ell_i(\xi)\mathrm{d}\xi \tag{11}$$

The Lagrange polynomials satisfy the conditions of

$$\text{Orthogonality:}\quad \int_0^1\ell_i(\tau)\ell_j(\tau)\,\mathrm{d}\tau=0\quad\text{for}\quad i\neq j \tag{12a}$$

$$\text{Punctuality:}\quad \ell_i(\tau_j)=\begin{cases}1 & \text{if}\quad j=i\\ 0 & \text{if}\quad j\neq i\end{cases} \tag{12b}$$

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

$$\dot{\hat{\mathbf{x}}}(t_k+\Delta t\cdot\tau_j)=\mathbf{f}\left(\hat{\mathbf{x}}(t_k+\Delta t\cdot\tau_j),\,\mathbf{u}\left(t_k+\Delta t\cdot\tau_j\right)\right),\qquad\text{in the explicit ODE case} \tag{13a}$$

$$\mathbf{F}\left(\dot{\hat{\mathbf{x}}}(t_k+\Delta t\cdot\tau_j),\hat{\mathbf{x}}(t_k+\Delta t\cdot\tau_j),\,\mathbf{u}\left(t_k+\Delta t\cdot\tau_j\right)\right)=0,\qquad\text{in the implicit ODE case} \tag{13b}$$

$$\mathbf{F}\left(\dot{\hat{\mathbf{x}}}(t_k+\Delta t\cdot\tau_j),\hat{\mathbf{z}}_j,\hat{\mathbf{x}}(t_k+\Delta t\cdot\tau_j),\,\mathbf{u}\left(t_k+\Delta t\cdot\tau_j\right)\right)=0,\qquad\text{in the fully-implicit DAE case} \tag{13c}$$

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

$$P_s\left(\tau\right)=\frac{1}{s!}\frac{\mathrm{d}^s}{\mathrm{d}\tau^s}\left[\left(\tau^2-\tau\right)^s\right] \tag{14}$$

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

$$\max_{\boldsymbol{\theta}}\quad\mathbb{P}\left[e_k=y_k-\hat{y}_k\quad\text{for}\quad k=1,\dots,N\mid\boldsymbol{\theta}\right] \tag{15}$$

If the noise sequence is uncorrelated, then

$$\mathbb{P}\left[e_k=y_k-\hat{y}_k\quad\text{for}\quad k=0,\dots,N\mid\boldsymbol{\theta}\right]=\prod_{k=1}^{N}\mathbb{P}\left[e_k=y_k-\hat{y}_k\,|\,\boldsymbol{\theta}\right] \tag{16}$$

- The solution of a linear least-squares problem

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{2} \|A\boldsymbol{\theta} - \mathbf{y}\|^2_{\Sigma_e^{-1}} \tag{17}$$

reads as:

$$\hat{\boldsymbol{\theta}} = \left(A^\top \Sigma_e^{-1} A\right)^{-1} A^\top \Sigma_e^{-1} \mathbf{y} \tag{18}$$

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

$$\Sigma_{\hat{\boldsymbol{\theta}}} = \left(A^\top \Sigma_e^{-1} A\right)^{-1} \tag{19}$$

- 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

$$H(z)\hat{y}(z) = G(z)u(z) + (H(z) - 1)y(z) \tag{20}$$

- The Gauss-Newton approximation in an optimization problem

$$\min_{\mathbf{x}} \quad J(\mathbf{x}) = \frac{1}{2} \|\mathbf{R}(\mathbf{x})\|^2 \tag{21}$$

uses the approximation:

$$\frac{\partial^2 J}{\partial \mathbf{x}^2} \approx \frac{\partial R}{\partial \mathbf{x}}^\top \frac{\partial R}{\partial \mathbf{x}} \tag{22}$$

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

$$\mathbf{x}(t) = e^{At}\mathbf{x}(0) + \int_0^t e^{A(t-\tau)} B\mathbf{u}(\tau)\mathrm{d}\tau \tag{23}$$

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

$$G(s) = C\left(sI - A\right)^{-1} B + D \tag{24}$$

- $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $det(A) = ad - bc$

- $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$, $det(A) = a.det(\begin{bmatrix} e & f \\ h & i \end{bmatrix}) - b.det(\begin{bmatrix} d & f \\ g & i \end{bmatrix}) + c.det(\begin{bmatrix} d & e \\ g & h \end{bmatrix})$

- $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $det(A) = ad - bc$, $A^{-1} = \frac{1}{det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

- $\alpha = \mathbf{x}^T A \mathbf{x}$, 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, $\frac{\partial \alpha}{\partial \mathbf{x}} = 2\mathbf{x}^T A$.