Modellierung eines Roboters

Nicolas Schäfer
Saarbrücken, 24. Oktober 2025

Inhaltsverzeichnis

1 Modellierung eines Roboters

Wir halten uns an [5] und [4]

$$\underbrace{f - mg}_{\text{Kräftebilanz}} = \underbrace{m\ddot{x}}_{\text{Newtons 2nd Law}} = \frac{\mathrm{d}}{\mathrm{d}t}[m\dot{x}]$$

$$f - \frac{\partial}{\partial x}\underbrace{[mgx(t)]}_{\text{Potenzielle Energie } V} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{x}}\underbrace{\left[m\frac{1}{2}\dot{x}(t)^2\right]}_{\text{kinetische Energie } T}$$

$$f - \frac{\partial V}{\partial x} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{x}} \Rightarrow f = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x}$$

Wobei $L = T - V = \frac{1}{2}m\dot{x}(t)^2 - mgx(t)$ gilt. Für generalisierte Koordinaten $\mathbf{q} = \theta \in \mathbb{R}^n$

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{\mathbf{q}}_i}\right) - \frac{\partial T}{\partial \mathbf{q}_i} + \frac{\partial V}{\partial \mathbf{q}_i} = \tau_i \quad i = 1, \cdots, n$$

Für die kinetische Energie eines starren Körpers gilt:

$$T = \frac{1}{2}m\mathbf{v}^T\mathbf{v} + \frac{1}{2}\boldsymbol{\omega}^T I\boldsymbol{\omega}$$

Wobei $\mathbf{v}$ die Geschwindigkeit des Schwerpunkts, $\boldsymbol{\omega}$ die Winkelgeschwindigkeit und I der Trägheitstensor bezüglich des Schwerpunkts ist. Für den Trägheitstensor gilt:

$$I = \int_V \rho(\mathbf{r})[(\mathbf{r}\cdot\mathbf{r})\mathbf{1} - \mathbf{r}\otimes\mathbf{r}]\mathrm{d}V \tag{1}$$

Für ein Mehrkörpersystem mit generalisierten Koordinaten $\mathbf{q}$ lässt sich die Gesamtenergie in kompakter Form schreiben als:

$$T = \frac{1}{2}\dot{\mathbf{q}}^T M(\mathbf{q})\dot{\mathbf{q}}$$

wobei $M(\mathbf{q})$ die konfigurationsabhängige Massenmatrix ist, die alle Massen und Trägheiten der einzelnen Glieder erfasst. Für die Geschwindigkeiten gilt mittels der Kettenregel:

$$\mathbf{v} = J_v(\mathbf{q})\dot{\mathbf{q}} \quad \text{und} \quad \boldsymbol{\omega} = J_\omega(\mathbf{q})\dot{\mathbf{q}} \quad J_v = \frac{\partial \mathbf{r}}{\partial \mathbf{q}} \quad \text{und} \quad J_\omega = \frac{\partial \boldsymbol{\omega}}{\partial \dot{\mathbf{q}}}$$

Einsetzen in die kinetische Energie liefert:

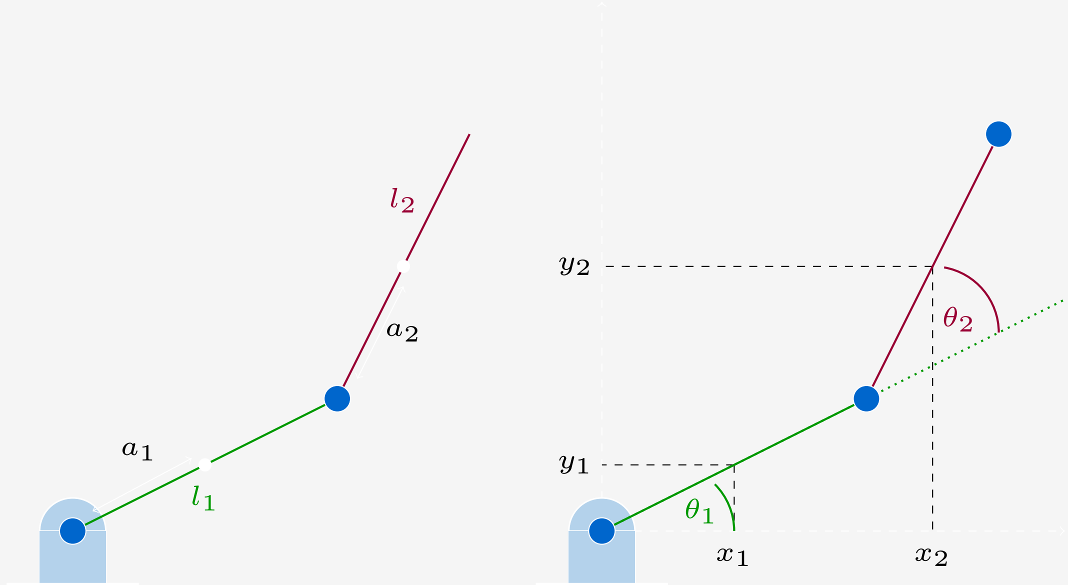
$$\begin{aligned} T &= \frac{1}{2}m(J_v\dot{\mathbf{q}})^T(J_v\dot{\mathbf{q}}) + \frac{1}{2}(J_\omega\dot{\mathbf{q}})^T I(J_\omega\dot{\mathbf{q}}) \\ &= \frac{1}{2}\dot{\mathbf{q}}^T(mJ_v^TJ_v)\dot{\mathbf{q}} + \frac{1}{2}\dot{\mathbf{q}}^T(J_\omega^T I J_\omega)\dot{\mathbf{q}} \\ &= \frac{1}{2}\dot{\mathbf{q}}^T\underbrace{\left(mJ_v^TJ_v + J_\omega^T I J_\omega\right)}_{M(\mathbf{q})}\dot{\mathbf{q}} \end{aligned}$$

Die Gesamtenergie setzt sich zusammen aus der Summe der Energien aller n Glieder:

$$M(\mathbf{q}) = \sum_{i=1}^{n} m_i J_{v_i}^T J_{v_i} + J_{\omega_i}^T I_i J_{\omega_i}$$

1.1 Two Link Revolute Manipulator

Wir bezeichnen $\mathbf{q} = \begin{bmatrix}\theta_1 \\ \theta_2\end{bmatrix}$ als die generalisierten Koordinaten des Systems. Die Schwerpunkte der beiden Links bezeichnen wir mit (x_1, y_1) und (x_2, y_2). Die Abstände der Schwerpunkte von den Gelenken werden mit a_1 und a_2 bezeichnet.



Für die Koordinaten der Schwerpunkte gilt:

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} a_1 \cdot \cos(\theta_1) \\ a_1 \cdot \sin(\theta_1) \\ 0 \end{pmatrix} \qquad \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} l_1 \cdot \cos(\theta_1) + a_2 \cos(\theta_1 + \theta_2) \\ l_1 \cdot \sin(\theta_1) + a_2 \sin(\theta_1 + \theta_2) \\ 0 \end{pmatrix}$$

1.1.1 Formulierung der Massenmatrix

Für die Massenmatrix gilt hier:

$$M(\mathbf{q}) = m_1 J_{v_1}^T J_{v_1} + J_{\omega_1}^T I_1 J_{\omega_1} + m_2 J_{v_2}^T J_{v_2} + J_{\omega_2}^T I_2 J_{\omega_2}$$

Hierbei bezeichnet J die Jacobi-Matrix, I_i das Trägheitstensor und m_i die Masse des i-ten Links.

$$J_{v_1} = \frac{\partial \mathbf{r}_1}{\partial \mathbf{q}} = \frac{\partial(x_1, y_1, z_1)}{\partial(\theta_1, \theta_2)} = \begin{pmatrix} \frac{\partial x_1}{\partial \theta_1} & \frac{\partial x_1}{\partial \theta_2} \\ \frac{\partial y_1}{\partial \theta_1} & \frac{\partial y_1}{\partial \theta_2} \\ \frac{\partial z_1}{\partial \theta_1} & \frac{\partial z_1}{\partial \theta_2} \end{pmatrix} = \begin{pmatrix} -a_1 \sin(\theta_1) & 0 \\ a_1 \cos(\theta_1) & 0 \\ 0 & 0 \end{pmatrix}$$

$$J_{v_2} = \frac{\partial \mathbf{r}_2}{\partial \mathbf{q}} = \frac{\partial(x_2, y_2, z_2)}{\partial(\theta_1, \theta_2)} = \begin{pmatrix} \frac{\partial x_2}{\partial \theta_1} & \frac{\partial x_2}{\partial \theta_2} \\ \frac{\partial y_2}{\partial \theta_1} & \frac{\partial y_2}{\partial \theta_2} \\ \frac{\partial z_2}{\partial \theta_1} & \frac{\partial z_2}{\partial \theta_2} \end{pmatrix} = \begin{pmatrix} -l_1 \sin(\theta_1) - a_2 \sin(\theta_1 + \theta_2) & -a_2 \sin(\theta_1 + \theta_2) \\ l_1 \cos(\theta_1) + a_2 \cos(\theta_1 + \theta_2) & a_2 \cos(\theta_1 + \theta_2) \\ 0 & 0 \end{pmatrix}$$

Für das Matrizenprodukt $J_{v_1}^T J_{v_1}$ gilt:

$$\begin{aligned} J_{v_1}^T J_{v_1} &= \begin{pmatrix} -a_1 \sin(\theta_1) & a_1 \cos(\theta_1) & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} -a_1 \sin(\theta_1) & 0 \\ a_1 \cos(\theta_1) & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} a_1^2 \sin^2(\theta_1) + a_1^2 \cos^2(\theta_1) & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} a_1^2 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Für das Matrizenprodukt $J_{v_2}^T J_{v_2}$ gilt:

$$
\begin{aligned}
J_{v_2}^T J_{v_2} &= \begin{pmatrix} -l_1\sin(\theta_1) - a_2\sin(\theta_1+\theta_2) & l_1\cos(\theta_1) + a_2\cos(\theta_1+\theta_2) & 0 \\ -a_2\sin(\theta_1+\theta_2) & a_2\cos(\theta_1+\theta_2) & 0 \end{pmatrix} \\
&\quad \cdot \begin{pmatrix} -l_1\sin(\theta_1) - a_2\sin(\theta_1+\theta_2) & -a_2\sin(\theta_1+\theta_2) \\ l_1\cos(\theta_1) + a_2\cos(\theta_1+\theta_2) & a_2\cos(\theta_1+\theta_2) \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} (l_1^2 + a_2^2 + 2l_1a_2\cos(\theta_2)) & (a_2^2 + l_1a_2\cos(\theta_2)) \\ (a_2^2 + l_1a_2\cos(\theta_2)) & a_2^2 \end{pmatrix}
\end{aligned}
$$

Für die Winkelgeschwindigkeiten setzen wir $\boldsymbol{\omega}_1 = \begin{pmatrix} 0 & 0 & \theta_1' \end{pmatrix}^T$ und $\boldsymbol{\omega}_2 = \begin{pmatrix} 0 & 0 & \theta_1' + \theta_2' \end{pmatrix}^T$.

$$
J_{\omega_1} = \frac{\partial \boldsymbol{\omega}_1}{\partial \dot{\mathbf{q}}} = \frac{\partial(\omega_1)}{\partial({\theta_1}', {\theta_2}')} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \frac{\partial \omega_1}{\partial {\theta_1}'} & \frac{\partial \omega_1}{\partial {\theta_2}'} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}
$$

$$
J_{\omega_2} = \frac{\partial \boldsymbol{\omega}_2}{\partial \dot{\mathbf{q}}} = \frac{\partial(\omega_2)}{\partial({\theta_1}', {\theta_2}')} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \frac{\partial \omega_2}{\partial {\theta_1}'} & \frac{\partial \omega_2}{\partial {\theta_2}'} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}
$$

Für den Trägheitstensor I_i gilt:

$$
I = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} = \begin{pmatrix} \int (y^2+z^2)\mathrm{d}m & -\int xy\mathrm{d}m & -\int xz\mathrm{d}m \\ -\int yx\mathrm{d}m & \int (x^2+z^2)\mathrm{d}m & -\int yz\mathrm{d}m \\ -\int zx\mathrm{d}m & -\int zy\mathrm{d}m & \int (x^2+y^2)\mathrm{d}m \end{pmatrix}
$$

Für symmetrische Körper (z.B. Zylinder, Kugel, Quader) gilt: $I_{xy} = I_{xz} = I_{yz} = 0$ also erhalten wir eine Diagonalmatrix.

$$
I_i = \begin{pmatrix} \int_y^2 \mathrm{d}m & 0 & 0 \\ 0 & \int x^2\mathrm{d}m & 0 \\ 0 & 0 & \int (x^2+y^2)\mathrm{d}m \end{pmatrix} = \begin{pmatrix} \int_V \rho y^2 \mathrm{d}V & 0 & 0 \\ 0 & \int_V \rho x^2 \mathrm{d}V & 0 \\ 0 & 0 & \int_V \rho(x^2+y^2)\mathrm{d}V \end{pmatrix}
$$

Für die Massenmatrix $M(\mathbf{q})$ gilt somit:

$$
\begin{aligned}
M(\mathbf{q}) &= m_1 J_{v_1}^T J_{v_1} + J_{\omega_1}^T I_1 J_{\omega_1} + m_2 J_{v_2}^T J_{v_2} + J_{\omega_2}^T I_2 J_{\omega_2} \\
&= m_1 \begin{pmatrix} a_1^2 & 0 \\ 0 & 0 \end{pmatrix} + \underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} I_{1,xx} & 0 & 0 \\ 0 & I_{1,yy} & 0 \\ 0 & 0 & I_{1,zz} \end{pmatrix}}_{= \begin{pmatrix} 0 & 0 & I_{1,zz} \\ 0 & 0 & 0 \end{pmatrix}} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} \\
&\quad + m_2 \begin{pmatrix} (l_1^2 + a_2^2 + 2l_1a_2\cos(\theta_2)) & (a_2^2 + l_1a_2\cos(\theta_2)) \\ (a_2^2 + l_1a_2\cos(\theta_2)) & a_2^2 \end{pmatrix} \\
&\quad + \underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} I_{2,xx} & 0 & 0 \\ 0 & I_{2,yy} & 0 \\ 0 & 0 & I_{2,zz} \end{pmatrix}}_{= \begin{pmatrix} 0 & 0 & I_{2,zz} \\ 0 & 0 & I_{2,zz} \end{pmatrix}} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} m_1a_1^2 + I_{1,zz} + m_2 \cdot (l_1^2 + a_2^2 + 2l_1a_2\cos(\theta_2)) + I_{2,zz} & m_2(a_2^2 + l_1a_2\cos(\theta_2)) + I_{2,zz} \\ m_2(a_2^2 + l_1a_2\cos(\theta_2)) + I_{2,zz} & m_2a_2^2 + I_{2,zz} \end{pmatrix}
\end{aligned}
$$

Die Trägheitselemente $I_{i,xx}$ und $I_{i,yy}$ sind nicht relevant. Für unser Model verwenden wir den Standardansatz eines dünnen Stabes der Länge l_i und Masse m_i. Somit gilt $z = y \approx 0$ und wir erhalten: $I_{i,xx} = I_{i,yy} = 0$

$$I_{i,zz} \approx \int_{-\frac{l_i}{2}}^{\frac{l_i}{2}} \underbrace{\rho}_{\frac{m_i}{l_i}} x^2 \mathrm{d}x = \frac{1}{3} \cdot \frac{m_i}{l_i} \cdot \left[\left(\frac{l_i}{2}\right)^3 - \left(-\frac{l_i}{2}\right)^3 \right] = \frac{1}{12} m_i l_i^2 \Rightarrow I_i = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{12} m_i l_i^2 \end{pmatrix}$$

1.1.2 Aufstellen der Coriolis-Matrix

Für die partiellen Ableitungen der kinetischen Energie $T(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} \dot{\mathbf{q}}^T M(\mathbf{q}) \dot{\mathbf{q}}$ gilt:

$$\begin{aligned}
\frac{\partial T}{\partial \theta_1} &= \frac{1}{2} \frac{\partial M_{11}}{\partial \theta_1} {\theta_1'}^2 + \frac{\partial M_{12}}{\partial \theta_1} \theta_1' \theta_2' + \frac{1}{2} \frac{\partial M_{22}}{\partial \theta_1} \theta_2'(t)^2 \\
\frac{\partial T}{\partial \theta_2} &= \frac{1}{2} \frac{\partial M_{11}}{\partial \theta_2} {\theta_1'}^2 + \frac{\partial M_{12}}{\partial \theta_2} \theta_1' \theta_2' + \frac{1}{2} \frac{\partial M_{22}}{\partial \theta_2} \theta_2'(t)^2 \\
\frac{\partial T}{\partial \theta_1'} &= M_{11} \theta_1' + M_{12} \theta_2' \\
\frac{\partial T}{\partial \theta_2'} &= M_{12} \theta_1' + M_{22} \theta_2'
\end{aligned}$$

Für die zeitlichen Ableitungen der Größen gilt

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial T}{\partial \theta_1'} &= \underbrace{\frac{\mathrm{d}}{\mathrm{d}t} M_{11}}_{\frac{\partial M_{11}}{\partial \theta_1} \theta_1' + \frac{\partial M_{11}}{\partial \theta_2} \theta_2'} \cdot \theta_1' + M_{11} \theta_1'' + \underbrace{\frac{\mathrm{d}}{\mathrm{d}t} M_{12}}_{\frac{\partial M_{12}}{\partial \theta_1} \theta_1' + \frac{\partial M_{12}}{\partial \theta_2} \theta_2'} \cdot \theta_2' + M_{12} \theta_2'' \\
&= \frac{\partial M_{11}}{\partial \theta_1} {\theta_1'}^2 + \frac{\partial M_{11}}{\partial \theta_2} \theta_1' \theta_2' + M_{11} \theta_1'' + \frac{\partial M_{12}}{\partial \theta_1} \theta_1' \theta_2' + \frac{\partial M_{12}}{\partial \theta_2} {\theta_2'}^2 + M_{12} \theta_2'' \\
&= \frac{\partial M_{11}}{\partial \theta_1} {\theta_1'}^2 + \left[\frac{\partial M_{11}}{\partial \theta_2} + \frac{\partial M_{12}}{\partial \theta_1} \right] \theta_1' \theta_2' + \frac{\partial M_{12}}{\partial \theta_2} {\theta_2'}^2 + M_{11} \theta_1'' + M_{12} \theta_2''
\end{aligned}$$

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial T}{\partial \theta_2'} &= \underbrace{\frac{\mathrm{d}}{\mathrm{d}t} M_{12}}_{\frac{\partial M_{12}}{\partial \theta_1} \theta_1' + \frac{\partial M_{12}}{\partial \theta_2} \theta_2'} \cdot \theta_1' + M_{12} \theta_1'' + \underbrace{\frac{\mathrm{d}}{\mathrm{d}t} M_{22}}_{\frac{\partial M_{22}}{\partial \theta_1} \theta_1' + \frac{\partial M_{22}}{\partial \theta_2} \theta_2'} \cdot \theta_2' + M_{22} \theta_2'' \\
&= \frac{\partial M_{12}}{\partial \theta_1} {\theta_1'}^2 + \frac{\partial M_{12}}{\partial \theta_2} \theta_1' \theta_2' + M_{12} \theta_1'' + \frac{\partial M_{22}}{\partial \theta_1} \theta_1' \theta_2' + \frac{\partial M_{22}}{\partial \theta_2} {\theta_2'}^2 + M_{22} \theta_2''
\end{aligned}$$

Es gilt

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial T}{\partial \theta_1'} - \frac{\partial T}{\partial \theta_1} &= \left[\frac{\partial M_{11}}{\partial \theta_1} - \frac{1}{2} \frac{\partial M_{11}}{\partial \theta_1} \right] {\theta_1'}^2 + \left[\frac{\partial M_{11}}{\partial \theta_2} + \frac{\partial M_{12}}{\partial \theta_1} - \frac{\partial M_{12}}{\partial \theta_1} \right] \theta_1' \theta_2' + \left[\frac{\partial M_{12}}{\partial \theta_2} - \frac{1}{2} \frac{\partial M_{22}}{\partial \theta_1} \right] {\theta_2'}^2 \\
&\quad + M_{11} \theta_1'' + M_{12} \theta_2''
\end{aligned}$$

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial T}{\partial \theta_2'} - \frac{\partial T}{\partial \theta_2} &= \left[\frac{\partial M_{12}}{\partial \theta_1} - \frac{1}{2} \frac{\partial M_{11}}{\partial \theta_2} \right] {\theta_1'}^2 + \left[\frac{\partial M_{12}}{\partial \theta_2} + \frac{\partial M_{22}}{\partial \theta_1} - \frac{\partial M_{12}}{\partial \theta_2} \right] \theta_1' \theta_2' + \left[\frac{\partial M_{22}}{\partial \theta_2} - \frac{1}{2} \frac{\partial M_{22}}{\partial \theta_2} \right] {\theta_2'}^2 \\
&\quad + M_{12} \theta_1'' + D_{22} \theta_2'' \\
&= \underbrace{\left[\frac{\partial M_{12}}{\partial \theta_1} \theta_1' - \frac{1}{2} \frac{\partial M_{11}}{\partial \theta_2} \theta_1' \right]}_{C_{21}} \theta_1' + \underbrace{\left[\frac{1}{2} \frac{\partial M_{22}}{\partial \theta_2} \theta_2' + \frac{\partial M_{22}}{\partial \theta_1} \theta_1' \right]}_{C_{22}} \theta_2' + M_{12} \theta_1'' + M_{22} \theta_2''
\end{aligned}$$

Für die Einträge der Matrix C gilt

$$
\begin{aligned}
C_{11} &= \frac{1}{2}\underbrace{\frac{\partial M_{11}}{\partial \theta_1}}_{=0}\theta_1' + \underbrace{\frac{\partial M_{11}}{\partial \theta_2}}_{=-2m_2l_1a_2\sin(\theta_2)}\theta_2' = -2m_2l_1a_2\sin(\theta_2)\theta_2' \\
C_{12} &= \underbrace{\frac{\partial M_{12}}{\partial \theta_2}}_{=-m_2l_1a_2\sin(\theta_2)}\theta_2' - \frac{1}{2}\underbrace{\frac{\partial M_{22}}{\partial \theta_1}}_{=0}\theta_2' = -m_2l_1a_2\sin(\theta_2)\theta_2' \\
C_{21} &= \underbrace{\frac{\partial M_{12}}{\partial \theta_1}}_{=0}\theta_1' - \frac{1}{2}\underbrace{\frac{\partial M_{11}}{\partial \theta_2}}_{=-2m_2l_1a_2\sin(\theta_2)}\theta_1' = m_2l_1a_2\sin(\theta_2)\theta_1' \\
C_{22} &= \frac{1}{2}\underbrace{\frac{\partial M_{22}}{\partial \theta_2}}_{=0}\theta_2' + \underbrace{\frac{\partial M_{22}}{\partial \theta_1}}_{=0}\theta_1' = 0
\end{aligned}
$$

Es gilt somit

$$
C(\mathbf{q},\dot{\mathbf{q}})\cdot\dot{\mathbf{q}} = \begin{pmatrix} -2m_2l_1a_2\sin(\theta_2)\theta_2' & -m_2l_1a_2\sin(\theta_2)\theta_2' \\ m_2l_1a_2\sin(\theta_2)\theta_1' & 0 \end{pmatrix}\begin{pmatrix}\theta_1' \\ \theta_2'\end{pmatrix}
$$

1.1.3 Gravitationsterme

Für die potenzielle Energie V gilt:

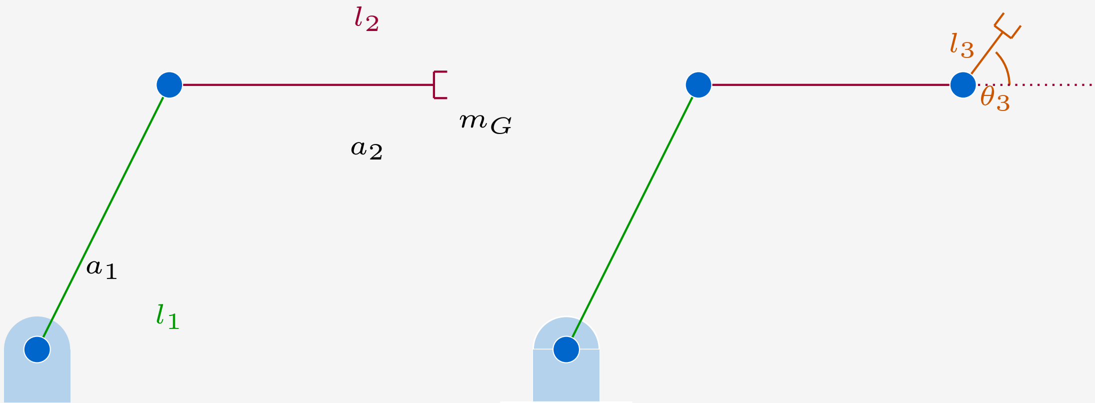
$$
V = m_1\cdot g\cdot y_1 + m_2\cdot g\cdot y_2 = m_1ga_1\sin(\theta_1) + m_2g\left(l_1\sin(\theta_1) + a_2\sin(\theta_1+\theta_2)\right)
$$

Für die partiellen Ableitungen erhalten wir

$$
G(\mathbf{q}) = \frac{\partial V}{\partial \mathbf{q}} = \begin{pmatrix}\frac{\partial V}{\partial \theta_1} \\ \frac{\partial V}{\partial \theta_2}\end{pmatrix} = \begin{pmatrix} m_1ga_1\cos(\theta_1) + m_2g\left(l_1\cos(\theta_1) + a_2\cos(\theta_1+\theta_2)\right) \\ m_2g\left(l_1\cos(\theta_1) + a_2\cos(\theta_1+\theta_2)\right)\end{pmatrix}
$$

1.2 Einfaches Greifobjekt

Wir modellieren ein Variables Greifobjekt, welches am Ende von Link 2 gehalten wird. Wir betrachten nachfolgend das Greifobjekt als Punktmasse mit Masse m_G.



Für die Masse von Link 2 inklusive Greifobjekt gilt:

$$m_i' = m_i + m_G$$

Für den Schwerpunkt von Link 2 inklusive Greifobjekt gilt:

$$\mathbf{a}_2' = \frac{m_i \cdot \mathbf{a}_2 + m_G \cdot \mathbf{d}}{m_i + m_G} \quad \mathbf{d}(\theta_3) = \begin{pmatrix} l_2 + r \cdot \cos(\theta_3) \\ r \cdot \sin(\theta_3) \\ 0 \end{pmatrix}$$

Für den Trägheitstensor von Link 2 inklusive Greifobjekt gilt:

$$I_i' = I_i + m_i \cdot S(\mathbf{a}_i - \mathbf{a}_i') + m_G \cdot S(\mathbf{d} - \mathbf{a}_i') \quad S(\mathbf{v}) = \begin{pmatrix} v_y^2 + v_z^2 & -v_x v_y & -v_x v_z \\ -v_x v_y & v_x^2 + v_z^2 & -v_y v_z \\ -v_x v_z & -v_y v_z & v_x^2 + v_y^2 \end{pmatrix}$$

Wobei $S(\mathbf{v})$ die Steiner-Matrix ist. Wir erhalten die Matrix direkt aus der Definition des Trägheitstensors (1) einer Punktmasse mit $\rho(\mathbf{r}) = m \cdot \delta(\mathrm{r} - \mathbf{d})$. Für die Verschiebungsvektoren gilt:

$$\mathbf{d} = \begin{pmatrix} l_2 \\ 0 \\ 0 \end{pmatrix} \; \mathbf{a}_2 = \begin{pmatrix} \frac{l_2}{2} \\ 0 \\ 0 \end{pmatrix} \Rightarrow \mathbf{a}_2' = \frac{l_2}{m_2 + m_G} \begin{pmatrix} \frac{m_2}{2} + m_G \\ 0 \\ 0 \end{pmatrix}$$

Für die Steiner-Matrizen gilt entsprechend

$$S(\mathbf{a}_2 - \mathbf{a}_2') = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \left[\frac{l_2}{2} - a_2'\right]^2 & 0 \\ 0 & 0 & \left[\frac{l_2}{2} - a_2'\right]^2 \end{pmatrix} \quad S(\mathbf{d} - \mathbf{a}_2') = \begin{pmatrix} 0 & 0 & 0 \\ 0 & [l_2^2 - a_2']^2 & 0 \\ 0 & 0 & [l_2^2 - a_2']^2 \end{pmatrix}$$

Somit erhalten wir für den Trägheitstensor von Link 2 inklusive Greifobjekt:

$$I_2' = \begin{pmatrix} I_{2,xx} & 0 & 0 \\ 0 & I_{2,yy} + m_2 \cdot \left[\frac{l_2^2}{2} - a_2'\right]^2 + m_G \cdot \left[\frac{l_2^2}{2} - a_2'\right]^2 & 0 \\ 0 & 0 & I_{2,zz} + m_2 \cdot \left[\frac{l_2^2}{2} - a_2'\right]^2 + m_G \cdot \left[\frac{l_2^2}{2} - a_2'\right]^2 \end{pmatrix}$$

Für die Koordinaten des Schwerpunktes von Link 3 gilt:

$$\mathbf{r}_3 = \begin{pmatrix} l_1 \cos(\theta_1) + l_2 \cos(\theta_1 + \theta_2) + l_3 \cos(\theta_1 + \theta_2 + \theta_3) \\ l_1 \sin(\theta_1) + l_2 \sin(\theta_1 + \theta_2) + l_3 \sin(\theta_1 + \theta_2 + \theta_3) \\ 0 \end{pmatrix}$$

Berechnung der Jacobi-Matrizen für Link 3 bezüglich der Geschwindigkeiten:

$$J_{v_3} = \frac{\partial \mathbf{r}_3}{\partial \mathbf{q}}$$

Für die Winkelgeschwindigkeit des Links 3 setzen wir $\boldsymbol{\omega_3} = \begin{pmatrix} 0 & 0 & \theta_1' + \theta_2' + \theta_3' \end{pmatrix}$ und erhalten für die Jacobi-Matrix bezüglich der Winkelgeschwindigkeit:

$$J_{\omega_3} = \frac{\partial \boldsymbol{\omega_3}}{\partial \dot{\mathbf{q}}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

2 Theorie optimaler Steuerungungsprobleme

Wir halten uns an die Definitionen in [1]

Definition 2.1 (Optimales Steuerungsproblem) *Seien $t_0 < t_f$ feste Zeiten und*

$$\begin{aligned}
&\Phi : \mathbb{R}^{n_{\boldsymbol{x}}} \to \mathbb{R},\\
&L : [t_0, t_f] \times \mathbb{R}^{n_{\boldsymbol{x}}} \to \mathbb{R},\\
&f : [t_0, t_f] \times \mathbb{R}^{n_{\boldsymbol{x}}} \times \mathbb{R}^{n_{\boldsymbol{u}}} \to \mathbb{R}^{n_{\boldsymbol{x}}}
\end{aligned}$$

hinreichend glatte Funktionen und $\mathcal{U} \subset \mathbb{R}^{n_{\boldsymbol{u}}}$ eine abgeschlossene konvexe nichtleere Menge.

$$\min_{\boldsymbol{u}} \int_{t_0}^{t_f} L(t, \boldsymbol{x}(t), \boldsymbol{u}(t))\,\mathrm{d}t + \Phi(\boldsymbol{x}(t_f))$$

mit $\boldsymbol{x} \in W_{1,\infty}^{n_{\boldsymbol{x}}}([t_0, t_f]), \boldsymbol{u} \in L_{\infty}^{n_{\boldsymbol{u}}}([t_0, t_f])$ und

$$\dot{\boldsymbol{x}}(t) = f(t, \boldsymbol{x}(t), \boldsymbol{u}(t)), \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0, \quad \boldsymbol{u}(t) \in \mathcal{U} \text{ f.ü. } t \in [t_0, t_f]$$

Wir betrachten ein Steuerungsproblem und definieren die Value-Function V mittels

$$V(x(t), t) := \min_{u} \left\{ \int_{t}^{t_f} L(s, x(s), u(s))\,\mathrm{d}s + \Phi(x(t_f)) \right\}$$

Satz 2.2 (Hamilton-Jacobi-Bellman Gleichung) *Die Value-Function erfüllt*

$$0 = \frac{\partial V(x,t)}{\partial t} + \min_{u} \left[L(t,x,u) + \frac{\partial V(x,t)}{\partial x} \cdot f(t,x,u) \right], \quad x \in \mathbb{R}^n,\ t_0 \le t \le t_f$$

mit den Randbedingungen

$$V(x, t_f) = \Phi(x(t_f)) \quad V(x, t_0) = \min_{u} \int_{t_0}^{t_f} L(t, x(t), u(t))\,\mathrm{d}t + \Phi(x(t_f))$$

Beweis:
Die Value-Function $V(x,t)$ beschreibt die minimalen Kosten von Zustand x zum Endzeitpunkt t_f. Für $s \in [t, t_f]$ gilt

- Systemdynamik

$$x'(s) = f(x(s), u(s)), \quad x(t) = x$$

- Kostenfunktional

$$J(u) = \int_{t}^{t_f} L(s, x(s), u(s))\,\mathrm{d}s + \Phi(x(t_f))$$

Angenommen, wir wechseln zum Zeitpunkt $t+h$ zu einer optimalen Steuerung, dann gilt für unser Zielfunktional

$$\begin{aligned}
J(u) &= \int_{t}^{t+h} L(s, x(s), u(s))\,\mathrm{d}s + \int_{t+h}^{t_f} L(s, x(s), u(s))\,\mathrm{d}s + \Phi(x(t_f))\\
&= \int_{t}^{t+h} L(s, x(s), u(s))\,\mathrm{d}s + V(x(t+h), t+h)
\end{aligned}$$

Da $V(x(t),t)$ die minimalen Kosten vom Zustand x zur Zeit t angibt, gilt:

$$\begin{aligned} V(x(t),t) &\leq \int_t^{t+h} L(s,x(s),u(s))\,\mathrm{d}s + V(x(t+h),t+h) \\ \Rightarrow 0 &\leq \int_t^{t+h} L(s,x(s),u(s))\,\mathrm{d}s + V(x(t+h),t+h) - V(x(t),t) \end{aligned}$$

Im Grenzübergang $h \to 0$ erhalten wir

$$\begin{aligned} 0 &\leq \lim_{h\to 0} \frac{1}{h} \int_t^{t+h} L(s,x(s),u(s))\,\mathrm{d}s + \lim_{h\to 0} \frac{V(x(t+h),t+h) - V(x,t)}{h} \\ &= L(t,x(t),u(t)) + \frac{\mathrm{d}}{\mathrm{d}t} V(x(t),t) \\ &= L(t,x(t),u(t)) + \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} \cdot x'(t) \\ &= L(t,x(t),u(t)) + \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} \cdot f(t,x(t),u(t)) \end{aligned}$$

Durch Minimierung bezüglich u erhalten wir die Gleichheit und somit

$$\begin{aligned} 0 &= \frac{\partial V}{\partial t} + \min_u \left[L(t,x(t),u(t)) + \frac{\partial V}{\partial x} \cdot f(t,x(t),u(t)) \right] \\ &= \frac{\partial V}{\partial t} + \mathsf{H}\left(t, x(t), \frac{\partial V}{\partial x}(t,x(t))\right) \end{aligned}$$

□

2.1 Herleitung des Minimumprinzips

Berechnung der totalen Ableitung der Value-Function V entlang der Charakteristik $x(t)$:

$$\frac{dV}{dt} = \frac{\partial V}{\partial x}\frac{dx}{dt} + \frac{\partial V}{\partial t} = p(x,t) \cdot x'(t) - H\left(t, x, \frac{\partial V}{\partial x}\right)$$

Berechnung der totalen Ableitung von $p = \frac{\partial V}{\partial x}$:

$$\frac{dp}{dt} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial V}{\partial x}\right) = \frac{\partial^2 V}{\partial x^2} \cdot x'(t) + \frac{\partial V}{\partial x \partial t}$$

Differenzieren der HJB bezüglich x liefert

$$0 = \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial V}{\partial t} + H\left(t, x, \frac{\partial V}{\partial x}\right)\right] = \frac{\partial V}{\partial t \partial x} + \frac{\partial H}{\partial x} + \frac{\partial H}{\partial p} \cdot \frac{\partial^2 V}{\partial x^2}$$

Einsetzen der Gleichung liefert:

$$\frac{dp}{dt} = \frac{\partial^2 V}{\partial x^2} \cdot x'(t) + \left(-\frac{\partial H}{\partial x} - \frac{\partial H}{\partial p} \cdot \frac{\partial^2 V}{\partial x^2}\right) = \frac{\partial^2 V}{\partial x^2} \cdot \left(x'(t) - \frac{\partial H}{\partial p}\right) - \frac{\partial H}{\partial x}$$

Wir erhalten somit die bekannten, notwendigen Optimalitätsbedingungen:

$$p'(t) = -H_x, \quad x'(t) = f(t,x,u^*), \quad u^* = \mathsf{argmin}\,[L(t,x,u) + p(t,x) f(t,x,u)]$$

2.2 Anwendung auf Linear-Quadratische Probleme

2.2.1 Finite time problem

Gegeben sei das Steuerungsproblem

$$\min_u \int_{t_0}^{t_f} x^T Q x + u^T R u \, \mathrm{d}t + \underbrace{(x(t_f) - x_f)^T S (x(t_f) - x_f)}_{x(t_f)^T S x(t_f) - 2x_f^T S x(t_f) + x_f^T S x_f} = V(t_0, x(t_0)) + x_f^T S x_f$$

unter der Nebenbedingung $x'(t) = Ax + Bu$ und $x(t_0) = x_0$. Der Term $x_f^T S x_f$ ist unabhängig von u und kann bei der Minimierung weggelassen werden.

$$V(x,t) = x^T K(t) x + 2s(t)^T x + r(t) \quad V(x,t_f) = x(t_f)^T K(t_f) x(t_f) + 2s(t_f)^T x(t_f) + r(t_f)$$

Wir erhalten für die Endwerte

$$K(t_f) = S, \quad s(t_f) = -Sx_f \quad r(t_f) = 0$$

$$V_t = x^T K'(t) x + 2s'(t)^T x + r'(t) \quad V_x = 2K(t)x + 2s(t)$$

$$\frac{\partial}{\partial u}\left[x^T Q x + u^T R u + V_x^T \cdot (Ax + Bu)\right] = 2Ru + B^T \cdot V_x \Rightarrow u^*(t) = -\frac{1}{2} R^{-1} B^T V_x$$

$$\begin{aligned}
(u^*)^T R u^* &= \left(-\frac{1}{2} R^{-1} B^T V_x\right)^T R \left(-\frac{1}{2} R^{-1} B^T V_x\right) \\
&= \frac{1}{4} \cdot \left(R^{-1} B^T V_x\right)^T R \left(R^{-1} B^T V_x\right) \\
&= \frac{1}{4} \cdot \left(2R^{-1} B^T (K(t)x + s(t))\right)^T R \left(2R^{-1} B^T (K(t)x + s(t))\right) \\
&= (K(t)x + s(t))^T B R^{-1} R R^{-1} B^T (K(t)x + s(t)) \\
&=< K(t)x + s(t), P(2K(t)x + s(t)) > \\
&=< K(t)x, PK(t)x > + \underbrace{< K(t)x, Ps(t) >}_{< s(t), P^T K(t)x >} + < s(t), PK(t)x > + < s(t), Ps(t) > \\
&= x^T K(t)^T P K(t) x + 2 \cdot s(t)^T P K(t) x + s(t)^T P s(t)
\end{aligned}$$

mit $P = BR^{-1}B^T$ und $P^T = P$.

$$\begin{aligned}
V_x^T \left(-\frac{1}{2} B R^{-1} B^T V_x\right) &= (2K(t)x + 2s(t))^T \left(-\frac{1}{2} B R^{-1} B^T (2K(t)x + 2s(t))\right) \\
&= -2 \cdot (K(t)x + s(t))^T P (K(t)x + s(t)) \\
&= -2 \cdot x^T K(t)^T P K(t) x - 4 \cdot s(t)^T P K(t) x - 2 \cdot s(t)^T P s(t)
\end{aligned}$$

$$V_x^T A x = 2(K(t)x + s(t))^T A x = 2x^T K(t)^T A x + 2s(t)^T A x = x^T (K(t)^T A + A^T K(t)) x + 2s(t)^T A x$$

$$-V_t = \min_u \left[x^T Qx + u^T Ru + V_x^T \cdot (Ax + Bu)\right]$$

$$-V_t = x^T Qx + (u^*)^T Ru^* + V_x^T \cdot \left(Ax - \frac{1}{2}BR^{-1}B^T V_x\right)$$

$$-x^T K'(t)x - 2s'(t)^T x - r'(t) = x^T(Q - KPK + KA + A^T K)x - 2 \cdot ((PK - A)^T s(t))^T x - s(t)^T Ps(t)$$

$$\begin{aligned}
K'(t) &= -Q + K(t)BR^{-1}B^T K(t) - K(t)A - A^T K(t) \quad K(t_f) = S \\
s'(t) &= (K(t)BR^{-1}B^T - A^T)s(t) \quad s(t_f) = -Sx_f \\
r'(t) &= s(t)^T BR^{-1}B^T s(t) \quad r(t_f) = 0
\end{aligned}$$

2.2.2 infinite time problem

Im Grenzfall $t \to \infty$ betrachten wir das Steuerungsproblem

$$\min_u \int_{t_0}^{\infty} (x - x_f)^T Q(x - x_f) + u^T Ru \,\mathrm{d}t = V(x(t_0))$$

unter der Nebenbedingung $x'(t) = Ax + Bu$ und $x(t_0) = x_0$. Als Ansatz für die Value Funktion wählen wir erneut

$$V(x) = x^T Kx + 2s^T x + r \Rightarrow V_x = 2Kx + 2s$$

Es gilt erneut

$$\frac{\partial}{\partial u}\left((x - x_f)^T Q(x - x_f) + u^T Ru + V_x^T(Ax + Bu)\right) = 2Ru + B^T V_x = 0 \Rightarrow u^* = -\frac{1}{2}R^{-1}B^T V_x$$

$$(x - x_f)^T Q(x - x_f) = x^T Qx - 2x_f^T Qx + x_f^T Qx_f$$

$$0 = x^T(Q - KPK + KA + A^T K)x - 2((PK - A)^T s + Qx_f)^T x - s^T(P)s + x_f^T Qx_f$$

3 Linearisierung der Dynamik

Wir gehen analog zu [2] vor. Wir definieren den Zustand $\mathbf{x} = \begin{bmatrix}\mathbf{q}\\ \dot{\mathbf{q}}\end{bmatrix}$ und schreiben unsere Dynamikgleichung wir folgt um

$$M(\mathbf{q})\ddot{\mathbf{q}} + C(\mathbf{q},\dot{\mathbf{q}})\dot{\mathbf{q}} + G(\mathbf{q}) = \boldsymbol{u} \Rightarrow \underbrace{\begin{bmatrix}\dot{\mathbf{q}}\\ \ddot{\mathbf{q}}\end{bmatrix}}_{\dot{\mathbf{x}}} = \underbrace{\begin{bmatrix}\dot{\mathbf{q}}\\ M^{-1}(\mathbf{q})\left[\boldsymbol{u} - C(\mathbf{q},\dot{\mathbf{q}})\dot{\mathbf{q}} - G(\mathbf{q})\right]\end{bmatrix}}_{f(\mathbf{x},\boldsymbol{u})}$$

Wir linearisieren um den Arbeitspunkt $(\boldsymbol{x}^*, \boldsymbol{u}^*)$

$$\begin{aligned}\dot{\boldsymbol{x}} &= f(\boldsymbol{x},\boldsymbol{u})\\ &\approx \underbrace{f(\boldsymbol{x}^*,\boldsymbol{u}^*)}_{=0} + \underbrace{\frac{\partial f}{\partial x}(\boldsymbol{x}^*,\boldsymbol{u}^*)}_{A(\boldsymbol{x}^*,\boldsymbol{u}^*)}(\boldsymbol{x}-\boldsymbol{x}^*) + \underbrace{\frac{\partial f}{\partial u}(\boldsymbol{x}^*,\boldsymbol{u}^*)}_{B(\boldsymbol{x}^*,\boldsymbol{u}^*)}(\boldsymbol{u}-\boldsymbol{u}^*)\\ &= A(\boldsymbol{x}^*,\boldsymbol{u}^*)\bar{\boldsymbol{x}} + B(\boldsymbol{x}^*,\boldsymbol{u}^*)\bar{\boldsymbol{u}}\end{aligned}$$

Wobei $\bar{\boldsymbol{x}}(t) = \boldsymbol{x}(t) - \boldsymbol{x}^*$ und $\bar{\boldsymbol{u}}(t) = \boldsymbol{u}(t) - \boldsymbol{u}^*$ gilt. Für die Ableitung der Größen gilt

$$\frac{\mathrm{d}\bar{\boldsymbol{x}}}{\mathrm{d}t} = \dot{\boldsymbol{x}} - \underbrace{\frac{\mathrm{d}\boldsymbol{x}^*}{\mathrm{d}t}}_{=0} = \dot{\boldsymbol{x}}(t) \Rightarrow \bar{\boldsymbol{x}}(t) = A\bar{\boldsymbol{x}}(t) + B\bar{\boldsymbol{u}}(t)$$

Somit ist die Rückführung von $\bar{\boldsymbol{x}}$ auf $\mathbf{0}$ äquivalten zu der Rückführung von $\boldsymbol{x}$ auf den Arbeitspunkt $\boldsymbol{x}^*$. Wir definieren $h(\boldsymbol{q},\dot{\boldsymbol{q}},\boldsymbol{u}) := \boldsymbol{u} - C(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}} - G(\boldsymbol{q})$. und wenden die Produktregel auf den Term $M(\boldsymbol{q})\ddot{\boldsymbol{q}}$ an, um die Ableitung von f nach $\boldsymbol{x}$ zu bestimmen.

$$\frac{\partial}{\partial \boldsymbol{x}}\left[M(\boldsymbol{q})\ddot{\boldsymbol{q}}\right] = \frac{\partial}{\partial \boldsymbol{x}}\left[h(\boldsymbol{q},\dot{\boldsymbol{q}},\boldsymbol{u})\right] \Rightarrow \frac{\partial M}{\partial \boldsymbol{x}}\ddot{\boldsymbol{q}} + M(\boldsymbol{q})\frac{\partial \ddot{\boldsymbol{q}}}{\partial \boldsymbol{x}} = \frac{\partial h}{\partial \boldsymbol{x}}$$

Werten wir h im Arbeitspunkt $(\boldsymbol{x}^*,\boldsymbol{u}^*)$ aus, so gilt $\ddot{\boldsymbol{q}} = \dot{\boldsymbol{q}} = \mathbf{0}$, somit fällt der erste Term weg und wir erhalten

$$\frac{\partial h}{\partial \boldsymbol{x}}(\boldsymbol{x}^*,\boldsymbol{u}^*) = \underbrace{\frac{\partial M}{\partial \boldsymbol{x}}\ddot{\boldsymbol{q}}}_{=0} + M(\boldsymbol{x}^*)\frac{\partial \ddot{\boldsymbol{q}}}{\partial \boldsymbol{x}} \Rightarrow \frac{\partial \ddot{\boldsymbol{q}}}{\partial \boldsymbol{x}} = M^{-1}(\boldsymbol{q}^*)\frac{\partial h}{\partial \boldsymbol{x}}(\boldsymbol{x}^*,\boldsymbol{u}^*)$$

Für die Ableitung der Funktion $h = \boldsymbol{u} - C(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}} - G(\boldsymbol{q})$ in $(\boldsymbol{x}^*,\boldsymbol{u}^*)$ gilt

$$\begin{aligned}\frac{\partial h}{\partial \boldsymbol{x}}(\boldsymbol{x}^*,\boldsymbol{u}^*) = \begin{bmatrix}\frac{\partial h}{\partial \boldsymbol{q}} & \frac{\partial h}{\partial \dot{\boldsymbol{q}}}\end{bmatrix} &= \begin{bmatrix}-\frac{\partial C}{\partial \boldsymbol{q}}\dot{\boldsymbol{q}} - \frac{\partial G}{\partial \boldsymbol{q}} & -C(\boldsymbol{q},\dot{\boldsymbol{q}}) - \frac{\partial C}{\partial \dot{\boldsymbol{q}}}\dot{\boldsymbol{q}}\end{bmatrix}(\boldsymbol{q}^*, \dot{\boldsymbol{q}}^* = \mathbf{0})\\ &= \begin{bmatrix}-\frac{\partial G}{\partial \boldsymbol{q}} & -C(\boldsymbol{q}^*,\mathbf{0})\end{bmatrix}\end{aligned}$$

Fassen wir die Ergebnisse zusammen, so gilt für die Ableitung von f nach $\boldsymbol{x}$ im Arbeitspunkt $(\boldsymbol{x}^*,\boldsymbol{u}^*)$

$$A(\boldsymbol{x}^*,\boldsymbol{u}^*) = \begin{bmatrix}\frac{\partial \dot{\boldsymbol{q}}}{\partial \boldsymbol{q}} & \frac{\partial \dot{\boldsymbol{q}}}{\partial \dot{\boldsymbol{q}}}\\ \frac{\partial \ddot{\boldsymbol{q}}}{\partial \boldsymbol{q}} & \frac{\partial \ddot{\boldsymbol{q}}}{\partial \dot{\boldsymbol{q}}}\end{bmatrix}(\boldsymbol{x}^*,\boldsymbol{u}^*) = \begin{bmatrix}0 & I\\ M^{-1}(\boldsymbol{q}^*)\left[-\frac{\partial G}{\partial \boldsymbol{q}}(\boldsymbol{q}^*)\right] & M^{-1}(\boldsymbol{q}^*)\left[-C(\boldsymbol{q}^*,\mathbf{0})\right]\end{bmatrix}$$

Für die Ableitung von f nach $\boldsymbol{u}$ gilt

$$B(\boldsymbol{x}^*,\boldsymbol{u}^*) = \begin{bmatrix}\frac{\partial \dot{\boldsymbol{q}}}{\partial \boldsymbol{u}}\\ \frac{\partial \ddot{\boldsymbol{q}}}{\partial \boldsymbol{u}}\end{bmatrix}(\boldsymbol{x}^*,\boldsymbol{u}^*) = \begin{bmatrix}0\\ M^{-1}(\boldsymbol{q}^*)\end{bmatrix}$$

Satz 3.1 *Gegeben sei die Dynamikgleichung*

$$M(\boldsymbol{q})\ddot{\boldsymbol{q}} + C(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + G(\boldsymbol{q}) = \boldsymbol{u} \Rightarrow \dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u})$$

und ein Arbeitspunkt $(\boldsymbol{x}^*, \boldsymbol{u}^*)$ *mit* $f(\boldsymbol{x}^*, \boldsymbol{u}^*) = \boldsymbol{0}$. *Dann gilt für die Linearisierung:*

$$\begin{aligned}\dot{\boldsymbol{x}} &= A(\boldsymbol{x}^*, \boldsymbol{u}^*)\boldsymbol{x} + B(\boldsymbol{x}^*, \boldsymbol{u}^*)\boldsymbol{u} \\ &= \begin{bmatrix} 0 & I \\ M^{-1}(\boldsymbol{q}^*)\left[-\frac{\partial G}{\partial \boldsymbol{q}}(\boldsymbol{q}^*)\right] & M^{-1}(\boldsymbol{q}^*)\left[-C(\boldsymbol{q}^*, \boldsymbol{0})\right] \end{bmatrix}\boldsymbol{x} + \begin{bmatrix} 0 \\ M^{-1}(\boldsymbol{q}^*) \end{bmatrix}\boldsymbol{u}\end{aligned}$$

3.1 Linearisierung des two link revolute manipulators

Wir definieren $x_1 = q_1$, $x_2 = q_2$, $x_3 = \dot{q}_1$, $x_4 = \dot{q}_2$ und erhalten $\dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u})$ mit

$$\begin{bmatrix} f_1(\boldsymbol{x}, \boldsymbol{u}) \\ f_2(\boldsymbol{x}, \boldsymbol{u}) \end{bmatrix} = \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \end{bmatrix} \quad \begin{bmatrix} f_3(\boldsymbol{x}, \boldsymbol{u}) \\ f_4(\boldsymbol{x}, \boldsymbol{u}) \end{bmatrix} = M(\boldsymbol{q})^{-1}\left[\boldsymbol{u} - C(\boldsymbol{x})\dot{\boldsymbol{q}} - G(\boldsymbol{q})\right]$$

Für die Corioliskraft im Punkt $(\boldsymbol{x}^*, \boldsymbol{u}^*)$, bzw. für die Matrix $C(\boldsymbol{q}, \dot{\boldsymbol{q}} = \boldsymbol{0})$ gilt:

$$C(\boldsymbol{q}^*, \boldsymbol{0}) = \begin{bmatrix} 0 & -m_2 l_1 a_2 \sin(q_2^*) \cdot 0 \\ m_2 l_1 a_2 \sin(q_2^*) \cdot 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Wir berechnen zunächst die Ableitung von h bzgl. $\boldsymbol{x}$ im Punkt $(\boldsymbol{x}^*, \boldsymbol{u}^*)$ und erhalten

$$\frac{\partial h}{\partial \boldsymbol{x}}(\boldsymbol{x}^*, \boldsymbol{u}^*) = \begin{bmatrix} -\frac{\partial G}{\partial \boldsymbol{q}} & -\underbrace{C(\boldsymbol{q}^*, \boldsymbol{0})}_{=\boldsymbol{0}} \end{bmatrix} = \begin{bmatrix} -\frac{\partial G_1}{\partial q_1}(\boldsymbol{q}^*) & -\frac{\partial G_1}{\partial q_2}(\boldsymbol{q}^*) & 0 & 0 \\ -\frac{\partial G_2}{\partial q_1}(\boldsymbol{q}^*) & -\frac{\partial G_2}{\partial q_2}(\boldsymbol{q}^*) & 0 & 0 \end{bmatrix}$$

Für die konkrete Matrix $A = f_{\boldsymbol{x}}(\boldsymbol{x}^*, \boldsymbol{u}^*)$ gilt

$$A(\boldsymbol{x}^*, \boldsymbol{u}^*) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} & \frac{\partial f_1}{\partial x_4} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} & \frac{\partial f_2}{\partial x_4} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} & \frac{\partial f_3}{\partial x_4} \\ \frac{\partial f_4}{\partial x_1} & \frac{\partial f_4}{\partial x_2} & \frac{\partial f_4}{\partial x_3} & \frac{\partial f_4}{\partial x_4} \end{bmatrix}(\boldsymbol{x}^*, \boldsymbol{u}^*) = \begin{bmatrix} \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \\ -M^{-1}(\boldsymbol{q}^*)\frac{\partial G}{\partial \boldsymbol{q}}(\boldsymbol{q}^*) & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \end{bmatrix}$$

Für die konkrete Matrix $B = f_{\boldsymbol{u}}(\boldsymbol{x}^*, \boldsymbol{u}^*)$ gilt

$$B(\boldsymbol{x}^*, \boldsymbol{u}^*) = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \frac{\partial f_1}{\partial u_2} \\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial u_2} \\ \frac{\partial f_3}{\partial u_1} & \frac{\partial f_3}{\partial u_2} \\ \frac{\partial f_4}{\partial u_1} & \frac{\partial f_4}{\partial u_2} \end{bmatrix}(\boldsymbol{x}^*, \boldsymbol{u}^*) = \begin{bmatrix} \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ M^{-1}(\boldsymbol{q}^*) \end{bmatrix}$$

Für die konkret Berechnung der partiellen Ableitungen der Gravitationskraft gilt

$$\begin{aligned} G(\boldsymbol{q}) &= \begin{bmatrix} m_1 g a_1 \cos(q_1) + m_2 g\left(l_1 \cos(q_1) + a_2 \cos(q_1 + q_2)\right) \\ m_2 g\left(l_1 \cos(q_1) + a_2 \cos(q_1 + q_2)\right) \end{bmatrix} \\ \frac{\partial G}{\partial \boldsymbol{q}} &= \begin{bmatrix} \frac{\partial G_1}{\partial q_1} & \frac{\partial G_1}{\partial q_2} \\ \frac{\partial G_2}{\partial q_1} & \frac{\partial G_2}{\partial q_2} \end{bmatrix} \\ &= \begin{bmatrix} -m_1 g a_1 \sin(q_1) - m_2 g\left(l_1 \sin(q_1) + a_2 \sin(q_1 + q_2)\right) & -m_2 g a_2 \sin(q_1 + q_2) \\ -m_2 g\left(l_1 \sin(q_1) + a_2 \sin(q_1 + q_2)\right) & -m_2 g a_2 \sin(q_1 + q_2) \end{bmatrix} \end{aligned}$$

Die Massenmatrix $M(\boldsymbol{q})$ ist gegeben durch den Ausdruck:

$$M(\boldsymbol{q}) = \begin{bmatrix} m_1 a_1^2 + I_{1,zz} + m_2 \cdot (l_1^2 + a_2^2 + 2 l_1 a_2 \cos(q_2)) + I_{2,zz} & m_2 \cdot (a_2^2 + l_1 a_2 \cos(q_2)) + I_{2,zz} \\ m_2 \cdot (a_2^2 + l_1 a_2 \cos(q_2)) + I_{2,zz} & m_2 a_2^2 + I_{2,zz} \end{bmatrix}$$

Für die Inverse der Massenmatrix M^{-1} gilt analytisch mittels $F_{q_2} := m_2 l_1 a_2 \cos(q_2)$

$$\begin{aligned}
\det(M) &= M_{11} M_{22} - M_{12}^2 \\
&= \left[m_1 a_1^2 + I_{1,zz} + m_2 \cdot (l_1^2 + a_2^2 + 2 l_1 a_2 \cos(q_2)) + I_{2,zz}\right] \cdot (m_2 a_2^2 + I_{2,zz}) \\
&\quad - (m_2 \cdot (a_2^2 + l_1 a_2 \cos(q_2)) + I_{2,zz})^2 \\
&= \left[m_1 a_1^2 + I_{1,zz} + I_{2,zz} + m_2 l_1^2 + m_2 a_2^2 + 2 \underbrace{m_2 l_1 a_2 \cos(q_2)}_{F_{q_2}} \right] (m_2 a_2^2 + I_{2,zz}) \\
&\quad - (m_2 a_2^2 + I_{2,zz})^2 - 2 \underbrace{m_2 l_1 a_2 \cos(q_2)}_{F_{q_2}} \cdot (m_2 a_2^2 + I_{2,zz}) - F_{q_2}^2 \\
&= (m_2 a_2^2 + I_{2,zz}) \cdot \left[m_1 a_1^2 + I_{1,zz} + m_2 l_1^2 + \underbrace{I_{2,zz} + m_2 a_2^2 - m_2 a_2^2 - I_{2,zz}}_{=0} \right] - \underbrace{(m_2 l_1 a_2)^2 (1 - \sin(q_2)^2)}_{F_{q_2}^2} \\
&= (m_2 a_2^2 + I_{2,zz}) \cdot (m_1 a_1^2 + I_1 + m_2 l_1^2) - (m_2 l_1 a_2)^2 + (m_2 l_1 a_2)^2 \sin(q_2)^2 \\
&= (m_2 a_2^2 + I_{2,zz}) \cdot (m_1 a_1^2 + I_1) + \underbrace{(m_2 a_2^2 + I_{2,zz}) \cdot m_2 l_1^2}_{m_2^2 l_1^2 a_2^2 + m_2 l_1^2 I_{2,zz}} - (m_2 l_1 a_2)^2 + (m_2 l_1 a_2)^2 \sin(q_2)^2 \\
&= (m_1 a_1^2 + I_1 + m_2 l_1^2) \cdot (m_2 a_2^2 + I_2) + m_2 l_1^2 I_{2,zz} + m_2^2 l_1^2 a_2^2 \sin(q_2)^2
\end{aligned}$$

Für die Inverse der Massenmatrix gilt somit

$$\begin{aligned}
M^{-1} &= \frac{1}{\det(M)} \begin{pmatrix} M_{22} & -M_{12} \\ -M_{21} & M_{11} \end{pmatrix} \\
&= \frac{1}{\det(M)} \begin{pmatrix} m_2 a_2^2 + I_2 & -(m_2 \cdot (a_2^2 + l_1 a_2 \cos(x_2)) + I_2) \\ -(m_2 \cdot (a_2^2 + l_1 a_2 \cos(x_2)) + I_2) & m_1 a_1^2 + I_1 + m_2 \cdot (l_1^2 + a_2^2 + 2 l_1 a_2 \cos(x_2)) + I_2 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
M_{11} &= m_1 a_1^2 + I_1 + m_2 \cdot (l_1^2 + a_2^2 + 2 l_1 a_2 \cos(x_2)) + I_2 \\
M_{12} &= m_2 \cdot (a_2^2 + l_1 a_2 \cos(x_2)) + I_2 \\
M_{21} &= M_{12} \\
M_{22} &= m_2 a_2^2 + I_2
\end{aligned}$$

Ursprungsproblem:

$$\min_u \int_0^\infty (\boldsymbol{q}-\boldsymbol{q}_f)^T Q(\boldsymbol{q}-\boldsymbol{q}_f) + (\boldsymbol{u}-\boldsymbol{u}_f)^T R(\boldsymbol{u}-\boldsymbol{u}_f)\,\mathrm{d}t$$

s.t. $M(\boldsymbol{q})\ddot{\boldsymbol{q}} + C(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + G(\boldsymbol{q}) = \boldsymbol{u}; \quad \boldsymbol{u}_f - G(\boldsymbol{q}_f) = \boldsymbol{0}$

Berechnung von $M^{-1}(\boldsymbol{q}_f)$ und $\frac{\partial G}{\partial \boldsymbol{q}}(\boldsymbol{q}_f)$:

$$A = \begin{bmatrix} \boldsymbol{0}_2 & I_2 \\ M^{-1}\frac{\partial G}{\partial \boldsymbol{q}} & \boldsymbol{0}_2 \end{bmatrix}, \quad B = \begin{bmatrix} \boldsymbol{0} \\ M^{-1} \end{bmatrix}$$

LQR mit $\bar{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{x}_f, \bar{\boldsymbol{u}} = \boldsymbol{u} - \boldsymbol{u}_f$

$$\min_{\bar{\boldsymbol{u}}} \int_0^\infty \bar{\boldsymbol{x}}^T Q \bar{\boldsymbol{x}} + \bar{\boldsymbol{u}}^T R \bar{\boldsymbol{u}}\,\mathrm{d}t \quad \dot{\bar{\boldsymbol{x}}} = A\bar{\boldsymbol{x}} + B\bar{\boldsymbol{u}};$$

Lösung der Riccati-Gleichung für $K \in \mathbb{R}^{2\times 2}$

$$Q - KBR^{-1}B^TK + KA + A^TK = 0$$

Closed-Loop Regelung:

$$\boldsymbol{u} = \boldsymbol{u}_f + \bar{\boldsymbol{u}} = G(\boldsymbol{q}_f) - R^{-1}B^TK(\boldsymbol{x} - \boldsymbol{x}_f)$$

Lösung Originalsystem

$$\boldsymbol{x}_0 = \boldsymbol{x}(0)$$
$$\boldsymbol{u}_i = G(\boldsymbol{q}_f) - R^{-1}B^TK(\boldsymbol{x}_i - \boldsymbol{x}_f)$$
$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + h\cdot f(\boldsymbol{x}_i, \boldsymbol{u}_i)$$

3.2 Riccati Regler

Es gilt

$$F(K) := Q - KPK + KA + A^TK \Rightarrow F(K) = 0$$

Definition 3.2 *Sei $X \in \mathbb{R}^{n\times n}$ dann definieren wir*

$$\mathsf{vec}(X) = \begin{pmatrix} X_{11} & X_{21} & \cdots & X_{n1} & X_{12} & X_{22} & \cdots X_{nn} \end{pmatrix}^T \in \mathbb{R}^{n^2}$$

Seien $A \in \mathbb{R}^{n\times m}, B \in \mathbb{R}^{p\times r}$, dann definieren wir

$$A \otimes B = \begin{pmatrix} A_{11}B & \cdots & A_{1m}B \\ A_{12}B & \dots & A_{2m}B \\ \vdots & \ddots & \vdots \\ A_{n1}B & \cdots & A_{nm}B \end{pmatrix} \in \mathbb{R}^{mp\times nr}$$

Es gilt die Formel

$$\mathsf{vec}(MXN) = (N^T \otimes M)\mathsf{vec}(X)$$

Betrachten wir den Fall $M = 1$ und $N = A$ bzw. $N = 1$ und $M = A^T$ erhalten wir mit $X = K$

$$\mathsf{vec}(KA) = (A^T \otimes 1)\mathsf{vec}(K) \quad \mathsf{vec}(A^TK) = (1 \otimes A^T)\mathsf{vec}(K)$$

Anwendung auf unser Problem liefert

$$\underbrace{\mathsf{vec}(F(K))}_{f(x=\mathsf{vec}(K))} = \underbrace{\mathsf{vec}(Q)}_{q} - \underbrace{\mathsf{vec}(KPK)}_{g(x=\mathsf{vec}(K))} + \underbrace{\mathsf{vec}(KA)}_{(A^T\otimes 1)\mathsf{vec}(K)} + \underbrace{\mathsf{vec}(A^TK)}_{(1\otimes A^T)\mathsf{vec}(K)}$$

Durch Einführung von $x = \mathsf{vec}(K) \in \mathbb{R}^{n^2}$ erhalten wir

$$f(x) = q - g(x) + (A^T \otimes 1)x + (1 \otimes A^T)x$$

Da $(KP)^T = KP$ gilt

$$g(\mathsf{vec}(K)) = \mathsf{vec}(KPK) = (1 \otimes KP)\mathsf{vec}(K)$$

Wir betrachten die Variation $K(\epsilon) = K + \epsilon \cdot \bar{K}$ und die Funktion $h : \mathbb{R} \to \mathbb{R}^{n^2}$ mit

$$\begin{aligned} h(\epsilon) := g(\mathsf{vec}(K(\epsilon))) &= \mathsf{vec}(K(\epsilon)PK(\epsilon)) \\ &= \mathsf{vec}((K + \epsilon \cdot \bar{K})P(K + \epsilon \cdot \bar{K})) \\ &= \mathsf{vec}(KPK + \epsilon KP\bar{K} + \epsilon \bar{K}PK + \epsilon^2 \bar{K}P\bar{K}) \end{aligned}$$

Ableitung nach ϵ und Auswertung in $\epsilon = 0$ ergibt

$$h'(0) = \frac{d}{d\epsilon} g(\mathsf{vec}(K + \epsilon\bar{K}))\Big|_{\epsilon=0} = \mathsf{vec}(KP\bar{K} + \bar{K}PK) = \mathsf{vec}(KP\bar{K}) + \mathsf{vec}(\bar{K}PK)$$

Umschreiben als Matrix-Vektor-Produkt

$$\mathsf{vec}(KP \cdot \bar{K} \cdot 1) = (1 \otimes KP)\mathsf{vec}(\bar{K}) \quad \mathsf{vec}(1 \cdot \bar{K} \cdot PK) = ((PK)^T \otimes 1)\mathsf{vec}(\bar{K})$$

Wir erhalten als Ableitung

$$g'(x) = (1 \otimes KP) + ((PK)^T \otimes 1) \Rightarrow f'(x) = -(1 \otimes KP) - ((PK)^T \otimes 1) + (A^T \otimes 1) + (1 \otimes A^T)$$

3.2.1 Implementierung

```
def vec(X):
    return X.reshape(-1, order='F')

def unvec(v, n):
    return v.reshape((n, n), order='F')
```

Hilfsfunktionen für die Vektorisierung

```python
# Newton Kleinman Verfahren
def solv_CARE(A,B,R,Q,tol=1e-8,max_iter=50):
  n = A.shape[0]
  I = np.eye(n)
  q = vec(Q)
  L = np.kron(I, A.T) + np.kron(A.T, I)
  P = B @ np.linalg.solve(R, B.T)
  # Startwert K0 hier Einheitsmatrix
  K = I
  x = vec(K)

  for i in range(max_iter):
    X = K @ P #KP
    vec_KPK = np.kron(I,X) @ x
    f = q - vec_KPK + L @ x
    Dg = np.kron(I,X)+np.kron(X,I)
    Df = L - Dg
    dx = np.linalg.solve(Df, -f)
    x_new = x + dx
    # Abbruch
    if np.linalg.norm(dx) / np.linalg.norm(x_new) < tol:
      x=x_new
      break

        x = x_new
        K = unvec(x, n)
  return K
```

Erklärung des Codes
Stabilität der Startlösung hier Gershgorin kreise und Stabilität erklären
Für Y mit $RY = B^T \Rightarrow Y = B^{-1}B^T$ und $B \cdot Y = P$

4 Gradientenverfahren

Satz 4.1 (Gradientenverfahren mit Linesearch) *Sei $f : \mathbb{R}^n \to \mathbb{R}$ differenzierbar,*

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \alpha_j \cdot \nabla f(\boldsymbol{x}_k) \quad \min \left\{ j : \underbrace{f(\boldsymbol{x}_i - \alpha_j \nabla f(\boldsymbol{x}_i))}_{\varphi(\alpha_j)} \leq \underbrace{f(\boldsymbol{x}_i)}_{\varphi(0)} \underbrace{- c \cdot \alpha_j \left\| \nabla f(\boldsymbol{x}_i) \right\|^2}_{c \cdot \varphi'(0) \cdot \alpha_j,} \right\}$$

konvergiert für $k \to \infty$ gegen die Lösung von $\min_x f(x)$.

Ziel Berechnung des Gradienten $J'(\boldsymbol{u})$ für das Steuerungsproblem

$$\begin{aligned} J^{aux}(\boldsymbol{u}) &= \int_{t_0}^{t_f} L(t, \boldsymbol{x}, \boldsymbol{u}) + \boldsymbol{p}(t) \cdot \Big[f(t, \boldsymbol{x}, \boldsymbol{u}) - \dot{\boldsymbol{x}}(t) \Big] \,\mathrm{d}t + \Phi(\boldsymbol{x}(t_f)) \\ &= \int_{t_0}^{t_f} L(t, \boldsymbol{x}, \boldsymbol{u}) + \boldsymbol{p}(t) \cdot f(t, \boldsymbol{x}, \boldsymbol{u}) \,\mathrm{d}t - \int_{t_0}^{t_f} \boldsymbol{p}(t) \dot{\boldsymbol{x}}(t) \,\mathrm{d}t + \Phi(\boldsymbol{x}(t_f)) \\ &= \int_{t_0}^{t_f} H[t] \,\mathrm{d}t - \Big[\boldsymbol{p}(t) \boldsymbol{x}(t) \Big]_{t_0}^{t_f} + \int_{t_0}^{t_f} \dot{\boldsymbol{p}}(t) \boldsymbol{x}(t) \,\mathrm{d}t + \Phi(\boldsymbol{x}(t_f)) \\ &= \int_{t_0}^{t_f} H[t] + \dot{\boldsymbol{p}}(t) \boldsymbol{x}(t) \,\mathrm{d}t - \Big[\boldsymbol{p}(t) \boldsymbol{x}(t) \Big]_{t_0}^{t_f} + \Phi(\boldsymbol{x}(t_f)) \\ &= \int_{t_0}^{t_f} H[t] + \dot{\boldsymbol{p}}(t) \boldsymbol{x}(t) \,\mathrm{d}t - \boldsymbol{p}(t_f) \boldsymbol{x}(t_f) + \boldsymbol{p}(t_0) \boldsymbol{x}(t_0) + \Phi(\boldsymbol{x}(t_f)) \end{aligned}$$

Der Term $\boldsymbol{p}(t_0)\boldsymbol{x}(t_0)$ ist unabhängig von $\boldsymbol{u}$ und fällt bei der Optimierung weg. Bilden der Gateau Ableitung in $\boldsymbol{u}$ entlang $\boldsymbol{h}$ liefert

$$\begin{aligned} DJ^{aux}(\boldsymbol{u}, \boldsymbol{h}) &= \int_{t_0}^{t_f} H_x[t] \cdot S(t) + H_u[t] \cdot h(t) + \dot{\boldsymbol{p}}(t) S(t) \,\mathrm{d}t - \boldsymbol{p}(t_f) S(t_f) + \Phi_x(\boldsymbol{x}(t_f)) S(1) \\ &= \int_{t_0}^{t_f} \big(\underbrace{H_x[t] + \dot{\boldsymbol{p}}(t)}_{=0} \big) \cdot S(t) + H_u[t] \cdot h(t) \,\mathrm{d}t + \big(\underbrace{\Phi_x(\boldsymbol{x}(t_f)) - \boldsymbol{p}(t_f)}_{=0} \big) \cdot S(t_f) \\ &= \int_0^T H_u[t] \cdot h(t) \,\mathrm{d}t \end{aligned}$$

Hierbei bezeichnet $S(t)$ die Sensitivity function von $\boldsymbol{x}$ in $\boldsymbol{u}$ in Richtung $\boldsymbol{h}$, also

$$S(t) = \frac{\partial \boldsymbol{x}(t)}{\partial \boldsymbol{u}} \cdot \boldsymbol{h}$$

Wir erhalten somit $\nabla J(\boldsymbol{u}) = H_u[t]$

4.1 Gradientenverfahren für den two link revolute Manipulator

Wir betrachten das Steuerungsproblem

$$
\begin{aligned}
\min_{\boldsymbol{u}} \int_{t_0}^{t_f} L(\boldsymbol{x},\boldsymbol{u})\,\mathrm{d}t + \Phi(\boldsymbol{x}(t_f))& \\
L(\boldsymbol{x},\boldsymbol{u}) &= (\boldsymbol{u}-\boldsymbol{u}_f)^T \cdot R \cdot (\boldsymbol{u}-\boldsymbol{u}_f) \\
\Phi(\boldsymbol{x}(t_f)) &= (\boldsymbol{q}(t_f)-\boldsymbol{q}_f)^T \cdot Q \cdot (\boldsymbol{q}(t_f)-\boldsymbol{q}_f) \\
\dot{\boldsymbol{x}} = f(\boldsymbol{x},\boldsymbol{u}) &= \begin{bmatrix} \dot{\boldsymbol{q}} \\ M^{-1}(\boldsymbol{q})\cdot(-C(\boldsymbol{q},\dot{\boldsymbol{q}})\cdot\dot{\boldsymbol{q}} - G(\boldsymbol{q}) + \boldsymbol{u}) \end{bmatrix} \\
\boldsymbol{x} &= \begin{bmatrix} \boldsymbol{q} \\ \dot{\boldsymbol{q}} \end{bmatrix} \quad \boldsymbol{x}(t_0) = \begin{bmatrix} \boldsymbol{q}(t_0) \\ \dot{\boldsymbol{q}}(t_0) \end{bmatrix}
\end{aligned}
$$

Für die adjungierte Gleichung gilt

$$
\dot{\boldsymbol{p}} = -H_{\boldsymbol{x}} = -\begin{bmatrix} \frac{\partial H}{\partial \boldsymbol{q}} \\ \frac{\partial H}{\partial \dot{\boldsymbol{q}}} \end{bmatrix}, \quad \boldsymbol{p}(t_f) = \Phi_{\boldsymbol{x}}(\boldsymbol{x}(t_f)) = \begin{bmatrix} 2Q(\boldsymbol{q}(t_f)-\boldsymbol{q}_f) \\ \boldsymbol{0} \end{bmatrix} \quad H[t] = L(\boldsymbol{u}) + \boldsymbol{p}^T \cdot f(\boldsymbol{x},\boldsymbol{u})
$$

Für den Gradienten gilt

$$
\nabla J(\boldsymbol{u}) = \frac{\partial H}{\partial \boldsymbol{u}} = 2R\cdot(\boldsymbol{u}-\boldsymbol{u}_f) + \frac{\partial f}{\partial \boldsymbol{u}}^T \cdot \boldsymbol{p}
$$

Wir nutzen die spezielle Struktur unseres Problems

$$
\frac{\partial f}{\partial \boldsymbol{u}} = \begin{bmatrix} \boldsymbol{0}_2 \\ M^{-1}(\boldsymbol{q}) \end{bmatrix} \Rightarrow \frac{\partial f}{\partial \boldsymbol{u}}^T \cdot \begin{bmatrix} \boldsymbol{p}_{\boldsymbol{q}} \\ \boldsymbol{p}_{\dot{\boldsymbol{q}}} \end{bmatrix} = M^{-1}(\boldsymbol{q})\cdot \boldsymbol{p}_{\dot{\boldsymbol{q}}} \Rightarrow \nabla J(\boldsymbol{u}) = 2R\cdot(\boldsymbol{u}-\boldsymbol{u}_f) + M^{-1}(\boldsymbol{q})\cdot\boldsymbol{p}_{\dot{\boldsymbol{q}}}
$$

1) Foward Integration der State Equation: $\dot{\boldsymbol{x}} = f(\boldsymbol{x},\boldsymbol{u})$

$$
\begin{aligned}
\boldsymbol{x}_0 &= \boldsymbol{x}(t_0) \\
\boldsymbol{x}_{i+1} &= \boldsymbol{x}_i + h\cdot \begin{bmatrix} \dot{\boldsymbol{q}}_i \\ M^{-1}(\boldsymbol{q}_i)\cdot(-C(\boldsymbol{q}_i,\dot{\boldsymbol{q}}_i)\cdot\dot{\boldsymbol{q}}_i - G(\boldsymbol{q}_i) + \boldsymbol{u}_i) \end{bmatrix} \quad i = 0,\dots,N-1
\end{aligned}
$$

Abspeichern Array $\boldsymbol{X} = [\boldsymbol{x}_0, \boldsymbol{x}_1, \dots, \boldsymbol{x}_N]$

2) Backward Integration der Adjungierten Gleichung: $\dot{\boldsymbol{p}} = -H_{\boldsymbol{x}} = -\frac{\partial f}{\partial \boldsymbol{x}}$

$$
\begin{aligned}
\boldsymbol{p}_N = \boldsymbol{p}(t_f) &= \begin{bmatrix} 2Q(\boldsymbol{q}(t_f)-\boldsymbol{q}_f) \\ \boldsymbol{0} \end{bmatrix} \\
\boldsymbol{p}_{i-1} &= \boldsymbol{p}_i + h\cdot \begin{bmatrix} \boldsymbol{0}_2 & I_2 \\ \frac{\partial \ddot{\boldsymbol{q}}}{\partial \boldsymbol{q}} & \frac{\partial \ddot{\boldsymbol{q}}}{\partial \dot{\boldsymbol{q}}} \end{bmatrix} \cdot \boldsymbol{p}_i \\
\frac{\partial \ddot{\boldsymbol{q}}}{\partial \boldsymbol{q}} &= \frac{\partial}{\partial \boldsymbol{q}} \left(M^{-1}(\boldsymbol{q}_i)\cdot(-C(\boldsymbol{q}_i,\dot{\boldsymbol{q}}_i)\cdot\dot{\boldsymbol{q}}_i - G(\boldsymbol{q}_i) + \boldsymbol{u}_i) \right) \\
\frac{\partial \ddot{\boldsymbol{q}}}{\partial \boldsymbol{q}} &= \frac{\partial}{\partial \dot{\boldsymbol{q}}} \left(M^{-1}(\boldsymbol{q}_i)\cdot(-C(\boldsymbol{q}_i,\dot{\boldsymbol{q}}_i)\cdot\dot{\boldsymbol{q}}_i - G(\boldsymbol{q}_i) + \boldsymbol{u}_i) \right)
\end{aligned}
$$

2.1) Finite Differenzen für $\frac{\partial \ddot{\boldsymbol{q}}}{\partial x_j}$ $j = 1,2,3,4$

$$
\begin{aligned}
\boldsymbol{x}_{plus} = \boldsymbol{x} + \epsilon\cdot\boldsymbol{e}_j &\quad \boldsymbol{x}_{minus} = \boldsymbol{x} - \epsilon\cdot\boldsymbol{e}_j \\
\ddot{\boldsymbol{q}}(\boldsymbol{x}) &= M^{-1}(\boldsymbol{q})\cdot(-C(\boldsymbol{q},\dot{\boldsymbol{q}})\cdot\dot{\boldsymbol{q}} - G(\boldsymbol{q}) + \boldsymbol{u}) \\
\frac{\partial \ddot{\boldsymbol{q}}}{\partial x_j} &\approx \frac{\ddot{\boldsymbol{q}}(\boldsymbol{x}_{plus}) - \ddot{\boldsymbol{q}}(\boldsymbol{x}_{minus})}{2\epsilon}
\end{aligned}
$$

Abspeichern Array $\boldsymbol{P} = [\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_N]$

3) Gradientenverfahren $\boldsymbol{u}_{i+1} = \boldsymbol{u}_i - \alpha_j \cdot \nabla J(\boldsymbol{u}_i)$

$$\boldsymbol{u}_{i+1} = \boldsymbol{u}_i - \alpha_j \cdot \left[2R \cdot (\boldsymbol{u}_i - \boldsymbol{u}_f) + M^{-1}(\boldsymbol{q}_i) \cdot \boldsymbol{P}_{\dot{\boldsymbol{q}},i}\right] \quad i = 0, \ldots, N-1$$

3.1) Berechnung der Schrittweite α_j über Linesearch

$$\begin{aligned}
&\alpha_{test} = \beta^m \cdot \alpha_0 \\
&\boldsymbol{u}_{test} = \boldsymbol{u}_k - \alpha_{test} \cdot \nabla J(\boldsymbol{u}_k) \\
&J_{test} = J(\boldsymbol{u}_{test}) \\
\text{IF } &J_{test} \leq J(\boldsymbol{u}_k) - c \cdot \alpha_{test} \cdot \|\nabla J(\boldsymbol{u}_k)\|^2 : \\
&\quad \alpha_k = \alpha_{test} \\
&\quad \Rightarrow \text{Break}
\end{aligned}$$

5 Foundations of Reinforcement Learning

Wir folgen der Notation aus [3].

Definition 5.1 *Wir definieren für den Markov decision process (MDP) mit Zustandsraum $\mathcal{S}$, Aktionsraum $\mathcal{A}$, Übergangswahrscheinlichkeiten $P : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$, Belohnungsfunktion $R : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to \mathbb{R}$ und Diskontfaktor $\gamma \in [0, 1)$ die optimale Value-Funktion als*

Definition 5.2 (Value-Funktion) *Wir definieren die Value-Funktion für eine Policy π als*

$$V^\pi(\boldsymbol{s}) = \mathsf{E}\left[\sum_{i=0}^{\infty} \gamma^i \cdot r_{t+i+1} \mid \boldsymbol{s}_t = \boldsymbol{s}\right]$$

$$\begin{aligned}
V^\pi(\boldsymbol{s}) &= \mathsf{E}\left[\sum_{i=0}^{\infty} \gamma^i \cdot r_{t+i+1} \mid \boldsymbol{s}_t = \boldsymbol{s}\right] \\
&= \mathsf{E}\left[r_{t+1} \mid \boldsymbol{s}_t = \boldsymbol{s}\right] + \gamma \cdot \mathsf{E}\left[\sum_{i=0}^{\infty} \gamma^i \cdot r_{t+i+2} \mid \boldsymbol{s}_t = \boldsymbol{s}\right] \\
&= \mathsf{E}\left[r_{t+1} \mid \boldsymbol{s}_t = \boldsymbol{s}\right] + \gamma \cdot \mathsf{E}\left[\underbrace{\mathsf{E}\left[\sum_{i=0}^{\infty} \gamma^i \cdot r_{t+i+2} \mid \boldsymbol{s}_{t+1}\right]}_{V^\pi(\boldsymbol{s}_{t+1})} \mid \boldsymbol{s}_t = \boldsymbol{s}\right] \\
\Rightarrow V^\pi(\boldsymbol{s}) &= \mathsf{E}\left[r_{t+1} + \gamma \cdot V^\pi(\boldsymbol{s}_{t+1}) \mid \boldsymbol{s}_t = \boldsymbol{s}\right]
\end{aligned}$$

Denn es gilt hier für die σ-Algebren, dass $\sigma(\boldsymbol{s}_t) \subseteq \sigma(\boldsymbol{s}_{t+1})$ gilt. Explizit ausgeschrieben in Erwartungswerten ergibt sich:

$$\begin{aligned}
V^\pi(\boldsymbol{s}) &= \mathsf{E}\left[r_{t+1} + \gamma V^\pi(\boldsymbol{s}_{t+1}) \mid \boldsymbol{s}_t = \boldsymbol{s}\right] \\
&= \mathsf{E}_{a_t \sim \pi(\cdot|\boldsymbol{s})}\left[\mathsf{E}\left[r_{t+1} + \gamma V^\pi(\boldsymbol{s}_{t+1}) \mid \boldsymbol{s}_t = \boldsymbol{s}, a_t = a\right]\right] \\
&= \mathsf{E}_{a_t \sim \pi(\cdot|\boldsymbol{s})}\left[\mathsf{E}_{\boldsymbol{s}_{t+1} \sim T_a(\boldsymbol{s})}\left[R_{a_t}(\boldsymbol{s}_t, \boldsymbol{s}_{t+1}) + \gamma V^\pi(\boldsymbol{s}_{t+1})\right]\right] \\
&= \mathsf{E}_{a \sim \pi(\cdot|\boldsymbol{s})}\left[\mathsf{E}_{\boldsymbol{s}' \sim T_a(\boldsymbol{s})}\left[R_a(\boldsymbol{s}, \boldsymbol{s}') + \gamma V^\pi(\boldsymbol{s}')\right]\right]
\end{aligned}$$

Definition 5.3 (Q-Function) *Wir definieren die Q-Function als*

$$Q^\pi(\boldsymbol{s}, a) = \mathsf{E}\left[\sum_{i=0}^{\infty} \gamma^i \cdot r_{t+i+1} \mid \boldsymbol{s}_t = \boldsymbol{s}, a_t = a\right]$$

Herleitung der Bellmann Gleichung für die Q-Funktion:

$$\begin{aligned}
Q^\pi(\boldsymbol{s}, a) &= \mathsf{E}\left[\sum_{i=0}^{\infty} \gamma^i \cdot r_{t+i+1} \mid \boldsymbol{s}_t = \boldsymbol{s}, a_t = a\right] \\
&= \mathsf{E}\left[\mathsf{E}\left[r_{t+1} + \sum_{i=1}^{\infty} \gamma^i \cdot r_{t+i+1} \mid \boldsymbol{s}_t = \boldsymbol{s}, a_t = a, \boldsymbol{s}_{t+1} = \boldsymbol{s}'\right] \mid \boldsymbol{s}_t = \boldsymbol{s}, a_t = a\right] \\
&= \mathsf{E}\left[r_{t+1} + \gamma \cdot \underbrace{\mathsf{E}\left[\sum_{i=0}^{\infty} \gamma^i \cdot r_{t+i+2} \mid \boldsymbol{s}_{t+1} = \boldsymbol{s}'\right]}_{V^\pi(\boldsymbol{s}_{t+1})} \mid \boldsymbol{s}_t = \boldsymbol{s}, a_t = a\right] \\
&= \mathsf{E}_{\boldsymbol{s}' \sim T_a(\boldsymbol{s})}\left[R_a(\boldsymbol{s}, \boldsymbol{s}') + \gamma V^\pi(\boldsymbol{s}')\right] \\
&= \mathsf{E}_{\boldsymbol{s}' \sim T_a(\boldsymbol{s})}\left[R_a(\boldsymbol{s}, \boldsymbol{s}') + \gamma \mathsf{E}_{a' \sim \pi(\cdot|\boldsymbol{s}')}\left[Q^\pi(\boldsymbol{s}', a')\right]\right]
\end{aligned}$$

Wobei wir im letzten Schritt die folgende Beziehung genutzt haben:

$$V^\pi(\boldsymbol{s}) = \mathsf{E}_{a \sim \pi(\cdot|\boldsymbol{s})}\left[Q^\pi(\boldsymbol{s}, a)\right]$$

5.1 Bellman Optimality Equations

Satz 5.4 (Optimality Equation für V) *Die optimale Value-Funktion $V^*(\boldsymbol{s})$ erfüllt:*

$$V^*(\boldsymbol{s}) = \max_a \mathsf{E}_{\boldsymbol{s}' \sim T_a(\boldsymbol{s})}\left[R_a(\boldsymbol{s}, \boldsymbol{s}') + \gamma V^*(\boldsymbol{s}')\right] \qquad \pi^*(\boldsymbol{s}) = \arg\max_a \mathsf{E}_{\boldsymbol{s}' \sim T_a(\boldsymbol{s})}\left[R_a(\boldsymbol{s}, \boldsymbol{s}') + \gamma V^*(\boldsymbol{s}')\right]$$

Beweis: Wir setzen $V^*(\boldsymbol{s}) = \max_\pi V^\pi(\boldsymbol{s})$. Für jede Policy π gilt nun $V^*(\boldsymbol{s}') \geq V^\pi(\boldsymbol{s}')$ und somit:

$$V^*(\boldsymbol{s}) = \max_\pi \mathsf{E}_{a \sim \pi(\cdot|\boldsymbol{s})}\left[\mathsf{E}_{\boldsymbol{s}' \sim T_a(\boldsymbol{s})}\left[R_a(\boldsymbol{s}, \boldsymbol{s}') + \gamma V^\pi(\boldsymbol{s}')\right]\right] \leq \max_\pi \mathsf{E}_{a \sim \pi(\cdot|\boldsymbol{s})}\left[\mathsf{E}_{\boldsymbol{s}' \sim T_a(\boldsymbol{s})}\left[R_a(\boldsymbol{s}, \boldsymbol{s}') + \gamma V^*(\boldsymbol{s}')\right]\right]$$

Das Maximum über alle Policies π wird erreicht durch die Policy, die in jedem Zustand $\boldsymbol{s}$ die Aktion a auswählt, die das Maximum in der inneren Erwartung erreicht. Somit gilt:

$$V^*(\boldsymbol{s}) \leq \max_a \mathsf{E}_{\boldsymbol{s}' \sim T_a(\boldsymbol{s})}\left[R_a(\boldsymbol{s}, \boldsymbol{s}') + \gamma V^*(\boldsymbol{s}')\right]$$

Wir konstruieren nun die greedy Policy π^* mittels:

$$\pi^*(\boldsymbol{s}) = \arg\max_a \mathsf{E}_{\boldsymbol{s}' \sim T_a(\boldsymbol{s})}\left[R_a(\boldsymbol{s}, \boldsymbol{s}') + \gamma V^*(\boldsymbol{s}')\right]$$

Für diese Policy π^* gilt nun:

$$V^{\pi^*}(\boldsymbol{s}) = \max_a \mathsf{E}_{\boldsymbol{s}' \sim T_a(\boldsymbol{s})}\left[R_a(\boldsymbol{s}, \boldsymbol{s}') + \gamma V^*(\boldsymbol{s}')\right] \geq V^*(\boldsymbol{s}) \Rightarrow V^{\pi^*}(\boldsymbol{s}) = V^*(\boldsymbol{s})$$

□

Satz 5.5 (Optimality Equation für Q-Funktion) *Die optimale Q-Funktion $Q^*(\boldsymbol{s}, a)$ erfüllt:*

$$Q^*(\boldsymbol{s}, a) = \mathsf{E}_{\boldsymbol{s}' \sim T_a(\boldsymbol{s})} \left[R_a(\boldsymbol{s}, \boldsymbol{s}') + \gamma \max_{a'} Q^*(\boldsymbol{s}', a') \right] \qquad \pi^*(\boldsymbol{s}) = \arg\max_a Q^*(\boldsymbol{s}, a)$$

Beweis: Für die optimale Q-Funktion gilt somit

$$Q^*(\boldsymbol{s}, a) = \max_{\pi} Q^{\pi}(\boldsymbol{s}, a) \Rightarrow Q^*(\boldsymbol{s}, a) = \mathsf{E}_{\boldsymbol{s}' \sim T_a(\boldsymbol{s})} \left[R_a(\boldsymbol{s}, \boldsymbol{s}') + \gamma \max_{a'} Q^*(\boldsymbol{s}', a') \right]$$

Für die optimale Policy gilt:

$$\pi^*(\boldsymbol{s}) = a^* = \arg\max_a Q^*(\boldsymbol{s}, a)$$

□

$$\tau_t^n := \{\boldsymbol{s}_t, a_t, r_t, \boldsymbol{s}_{t+1}, \cdots, \boldsymbol{s}_{t+n}, a_{t+n}, r_{t+n}, \boldsymbol{s}_{t+n+1}\} = \left\{ \boldsymbol{s}_{t+i}, a_{t+i}, r_{t+i}, \boldsymbol{s}_{t+i+1} \right\}_{i=0,\dots,n}$$

Für die Wahrscheinlichkeit jeder Trajektorie gilt

$$P(\tau_0) = P(\boldsymbol{s}_0) \cdot \prod_{t=0}^{\infty} \pi(a_t | \boldsymbol{s}_t) \cdot T_{a_t}(\boldsymbol{s}_t, \boldsymbol{s}_{t+1})$$

Wir folgen der Notation aus [6] und definieren:

Definition 5.6 (Objective Function) *Wir definieren die Objective Function als*

$$J(\pi) = V^{\pi}(\boldsymbol{s}_0) = \mathsf{E}_{\tau_0 \sim P(\tau_0|\pi)} \left[R(\tau_0) \right]$$

und die optimal policy π^* als

$$\pi^* = \arg\max_{\pi} J(\pi)$$

5.2 Value Iteration

Wir definieren die Abbildung $\mathcal{F}$ als

$$\mathcal{F}[Q](\boldsymbol{s}, a) = \mathsf{E}_{\boldsymbol{s}' \sim T_a(\boldsymbol{s})} \left[R_a(\boldsymbol{s}, \boldsymbol{s}') + \gamma \max_{a'} Q(\boldsymbol{s}', a') \right]$$

Die Bellmanngleichung sagt nun aus, dass Q^* der Fixpunkt der Abbildung $\mathcal{F}$ ist, also

$$Q_{k+1} = \mathcal{F}[Q_k]$$

Algorithm 1 Value Iteration

Initialize $V_0(s)$ arbitrarily for all $s \in \mathcal{S}$
$k \leftarrow 0$
repeat
  for each $s \in \mathcal{S}$ **do**
    $V_{k+1}(s) \leftarrow \max_a \underbrace{\mathsf{E}_{s' \sim T_a(s)} \left[R_a(s, s') + \gamma V_k(s') \right]}_{Q_k(s,a)} = \max_a \sum_{s'} T_a(s, s') \left[R_a(s, s') + \gamma V_k(s') \right]$
  end for
  $k \leftarrow k + 1$
until $|V_k - V_{k-1}| < \varepsilon$
return V_k

5.3 Q-Learning

5.4 SARSA

Literatur

[1] Matthias Gerdts. *Optimal Control of ODEs and DAEs*. Berlin: De Gruyter, 2012.

[2] Amit Kumar, Shrey Kasera und L. B. Prasad. "Optimal Control of 2-Link Underactuated Robot Manipulator". In: *2017 International Conference on Innovations in Information Embedded and Communication Systems (ICIIECS)*. Gorakhpur, India, 2017.

[3] Aske Plaat. *Deep Reinforcement Learning*. Springer, 2022. ISBN: 978-981-19-0637-4.

[4] Bruno Siciliano u. a. *Robotics: Modelling, Planning and Control*. London: Springer, 2010. ISBN: 978-1849964507.

[5] Mark W. Spong, Seth Hutchinson und M. Vidyasagar. *Robot Dynamics and Control*. 2nd. Hoboken, NJ: Wiley, 2004. ISBN: 978-0471649908.

[6] Richard S. Sutton und Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 2018.