

# 4 Dynamische Optimierung

## 4.1 Grundlagen der Variationsrechnung

### 4.1.1 Problemformulierung

Im Gegensatz zu den bisher betrachteten statischen Optimierungsproblemen, bei denen die Optimierungsvariablen  $\mathbf{x}$  in einem *finite-dimensionalen Euklidischen Vektorraum*  $\mathbb{R}^n$  definiert sind, wird bei dynamischen Optimierungsaufgaben nach dem Minimum (Maximum) eines *Kostenfunktionals*  $J : \mathcal{X} \rightarrow \mathbb{R}$  bezüglich einer (reellen vektorwertigen) Funktion  $\mathbf{x}(t)$  aus einem geeigneten *Funktionsraum*  $\mathcal{X}$  gesucht. Typischerweise hat das Kostenfunktional die Form (*Lagrange Problem der Variationsrechnung*)

$$J(\mathbf{x}) = \int_{t_0}^{t_1} l(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) dt \quad (4.1)$$

oder (*Bolza Problem der Variationsrechnung*)

$$J(\mathbf{x}) = \varphi(t_0, \mathbf{x}(t_0), t_1, \mathbf{x}(t_1)) + \int_{t_0}^{t_1} l(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) dt . \quad (4.2)$$

Dabei ist  $t$  die *unabhängige Variable* (zumeist die Zeit),  $\mathbf{x}(t) = [x_1(t) \ \dots \ x_n(t)]^T$ ,  $t_0 \leq t \leq t_1$  wird häufig als *Trajektorie* bezeichnet und für festes  $t$  wird angenommen, dass gilt  $\mathbf{x}(t) \in \mathbb{R}^n$ . Im Weiteren bezeichnet  $\dot{\mathbf{x}}(t) \in \mathbb{R}^n$  die totale Ableitung von  $\mathbf{x}(t)$  bezüglich der unabhängigen Variable  $t$ , die reellwertige Funktion  $l(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  nennt man *Lagrangefunktion* oder besser *Lagrangesche Dichte* und  $\varphi(t_0, \mathbf{x}(t_0), t_1, \mathbf{x}(t_1)) : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$  beschreibt die *Randkostenfunktion* (Englisch: *terminal costs*).

Man nennt eine Trajektorie  $\mathbf{x}(t)$  *zulässig*, wenn im Intervall  $[t_0, t_1]$  sämtliche Beschränkungen eingehalten werden. Die Menge aller zulässigen Trajektorien wird im Weiteren mit  $\mathcal{X}_{\text{ad}}$  bezeichnet. Die einfachste Form solcher Beschränkungen ist, dass beide Endpunkte fixiert sind, d. h.  $\mathbf{x}(t_0) = \mathbf{x}_0$  und  $\mathbf{x}(t_1) = \mathbf{x}_1$  bzw.  $\mathcal{X}_{\text{ad}} = \{\mathbf{x}(t) \in \mathcal{X} : \mathbf{x}(t_0) = \mathbf{x}_0, \mathbf{x}(t_1) = \mathbf{x}_1\}$ . Eine weitere Möglichkeit besteht darin, dass die Trajektorie zwar an einem festen Endpunkt  $(t_0, \mathbf{x}_0)$  startet aber zum Zeitpunkt  $t_1$  auf einer vorgegebenen Kurve  $\Gamma : \mathbf{x} = \mathbf{g}(t)$ ,  $t_0 \leq t \leq T$  zu liegen kommen muss. In diesem Fall ist der optimale Wert der Endzeit  $t_1$  im Intervall  $t_0 \leq t_1 \leq T$  *frei* und die zulässige Menge lautet  $\mathcal{X}_{\text{ad}} = \{\mathbf{x}(t) \in \mathcal{X} : \mathbf{x}(t_0) = \mathbf{x}_0, \mathbf{x}(t_1) = \mathbf{g}(t_1)\}$ . Andere mögliche Beschränkungen sind so genannte *Pfadbeschränkungen* (Englisch: *path constraints*) oder *Beschränkungen der Lagrangeschen Form*

$$\psi(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) = 0 \quad \text{bzw.} \quad \psi(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) \leq 0, \quad t \in I \subset [t_0, t_1] \quad (4.3)$$

und *isoperimetrische Beschränkungen* der Form

$$\int_{t_0}^{t_1} \psi_k(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) dt = C_k, \quad k = 1, \dots, m < n . \quad (4.4)$$

Die bei der Variationsrechnung typischerweise betrachteten Funktionenräume sind die im Intervall  $[t_0, t_1]$  stetig differenzierbaren Funktionen  $C^1[t_0, t_1]$  und die stückweise stetig differenzierbaren Funktionen, im Weiteren als  $\hat{C}^1[t_0, t_1]$  bezeichnet. Elemente des Funktionenraumes  $\hat{C}^1[t_0, t_1]$  werden dabei im Folgenden auch als global stetig angenommen. Die Definition des globalen Minimums  $\mathbf{x}^*$  eines Kostenfunktional  $J(\mathbf{x})$  lässt sich ohne Angabe einer Norm direkt in der Form

$$J(\mathbf{x}^*) \leq J(\mathbf{x}), \quad \forall \mathbf{x} \in \mathcal{X}_{\text{ad}} \quad (4.5)$$

angeben. Die Beschreibung des lokalen Verhaltens in der Umgebung des Punktes  $\mathbf{x}^*$  hingegen verlangt die Definition einer Norm. Der Punkt  $\mathbf{x}^*$  ist ein lokales Minimum in  $\mathcal{X}_{\text{ad}}$  bezüglich der Norm  $\|\cdot\|$ , wenn gilt

$$\exists \gamma > 0 \text{ so, dass gilt } J(\mathbf{x}^*) \leq J(\mathbf{x}), \quad \forall \mathbf{x} \in \mathcal{X}_{\text{ad}} \cap B_\gamma(\mathbf{x}^*) \quad (4.6)$$

mit  $B_\gamma(\mathbf{x}^*) = \{\mathbf{x} \in \mathcal{X} : \|\mathbf{x} - \mathbf{x}^*\| < \gamma\}$ . Dadurch, dass in infinit-dimensionalen Vektorräumen die Normen nicht äquivalent sind, kann  $\mathbf{x}^*$  zwar bezüglich einer Norm ein lokales Minimum sein, aber bezüglich einer anderen Norm nicht. Für den Funktionenraum  $C^1[t_0, t_1]$  sind die gebräuchlichsten Normen in diesem Zusammenhang

$$\|x(t)\|_\infty := \max_{t_0 \leq t \leq t_1} |x(t)| \quad \text{und} \quad \|x(t)\|_{1,\infty} := \max_{t_0 \leq t \leq t_1} |x(t)| + \max_{t_0 \leq t \leq t_1} |\dot{x}(t)|. \quad (4.7)$$

Wenn  $\mathbf{x}(t)$  vektorwertig ist und im Funktionenraum  $(C^1[t_0, t_1])^n$  definiert ist, dann lauten die zugehörigen Normen

$$\|\mathbf{x}(t)\|_\infty := \max_{t_0 \leq t \leq t_1} \|\mathbf{x}(t)\| \quad \text{und} \quad \|\mathbf{x}(t)\|_{1,\infty} := \max_{t_0 \leq t \leq t_1} \|\mathbf{x}(t)\| + \max_{t_0 \leq t \leq t_1} \|\dot{\mathbf{x}}(t)\|, \quad (4.8)$$

wobei  $\|\mathbf{x}(t)\|$  eine Norm im finit-dimensionalen Vektorraum  $\mathbb{R}^n$  beschreibt.

### 4.1.2 Optimalitätsbedingungen

Zur Herleitung der notwendigen Optimalitätsbedingungen benötigt man den Begriff der *Variation eines Funktionals*.

**Definition 4.1** (Variation eines Funktionals, Gâteaux Ableitung). Die erste Variation des Funktionals  $J(\mathbf{x})$  am Punkt  $\mathbf{x} \in \mathcal{X}$  in Richtung  $\boldsymbol{\xi} \in \mathcal{X}$ , auch als die Gâteaux Ableitung von  $J(\mathbf{x})$  bezüglich  $\boldsymbol{\xi}$  am Punkt  $\mathbf{x}$  bezeichnet, ist in der Form

$$\delta J(\mathbf{x}; \boldsymbol{\xi}) := \lim_{\eta \rightarrow 0} \frac{J(\mathbf{x} + \eta \boldsymbol{\xi}) - J(\mathbf{x})}{\eta} = \left. \frac{d}{d\eta} J(\mathbf{x} + \eta \boldsymbol{\xi}) \right|_{\eta=0} \quad (4.9)$$

definiert. Falls  $\delta J(\mathbf{x}; \boldsymbol{\xi})$  für alle  $\boldsymbol{\xi} \in \mathcal{X}$  definiert ist, dann nennt man  $J(\mathbf{x})$  Gâteaux differenzierbar am Punkt  $\mathbf{x}$ .

Offensichtlich muss für die Existenz der Gâteaux Ableitung das Funktional  $J(\mathbf{x})$  definiert sein und die Ableitung von  $J(\mathbf{x} + \eta \boldsymbol{\xi})$  bezüglich  $\eta$  an der Stelle  $\eta = 0$  existieren.

**Beispiel 4.1.** Die Gâteaux Ableitung des Funktionals  $J(x) = \int_{t_0}^{t_1} x^2(t) dt$ ,  $x \in C^1[t_0, t_1]$  lautet

$$\delta J(x; \xi) = \lim_{\eta \rightarrow 0} \left( \int_{t_0}^{t_1} 2x(t)\xi(t) dt + \eta \int_{t_0}^{t_1} \xi^2(t) dt \right) = 2 \int_{t_0}^{t_1} x(t)\xi(t) dt \quad (4.10)$$

für alle  $\xi \in C^1[t_0, t_1]$ , weshalb  $J(x)$  an jedem Punkt  $x \in C^1[t_0, t_1]$  Gâteaux differenzierbar ist.

**Beispiel 4.2.** Man betrachte das Funktional  $J(x) = \int_0^1 |x(t)| dt$ ,  $x \in C^1[0, 1]$ , welches für jedes  $x \in C^1[0, 1]$  im endlichen Intervall  $[0, 1]$  einen finiten Wert liefert. Für  $x_0(t) = 0$  und  $\xi_0(t) = t$  lautet die Gâteaux Ableitung (4.9)

$$\delta J(x_0; \xi_0) = \lim_{\eta \rightarrow 0} \frac{1}{\eta} \left( \int_0^1 |x_0 + \eta \xi_0| dt - \int_0^1 |x_0| dt \right) = \quad (4.11a)$$

$$= \lim_{\eta \rightarrow 0} \operatorname{sgn}(\eta) \int_0^1 |t| dt = \begin{cases} \frac{1}{2}, & \eta \rightarrow +0 \\ -\frac{1}{2}, & \eta \rightarrow -0. \end{cases} \quad (4.11b)$$

Dabei erkennt man, dass in Richtung  $\xi_0 = t$  an der Stelle  $x_0 = 0$  die Gâteaux Ableitung nicht existiert.

Die Gâteaux Ableitung ist eine *lineare Operation*, weshalb gilt

$$\delta(J_1 + J_2)(\mathbf{x}; \boldsymbol{\xi}) = \delta J_1(\mathbf{x}; \boldsymbol{\xi}) + \delta J_2(\mathbf{x}; \boldsymbol{\xi}) \quad (4.12)$$

und für jedes reelle  $\alpha$  gilt die Beziehung

$$\delta J(\mathbf{x}; \alpha \boldsymbol{\xi}) = \alpha \delta J(\mathbf{x}; \boldsymbol{\xi}) . \quad (4.13)$$

Man beachte an dieser Stelle, dass eine Richtung  $\boldsymbol{\xi} \in \mathcal{X}_{\text{ad}}$ , für die gilt  $\delta J(\bar{\mathbf{x}}; \boldsymbol{\xi}) < 0$ , eine *Abstiegsrichtung* des Funktionals  $J$  am Punkt  $\bar{\mathbf{x}}$  bezeichnet und damit eine Generalisierung der Abstiegsrichtung  $\mathbf{d}$  vom finit-dimensionalen Fall mit  $\mathbf{d}^T(\nabla f)(\bar{\mathbf{x}}) < 0$  der Kostenfunktion  $f(\mathbf{x})$  am Punkt  $\bar{\mathbf{x}}$  gemäß Satz 2.1 beschreibt. Es gilt nun folgender Hilfssatz.

**Lemma 4.1 (Ausschluss eines Minimums).** Wenn  $J$  ein Funktional in einem normierten linearen Vektorraum  $(\mathcal{X}, \|\cdot\|)$  beschreibt und an einem Punkt  $\bar{\mathbf{x}} \in \mathcal{X}$  eine Richtung  $\boldsymbol{\xi} \in \mathcal{X}$  so existiert, dass gilt  $\delta J(\bar{\mathbf{x}}; \boldsymbol{\xi}) < 0$ , dann kann  $\bar{\mathbf{x}}$  kein lokales Minimum sein.

*Beweisskizze:* Gemäß Definition 4.1 existiert ein  $\gamma > 0$  so, dass gilt

$$\delta J(\bar{\mathbf{x}}; \boldsymbol{\xi}) = \lim_{\eta \rightarrow 0} \frac{J(\bar{\mathbf{x}} + \eta \boldsymbol{\xi}) - J(\bar{\mathbf{x}})}{\eta} < 0 \quad (4.14)$$

bzw.

$$J(\bar{\mathbf{x}} + \eta \boldsymbol{\xi}) < J(\bar{\mathbf{x}}), \quad \forall \eta \in (0, \gamma) . \quad (4.15)$$

Da  $\|\bar{\mathbf{x}} + \eta \boldsymbol{\xi} - \bar{\mathbf{x}}\| = \|\eta \boldsymbol{\xi}\| \rightarrow 0$  für  $\eta \rightarrow 0$ , ist zu erwarten, dass unabhängig von

der gewählten Norm  $\|\cdot\|$  Punkte  $\bar{\mathbf{x}} + \eta\xi$  in der Umgebung von  $\bar{\mathbf{x}}$  zu liegen kommen, weshalb in Richtung  $\xi$  am Punkt  $\bar{\mathbf{x}}$  das Funktional weiter verkleinert werden kann und damit  $\bar{\mathbf{x}}$  kein lokales Minimum beschreibt.  $\square$

Die notwendigen Bedingungen erster Ordnung für ein lokales Minimum eines Funktionals lassen sich nun wie folgt formulieren.

**Satz 4.1 (Notwendige Bedingungen erster Ordnung).** *Angenommen  $\mathbf{x}^* \in \mathcal{X}_{\text{ad}}$  ist ein (lokales) Minimum des Funktionals  $J$ , welches in einer Teilmenge  $\mathcal{X}_{\text{ad}}$  eines normierten linearen Vektorraums  $(\mathcal{X}, \|\cdot\|)$  definiert ist. Dann gilt*

$$\delta J(\mathbf{x}^*; \xi) = 0 \quad (4.16)$$

für alle zulässigen Richtungen  $\xi$  an der Stelle  $\mathbf{x}^*$ . Man nennt dabei eine Richtung  $\xi \in \mathcal{X}, \xi \neq \mathbf{0}$  an der Stelle  $\mathbf{x}^*$  zulässig, wenn gilt

(a)  $\delta J(\mathbf{x}^*; \xi)$  existiert und

(b)  $\mathbf{x}^* + \eta\xi \in \mathcal{X}_{\text{ad}}$  für alle  $\eta \in (-\varepsilon, \varepsilon)$  und hinreichend kleines  $\varepsilon > 0$ .

Bedingung (b) impliziert, dass  $\mathbf{x}^*$  im Inneren von  $\mathcal{X}_{\text{ad}}$  zu liegen kommt.

Im nächsten Schritt betrachte man das Lagrange Problem der Variationsrechnung gemäß (4.1) mit festem Anfangs- und Endpunkt.

**Satz 4.2 (Euler-Lagrange Gleichungen).** *Gegeben sei das Funktional*

$$J(\mathbf{x}) = \int_{t_0}^{t_1} l(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) dt \quad (4.17)$$

mit der zulässigen Menge  $\mathcal{X}_{\text{ad}} = \{\mathbf{x}(t) \in (C^1[t_0, t_1])^n : \mathbf{x}(t_0) = \mathbf{x}_0, \mathbf{x}(t_1) = \mathbf{x}_1\}$  und der stetig differenzierbaren Lagrangeschen Dichte  $l : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ . Wenn  $\mathbf{x}^*(t)$  ein (lokales) Minimum von  $J(\mathbf{x})$  auf  $\mathcal{X}_{\text{ad}}$  bezeichnet, dann erfüllt  $\mathbf{x}^*(t)$  die Euler-Lagrange Gleichungen

$$\frac{d}{dt} \left( \frac{\partial}{\partial \dot{x}_i} l \right) (t, \mathbf{x}^*(t), \dot{\mathbf{x}}^*(t)) - \left( \frac{\partial}{\partial x_i} l \right) (t, \mathbf{x}^*(t), \dot{\mathbf{x}}^*(t)) = 0 \quad (4.18)$$

für alle  $t \in [t_0, t_1]$  und  $i = 1, \dots, n$ .

*Beweis.* Da  $\mathbf{x}^*$  ein Minimum ist, muss wegen Satz 4.1 gelten

$$\begin{aligned} \delta J(\mathbf{x}^*; \xi) &= \left. \frac{\partial}{\partial \eta} J(\mathbf{x}^* + \eta\xi) \right|_{\eta=0} = \int_{t_0}^{t_1} \left. \frac{\partial}{\partial \eta} l(t, \mathbf{x}^*(t) + \eta\xi(t), \dot{\mathbf{x}}^*(t) + \eta\dot{\xi}(t)) \right|_{\eta=0} dt \\ &= \int_{t_0}^{t_1} \left[ \left( \frac{\partial}{\partial \mathbf{x}} l \right) (t, \mathbf{x}^*(t), \dot{\mathbf{x}}^*(t)) \xi + \left( \frac{\partial}{\partial \dot{\mathbf{x}}} l \right) (t, \mathbf{x}^*(t), \dot{\mathbf{x}}^*(t)) \dot{\xi} \right] dt = 0. \end{aligned} \quad (4.19)$$

Wegen der stetigen Differenzierbarkeit der Lagrangeschen Dichte  $l$  und da  $\boldsymbol{\xi} \in (C^1[t_0, t_1])^n$  ist der Integrand von (4.19) im Intervall  $[t_0, t_1]$  stetig und daher ist das Funktional  $J(\mathbf{x})$  an allen Punkten  $\mathbf{x} \in (C^1[t_0, t_1])^n$  Gâteaux differenzierbar. Eine nach Satz 4.1 zulässige Richtung  $\boldsymbol{\xi}$  muss die Bedingungen  $\boldsymbol{\xi}(t_0) = \mathbf{0}$  und  $\boldsymbol{\xi}(t_1) = \mathbf{0}$  erfüllen. Führt man für den zweiten Summanden in der zweiten Zeile von (4.19) eine partielle Integration durch, so erhält man

$$\int_{t_0}^{t_1} \left( \frac{\partial}{\partial \mathbf{x}} l \right) \boldsymbol{\xi} + \left( \frac{\partial}{\partial \dot{\mathbf{x}}} l \right) \dot{\boldsymbol{\xi}} dt = \int_{t_0}^{t_1} \left[ \left( \frac{\partial}{\partial \mathbf{x}} l \right) - \frac{d}{dt} \left( \frac{\partial}{\partial \dot{\mathbf{x}}} l \right) \right] \boldsymbol{\xi} dt + \underbrace{\left[ \left( \frac{\partial}{\partial \dot{\mathbf{x}}} l \right) \boldsymbol{\xi} \right]_{t_0}^{t_1}}_{=0} = 0. \quad (4.20)$$

Wählt man nun für ein festes  $i = 1, \dots, n$  die Richtung  $\boldsymbol{\xi} = [\xi_1 \ \dots \ \xi_n]^T \in (C^1[t_0, t_1])^n$  so, dass gilt  $\xi_j = 0$  für  $j \neq i$  und  $\xi_i(t_0) = \xi_i(t_1) = 0$ , dann ergibt sich

$$\int_{t_0}^{t_1} \left[ \left( \frac{\partial}{\partial x_i} l \right) - \frac{d}{dt} \left( \frac{\partial}{\partial \dot{x}_i} l \right) \right] \xi_i dt = 0 \quad (4.21)$$

und nach dem *Fundamentallemma der Variationsrechnung* folgt unmittelbar das Ergebnis (4.18).  $\square$

**Lemma 4.2 (Fundamentallemma der Variationsrechnung).** *Angenommen  $g(t)$  ist eine stückweise stetige Funktion auf dem Intervall  $[t_0, t_1]$  und es gilt*

$$\int_{t_0}^{t_1} g(t) \xi_i(t) dt = 0 \quad (4.22)$$

*für alle stückweise stetigen Funktionen  $\xi_i(t)$  im Intervall  $[t_0, t_1]$ , dann folgt fast überall (abgesehen von einer abzählbaren Menge von Punkten)  $g(t) = 0$ ,  $t \in [t_0, t_1]$ .*

Eine Funktion  $\bar{\mathbf{x}}(t)$ , die die Euler-Lagrange Gleichungen (4.18) erfüllt, wird auch als *stationäre Funktion der Lagrangeschen Dichte  $l$*  bezeichnet. In manchen Literaturstellen werden diese Funktionen auch als *extremale Funktionen* oder nur *Extremale* bezeichnet, obwohl es sein kann, dass sie weder ein Minimum noch ein Maximum des Kostenfunktionals beschreiben.

Die Lösung der Euler-Lagrange Gleichungen (4.18) kann für Spezialfälle mit Hilfe so genannter *erster Integrale* formuliert werden:

- (a) Die Lagrangesche Dichte hängt nicht von der unabhängigen Variablen  $t$  ab, d. h.  $l = l(\mathbf{x}, \dot{\mathbf{x}})$ . Mit der *Hamiltonfunktion*

$$H = \left( \frac{\partial}{\partial \dot{\mathbf{x}}} l \right) \dot{\mathbf{x}} - l(\mathbf{x}, \dot{\mathbf{x}}) \quad (4.23)$$

lassen sich die Euler-Lagrange Gleichungen (4.18) wie folgt

$$\frac{d}{dt}H = \frac{d}{dt}\left(\frac{\partial}{\partial \dot{\mathbf{x}}}\right)\dot{\mathbf{x}} + \left(\frac{\partial}{\partial \dot{\mathbf{x}}}\right)\ddot{\mathbf{x}} - \left(\frac{\partial}{\partial \mathbf{x}}\right)\dot{\mathbf{x}} - \left(\frac{\partial}{\partial \dot{\mathbf{x}}}\right)\ddot{\mathbf{x}} = \left[\frac{d}{dt}\left(\frac{\partial}{\partial \dot{\mathbf{x}}}\right) - \left(\frac{\partial}{\partial \mathbf{x}}\right)\right]\dot{\mathbf{x}} \quad (4.24)$$

anschreiben. Wie man aus (4.24) unmittelbar erkennt, ist die Hamiltonfunktion  $H$  entlang von stationären Funktionen konstant und bildet damit eine *Invariante* des Systems.

- (b) Die Lagrangesche Dichte hängt nicht von  $\mathbf{x}$  ab, d. h.  $l = l(t, \dot{\mathbf{x}})$ . Dann folgt aus den Euler-Lagrange Gleichungen (4.18), dass  $\frac{\partial}{\partial \dot{x}_i}l, i = 1, \dots, n$  *Invarianten* des Systems sind, denn es gilt

$$\frac{d}{dt}\left(\frac{\partial}{\partial \dot{x}_i}l\right) = 0. \quad (4.25)$$

**Aufgabe 4.1.** Nehmen Sie an, dass  $l(\mathbf{x}, \dot{\mathbf{x}})$  die Lagrangefunktion eines Starrkörpersystems ist (Skript Regelungssysteme 2) und  $\mathbf{x}$  bzw.  $\dot{\mathbf{x}}$  die generalisierten Lagekoordinaten und Geschwindigkeiten bezeichnen. Geben Sie eine physikalische Interpretation der Hamiltonfunktion  $H$  von (4.23) und der Größen  $\frac{\partial}{\partial \dot{x}_i}l, i = 1, \dots, n$  an.

**Beispiel 4.3 (Galileo's hängende Kette).** Galileo formulierte die Frage, welche Form eine an beiden Enden aufgehängte Kette im Schwerfeld der Erde besitzt. Galileo nahm dabei fälschlicherweise an, dass die Kette eine Parabel beschreiben würde. Das Problem wurde um 1690 von *Johann Bernoulli*, *Christiaan Huygens* und *Gottfried Leibniz* mit Hilfe der Variationsrechnung gelöst.

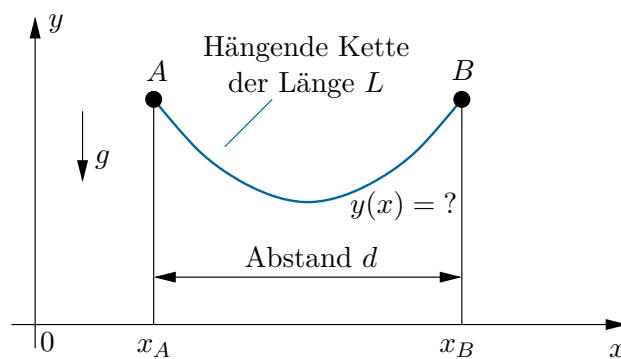


Abbildung 4.1: Galileos Problem: Bestimmung der Form  $y(x)$  einer hängenden Kette.

Betrachtet wird eine Kette der Länge  $L$ , deren Enden in den Punkten  $A$  und  $B$  aufgehängt sind, siehe Abbildung 4.1. Aus dem *Hamiltonschen Prinzip* der Mechanik folgt, dass die potentielle Energie der Kette im statischen Gleichgewicht minimal sein

muss. Mit Hilfe der Bogenlänge  $s$  und der Beziehung

$$ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + (y'(x))^2} dx \quad (4.26)$$

mit  $(\cdot)' = \frac{d}{dx}(\cdot)$  ergibt sich die potentielle Energie zu

$$V = \rho g \int_0^L y ds = \rho g \int_{x_A}^{x_B} y \sqrt{1 + (y'(x))^2} dx, \quad (4.27)$$

wobei  $g$  die Erdbeschleunigung und  $\rho$  die Linienmassendichte bezeichnet. Im Hinblick auf die Minimierung der potentiellen Energie  $V$  kann der konstante Faktor  $\rho g$  vernachlässigt werden, womit sich das Kostenfunktional und die Randbedingungen zu

$$J(y) = \int_{x_A}^{x_B} y \sqrt{1 + (y'(x))^2} dx, \quad y(x_A) = y_A, \quad y(x_B) = y_B \quad (4.28)$$

ergeben. Da in der Lagrangeschen Dichte  $l(y, y') = y \sqrt{1 + (y'(x))^2}$  die unabhängige Variable  $x$  nicht explizit auftritt, muss gemäß (4.24) die Hamiltonfunktion eine Invariante des Systems sein, d. h.

$$H = y' \left( \frac{\partial}{\partial y'} l \right) (y, y') - l(y, y') = \frac{y y'^2}{\sqrt{1 + y'^2}} - y \sqrt{1 + y'^2} = \frac{-y}{\sqrt{1 + y'^2}} = -c_1 \quad (4.29)$$

mit einer Konstanten  $c_1$ . Unter der Annahme  $c_1 \neq 0$  kann die letzte Gleichung in die Form

$$y' = \sqrt{\frac{y^2}{c_1^2} - 1} \quad (4.30)$$

gebracht werden. Unter Berücksichtigung von  $y' = \frac{d}{dx}y$  wird  $x$  zunächst mittels Integration als Funktion von  $y$  ausgedrückt

$$x = \int \frac{1}{\sqrt{\frac{y^2}{c_1^2} - 1}} dy = c_1 \ln \left( \frac{y + \sqrt{y^2 - c_1^2}}{c_1} \right) + c_2, \quad (4.31)$$

wobei  $c_2$  eine Integrationskonstante darstellt. Durch Umformung und Quadrierung,

$$\left( c_1 e^{(x-c_2)/c_1} - y \right)^2 = y^2 - c_1^2, \quad (4.32)$$

kann  $y$  in folgender Form

$$y = \frac{1}{2} c_1 \frac{e^{2(x-c_2)/c_1} + 1}{e^{(x-c_2)/c_1}} = \frac{1}{2} c_1 \left( e^{(x-c_2)/c_1} + e^{-(x-c_2)/c_1} \right). \quad (4.33)$$

geschrieben werden. Mit  $\cosh(b) = \frac{1}{2}(e^b + e^{-b})$  ergibt sich die *Kettenlinie* zu

$$y(x) = c_1 \cosh\left(\frac{x - c_2}{c_1}\right). \quad (4.34)$$

Gegebenenfalls kann noch eine dritte Konstante  $c_3$  zu (4.34) addiert werden, um die Kettenlinie für eine in den Punkten  $A$  und  $B$  aufgehängte Kette durch (numerische) Lösung der Gleichungen

$$y(x_A) = y_A, \quad y(x_B) = y_B, \quad \int_{x_A}^{x_B} \sqrt{1 + (y'(x))^2} \, dx = L \quad (4.35)$$

in Abhängigkeit der Kettenlänge  $L$  zu berechnen. Abbildung 4.2 zeigt die Kettenlinie  $y(x)$  mit der Länge  $L = 1.5$  für verschiedene Aufhängepunkte im Vergleich zu der von Galileo vermuteten Parabel.

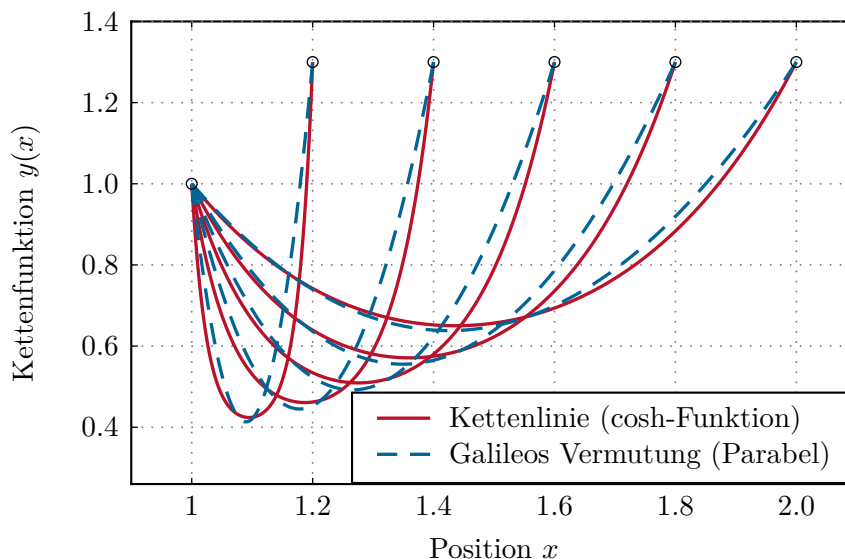


Abbildung 4.2: Hängende Kette der Länge  $L = 1.5$  für verschiedene Aufhängepunkte.

Analog zum finit-dimensionalen Fall, siehe Satz 2.2, können auch für die Minimierung von Funktionalen notwendige Bedingungen zweiter Ordnung formuliert werden.

**Satz 4.3** (Notwendige Bedingungen zweiter Ordnung - Legendre Bedingung). *Angenommen  $\mathbf{x}^* \in \mathcal{X}_{\text{ad}}$  ist ein (lokales) Minimum des Funktionals*

$$J(\mathbf{x}) = \int_{t_0}^{t_1} l(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) \, dt \quad (4.36)$$

*mit der zulässigen Menge  $\mathcal{X}_{\text{ad}} = \{\mathbf{x}(t) \in (C^1[t_0, t_1])^n : \mathbf{x}(t_0) = \mathbf{x}_0, \mathbf{x}(t_1) = \mathbf{x}_1\}$  und der zweifach stetig differenzierbaren Lagrangeschen Dichte  $l : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ , dann erfüllt  $\mathbf{x}^*$  die Euler-Lagrange Gleichungen (4.18) und die so genannte Legendre*



Bedingung

$$\left( \frac{\partial^2}{\partial \dot{\mathbf{x}}^2} l \right) (t, \mathbf{x}^*(t), \dot{\mathbf{x}}^*(t)) \text{ ist positiv semi-definit} \quad (4.37)$$

für alle  $t \in [t_0, t_1]$ .

Satz 4.2 behandelt das Lagrange Problem der Variationsrechnung (4.1). Im nächsten Schritt soll das *Bolza Problem der Variationsrechnung* (4.2) mit freier Endzeit näher untersucht werden.

**Satz 4.4 (Euler-Lagrange Gleichungen für freie Endzeit).** Gegeben sei das Funktional

$$J(t_1, \mathbf{x}) = \varphi(t_1, \mathbf{x}(t_1)) + \int_{t_0}^{t_1} l(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) dt \quad (4.38)$$

mit der zulässigen Menge  $\mathcal{X}_{\text{ad}} = \{(t_1, \mathbf{x}(t)) \in (t_0, T) \times (C^1[t_0, T])^n : \mathbf{x}(t_0) = \mathbf{x}_0\}$ , der hinreichend großen Zeit  $T \gg t_1$ , der stetig differenzierbaren Lagrangeschen Dichte  $l : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  und der stetig differenzierbaren Randkostenfunktion  $\varphi : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$ . Wenn  $(t_1^*, \mathbf{x}^*(t))$  ein (lokales) Minimum von  $J(\mathbf{x})$  auf  $\mathcal{X}_{\text{ad}}$  bezeichnet, dann erfüllt  $\mathbf{x}^*(t)$  die Euler-Lagrange Gleichungen (4.18) im Intervall  $[t_0, t_1^*]$  und es gelten die Anfangsbedingung  $\mathbf{x}^*(t_0) = \mathbf{x}_0$  sowie die Transversalitätsbedingungen

$$\left[ \frac{\partial}{\partial \dot{\mathbf{x}}} l + \frac{\partial}{\partial \mathbf{x}} \varphi \right]_{t=t_1^*, \mathbf{x}=\mathbf{x}^*} = \mathbf{0}^T \quad (4.39a)$$

$$\left[ l - \left( \frac{\partial}{\partial \dot{\mathbf{x}}} l \right) \dot{\mathbf{x}} + \frac{\partial}{\partial t} \varphi \right]_{t=t_1^*, \mathbf{x}=\mathbf{x}^*} = 0. \quad (4.39b)$$

*Beweis.* Wenn man die Endzeit  $t_1^*$  fixiert, dann folgt aus Satz 4.2 unmittelbar, dass die optimale Lösung  $\mathbf{x}^*(t)$  im Intervall  $[t_0, t_1^*]$  die Euler-Lagrange Gleichungen (4.18) erfüllt. Um die optimale Endzeit  $t_1^*$  zu berechnen, nimmt man an, dass  $\mathbf{x}(t)$  in einem hinreichend großen Intervall  $[t_0, T]$ ,  $T \gg t_1^*$  definiert ist und betrachtet den linearen Funktionenraum  $\mathbb{R} \times (C^1[t_0, T])^n$ . Die Gâteaux Ableitung (4.9) von Definition 4.1 wird dann in der Form

$$\delta J(t_1, \mathbf{x}; \tau, \boldsymbol{\xi}) := \lim_{\eta \rightarrow 0} \frac{J(t_1 + \eta\tau, \mathbf{x} + \eta\boldsymbol{\xi}) - J(t_1, \mathbf{x})}{\eta} = \left. \frac{\partial}{\partial \eta} J(t_1 + \eta\tau, \mathbf{x} + \eta\boldsymbol{\xi}) \right|_{\eta=0} \quad (4.40)$$

erweitert und die notwendige Bedingung für ein Minimum (4.16) von Satz 4.1 ausge-

wertet. Wendet man nun (4.40) auf (4.38) an, so erhält man

$$\begin{aligned}
& \frac{\partial}{\partial \eta} J(t_1^* + \eta\tau, \mathbf{x}^* + \eta\xi) = \\
& = \frac{\partial}{\partial \eta} \varphi(t_1^* + \eta\tau, \mathbf{x}^*(t_1^* + \eta\tau) + \eta\xi(t_1^* + \eta\tau)) + \frac{\partial}{\partial \eta} \int_{t_0}^{t_1^* + \eta\tau} l(t, \mathbf{x}^* + \eta\xi, \dot{\mathbf{x}}^* + \eta\dot{\xi}) dt \\
& = \left[ \left( \frac{\partial}{\partial t} \varphi \right) \tau + \frac{\partial}{\partial \mathbf{x}} \varphi \left[ \underbrace{\frac{\partial}{\partial t} \mathbf{x} + \eta \frac{\partial}{\partial t} \xi}_{\dot{\mathbf{x}}} \right] \tau + \frac{\partial}{\partial \mathbf{x}} \varphi \xi \right]_{t=t_1^* + \eta\tau, \mathbf{x}=\mathbf{x}^* + \eta\xi} \\
& \quad + \int_{t_0}^{t_1^* + \eta\tau} \frac{\partial}{\partial \eta} l(t, \mathbf{x}^* + \eta\xi, \dot{\mathbf{x}}^* + \eta\dot{\xi}) dt + \tau \left[ l(t, \mathbf{x}^* + \eta\xi, \dot{\mathbf{x}}^* + \eta\dot{\xi}) \right]_{t=t_1^* + \eta\tau} \\
& = \left[ \left( \frac{\partial}{\partial t} \varphi \right) \tau + \frac{\partial}{\partial \mathbf{x}} \varphi [\dot{\mathbf{x}} + \eta\dot{\xi}] \tau + \frac{\partial}{\partial \mathbf{x}} \varphi \xi \right]_{t=t_1^* + \eta\tau, \mathbf{x}=\mathbf{x}^* + \eta\xi} \\
& \quad + \tau \left[ l(t, \mathbf{x}^* + \eta\xi, \dot{\mathbf{x}}^* + \eta\dot{\xi}) \right]_{t=t_1^* + \eta\tau} + \left[ \left( \frac{\partial}{\partial \dot{\mathbf{x}}} l \right) (t, \mathbf{x}^* + \eta\xi, \dot{\mathbf{x}}^* + \eta\dot{\xi}) \xi \right]_{t_0}^{t_1^* + \eta\tau} \\
& \quad + \int_{t_0}^{t_1^* + \eta\tau} \left[ \left( \frac{\partial}{\partial \mathbf{x}} l \right) (t, \mathbf{x}^* + \eta\xi, \dot{\mathbf{x}}^* + \eta\dot{\xi}) - \frac{d}{dt} \left( \frac{\partial}{\partial \dot{\mathbf{x}}} l \right) (t, \mathbf{x}^* + \eta\xi, \dot{\mathbf{x}}^* + \eta\dot{\xi}) \right] \xi dt .
\end{aligned} \tag{4.41}$$

Wertet man (4.41) für  $\eta = 0$  aus, so lautet die notwendige Optimalitätsbedingung

$$\begin{aligned}
\delta J(t_1^*, \tau; \mathbf{x}^*, \xi) & = \frac{\partial}{\partial \eta} J(t_1^* + \eta\tau, \mathbf{x}^* + \eta\xi) \Big|_{\eta=0} \\
& = \left[ \left( \frac{\partial}{\partial t} \varphi \right) \tau + \frac{\partial}{\partial \mathbf{x}} \varphi (\dot{\mathbf{x}}\tau + \xi) + \tau l \right]_{t=t_1^*, \mathbf{x}=\mathbf{x}^*} + \left[ \left( \frac{\partial}{\partial \dot{\mathbf{x}}} l \right) \xi \right]_{t=t_1^*, \mathbf{x}=\mathbf{x}^*} \\
& \quad - \left[ \left( \frac{\partial}{\partial \dot{\mathbf{x}}} l \right) \xi \right]_{t=t_0, \mathbf{x}=\mathbf{x}^*} + \int_{t_0}^{t_1^*} \left[ \left( \frac{\partial}{\partial \mathbf{x}} l \right) (t, \mathbf{x}^*, \dot{\mathbf{x}}^*) - \frac{d}{dt} \left( \frac{\partial}{\partial \dot{\mathbf{x}}} l \right) (t, \mathbf{x}^*, \dot{\mathbf{x}}^*) \right] \xi dt = 0 .
\end{aligned} \tag{4.42}$$

Da der Anfangswert mit  $\mathbf{x}(t_0) = \mathbf{x}_0$  festgelegt ist, muss für eine zulässige Richtung  $\xi$  die Bedingung  $\xi(t_0) = \mathbf{0}$  gelten. Im Weiteren erfüllt die optimale Lösung  $\mathbf{x}^*(t)$  im Intervall  $[t_0, t_1^*]$  die Euler-Lagrange Gleichungen (4.18), weshalb sich (4.42) zu

$$\tau \left[ l - \left( \frac{\partial}{\partial \dot{\mathbf{x}}} l \right) \dot{\mathbf{x}} + \frac{\partial}{\partial t} \varphi \right]_{t=t_1^*, \mathbf{x}=\mathbf{x}^*} + \left[ \frac{\partial}{\partial \mathbf{x}} \varphi + \frac{\partial}{\partial \dot{\mathbf{x}}} l \right]_{t=t_1^*, \mathbf{x}=\mathbf{x}^*} (\dot{\mathbf{x}}^*(t_1^*)\tau + \xi(t_1^*)) = 0 \tag{4.43}$$

vereinfacht. Wenn die Endzeit  $t_1$  und der Endwert  $\mathbf{x}(t_1)$  frei sind, dann sind  $\tau$  und  $\xi(t_1^*)$  unabhängig voneinander frei wählbar, weshalb (4.43) nur dann Null ist, wenn die *Transversalitätsbedingungen* (4.39) erfüllt sind.  $\square$

Das Ergebnis von Satz 4.4 lässt sich nun wie folgt verallgemeinern.

- (a) Wenn die *Endzeit fest ist*, dann gilt  $t_1 = t_1^*$  und damit  $\tau = 0$ , womit automatisch der erste Term von (4.43) verschwindet. Es liegt somit keine Transversalitätsbedingung (4.39b) vor.
- (i) Wenn für eine Komponente  $x_j(t)$ ,  $j \in \{1, \dots, n\}$  von  $\mathbf{x}(t)$  gilt, dass *der Endwert*  $x_j(t_1^*) = x_j^*(t_1^*) = x_{j1}$  *fest ist*, dann muss für diese Komponente  $\xi_j(t_1^*) = 0$  gelten, womit der zugehörige Eintrag im zweiten Term von (4.43) automatisch verschwindet und keine Transversalitätsbedingung für diese Komponente vorliegt. Dieser Fall entspricht dem Ergebnis von Satz 4.2.
- (ii) Wenn für eine Komponente  $x_j(t)$ ,  $j \in \{1, \dots, n\}$  von  $\mathbf{x}(t)$  gilt, dass *der Endwert*  $x_j(t_1^*)$  *frei ist*, dann lautet, wie man aus (4.43) erkennen kann, die *Transversalitätsbedingung* für diese Komponente

$$\left[ \frac{\partial}{\partial \dot{x}_j} l + \frac{\partial}{\partial x_j} \varphi \right]_{t=t_1^*, \mathbf{x}=\mathbf{x}^*} = 0. \quad (4.44)$$

- (b) Wenn *die Endzeit frei ist*, dann muss die Transversalitätsbedingung (4.39b)

$$\left[ l - \left( \frac{\partial}{\partial \dot{\mathbf{x}}} l \right) \dot{\mathbf{x}} + \frac{\partial}{\partial t} \varphi \right]_{t=t_1^*, \mathbf{x}=\mathbf{x}^*} = 0 \quad (4.45)$$

gelten.

- (i) Wenn für eine Komponente  $x_j(t)$ ,  $j \in \{1, \dots, n\}$  von  $\mathbf{x}(t)$  gilt, dass *der Endwert*  $x_j^*(t_1^*) = x_{j1}$  *fest ist*, dann muss für diese Komponente eine zulässige Richtung  $(\tau, \xi_j)$  die Bedingung

$$x_{j1} = x_j^*(t_1^* + \eta\tau) + \eta\xi_j(t_1^* + \eta\tau) \quad (4.46)$$

bzw.

$$0 = \frac{\partial}{\partial \eta} x_{j1} \Big|_{\eta=0} = \xi_j(t_1^*) + \tau \dot{x}_j^*(t_1^*) \quad (4.47)$$

erfüllen. Damit verschwindet der zugehörige Eintrag im zweiten Term von (4.43) und es liegt keine weitere Transversalitätsbedingung für diese Komponente vor.

- (ii) Wenn für eine Komponente  $x_j(t)$ ,  $j \in \{1, \dots, n\}$  von  $\mathbf{x}(t)$  gilt, dass *der Endwert*  $x_j^*(t_1^*)$  *frei ist*, dann lautet, analog zum Fall (a)(ii), die *Transversalitätsbedingung* für diese Komponente

$$\left[ \frac{\partial}{\partial \dot{x}_j} l + \frac{\partial}{\partial x_j} \varphi \right]_{t=t_1^*, \mathbf{x}=\mathbf{x}^*} = 0. \quad (4.48)$$

### 4.1.3 Stückweise stetig differenzierbare Extremale

Bei den bisherigen Betrachtungen, siehe im Speziellen Satz 4.2, 4.3 und 4.4, wurde stets angenommen, dass  $\mathbf{x}(t)$  im Funktionenraum der im Intervall  $[t_0, t_1]$  (vektorwertigen) stetig differenzierbaren Funktionen  $(C^1[t_0, T])^n$  definiert ist. Im Weiteren soll dies auf den Funktionenraum der stückweise stetig differenzierbaren Funktionen  $(\hat{C}^1[t_0, T])^n$  erweitert werden, wobei zusätzlich die globale Stetigkeit vorausgesetzt wird. Man nennt nun eine reellwertige Funktion  $x(t) \in \hat{C}^1[t_0, t_1]$  *stückweise stetig differenzierbar*, wenn eine *finite Partitionierung*  $t_0 = c_0 < c_1 < \dots < c_{N+1} = t_1$  so existiert, dass die Funktion  $x(t)$  in allen Intervallen  $(c_k, c_{k+1})$ ,  $k = 0, \dots, N$  stetig differenzierbar ist, siehe Abbildung 4.3. Die inneren Punkte  $c_1, \dots, c_N$  werden als *Eckpunkte von  $x(t)$*  bezeichnet. Für stückweise

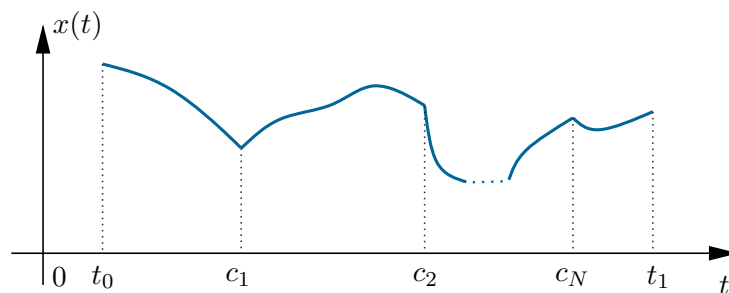


Abbildung 4.3: Beispiel einer Funktion  $x(t) \in \hat{C}^1[t_0, t_1]$ .

stetig differenzierbare Funktionen  $\hat{x}(t) \in \hat{C}^1[t_0, t_1]$  lauten die Normen gemäß (4.7)

$$\|\hat{x}(t)\|_\infty := \max_{t_0 \leq t \leq t_1} |\hat{x}(t)| \quad \text{und} \quad \|\hat{x}(t)\|_{1,\infty} := \max_{t_0 \leq t \leq t_1} |\hat{x}(t)| + \sup_{t \in \bigcup_{k=0}^N (c_k, c_{k+1})} \left| \frac{d}{dt} \hat{x}(t) \right|. \quad (4.49)$$

Es gilt nun folgender Satz.

**Satz 4.5 (Stückweise stetig vs. stetig differenzierbare Extremale).** *Angenommen  $\mathbf{x}^* \in \mathcal{X}_{\text{ad}}$  ist ein (lokales) Minimum des Funktionals*

$$J(\mathbf{x}) = \int_{t_0}^{t_1} l(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) dt \quad (4.50)$$

*mit der zulässigen Menge  $\mathcal{X}_{\text{ad}} = \{\mathbf{x}(t) \in (C^1[t_0, t_1])^n : \mathbf{x}(t_0) = \mathbf{x}_0, \mathbf{x}(t_1) = \mathbf{x}_1\}$  und der stetig differenzierbaren Lagrangeschen Dichte  $l : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ , dann ist  $\mathbf{x}^* \in \hat{\mathcal{X}}_{\text{ad}}$  auch ein (lokales) Minimum des Funktionals (4.50) in der zulässigen Menge  $\hat{\mathcal{X}}_{\text{ad}} = \{\mathbf{x}(t) \in (\hat{C}^1[t_0, t_1])^n : \mathbf{x}(t_0) = \mathbf{x}_0, \mathbf{x}(t_1) = \mathbf{x}_1\}$  bezüglich der gleichen Norm  $\|\cdot\|_\infty$  bzw.  $\|\cdot\|_{1,\infty}$ .*

Man kann nun zeigen, dass eine extremale Lösung  $\hat{\mathbf{x}}^*(t) \in (\hat{C}^1[t_0, t_1])^n$  im gesamten Intervall  $[t_0, t_1]$  außer an den Eckpunkten  $c_1, \dots, c_N$  die Euler-Lagrange Gleichungen (4.18) und die Legendre-Bedingung (4.37) erfüllt. Die Transversalitätsbedingungen (4.44), (4.45) und (4.48) bleiben im Falle stückweise stetig differenzierbarer Extremale *unverändert*. Die

Unstetigkeiten von  $\frac{d}{dt}\hat{\mathbf{x}}^*(t)$  an den Eckpunkten  $t = c_k$ ,  $k = 1, \dots, N$  unterliegen nun folgenden Einschränkungen:

**Satz 4.6 (Erste Weierstrass-Erdmann Bedingung).** Angenommen  $\hat{\mathbf{x}}^* \in \hat{\mathcal{X}}_{\text{ad}}$  ist ein (lokales) Minimum des Funktionals

$$J(\hat{\mathbf{x}}) = \int_{t_0}^{t_1} l(t, \hat{\mathbf{x}}(t), \dot{\hat{\mathbf{x}}}(t)) dt \quad (4.51)$$

mit der zulässigen Menge  $\hat{\mathcal{X}}_{\text{ad}} = \{ \hat{\mathbf{x}}(t) \in (\hat{C}^1[t_0, t_1])^n : \hat{\mathbf{x}}(t_0) = \hat{\mathbf{x}}_0, \hat{\mathbf{x}}(t_1) = \hat{\mathbf{x}}_1 \}$ , wobei die Lagrangesche Dichte  $l$  sowie die partiellen Ableitungen  $\frac{\partial}{\partial \hat{x}_i} l$  und  $\frac{\partial}{\partial \dot{\hat{x}}_i} l$  stetig auf  $[t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n$  sind. Dann gilt für jeden Eckpunkt  $c \in (t_0, t_1)$  von  $\hat{\mathbf{x}}^*(t)$ , dass die Bedingung

$$\left( \frac{\partial}{\partial \dot{\hat{\mathbf{x}}}} l \right) (c, \hat{\mathbf{x}}^*(c), \dot{\hat{\mathbf{x}}}^*(c^-)) = \left( \frac{\partial}{\partial \dot{\hat{\mathbf{x}}}} l \right) (c, \hat{\mathbf{x}}^*(c), \dot{\hat{\mathbf{x}}}^*(c^+)) \quad (4.52)$$

erfüllt ist, wobei  $\dot{\hat{\mathbf{x}}}^*(c^-)$  und  $\dot{\hat{\mathbf{x}}}^*(c^+)$  den links- bzw. rechtsseitigen Grenzwert von  $\dot{\hat{\mathbf{x}}}^*(t)$  an der Stelle  $t = c$  bezeichnen.

Die erste Weierstrass-Erdmann Bedingung besagt also, dass an den Eckpunkten einer (lokal) extremalen Trajektorie  $\hat{\mathbf{x}}^*(t) \in (\hat{C}^1[t_0, t_1])^n$  nur jene Unstetigkeiten von  $\dot{\hat{\mathbf{x}}}$  erlaubt sind, die die Stetigkeit von  $\frac{\partial}{\partial \dot{\hat{\mathbf{x}}}} l$  erhalten. Im Weiteren kann gezeigt werden, dass an den Eckpunkten auch die *Stetigkeit der Hamiltonfunktion*  $H = \left( \frac{\partial}{\partial \dot{\hat{\mathbf{x}}}} l \right) \dot{\hat{\mathbf{x}}} - l$  gewährleistet ist. Diese letztere Bedingung ist auch unter dem Namen *zweite Weierstrass-Erdmann Bedingung* bekannt.

**Satz 4.7 (Zweite Weierstrass-Erdmann Bedingung).** Angenommen  $\hat{\mathbf{x}}^* \in \hat{\mathcal{X}}_{\text{ad}}$  ist ein (lokales) Minimum des Funktionals

$$J(\hat{\mathbf{x}}) = \int_{t_0}^{t_1} l(t, \hat{\mathbf{x}}(t), \dot{\hat{\mathbf{x}}}(t)) dt \quad (4.53)$$

mit der zulässigen Menge  $\hat{\mathcal{X}}_{\text{ad}} = \{ \hat{\mathbf{x}}(t) \in (\hat{C}^1[t_0, t_1])^n : \hat{\mathbf{x}}(t_0) = \hat{\mathbf{x}}_0, \hat{\mathbf{x}}(t_1) = \hat{\mathbf{x}}_1 \}$ , wobei die Lagrangesche Dichte  $l$  sowie die partiellen Ableitungen  $\frac{\partial}{\partial \hat{x}_i} l$  und  $\frac{\partial}{\partial \dot{\hat{x}}_i} l$  stetig auf  $[t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n$  sind. Dann gilt für jeden Eckpunkt  $c \in (t_0, t_1)$  von  $\hat{\mathbf{x}}^*(t)$ , dass mit der Hamiltonfunktion

$$H(t, \hat{\mathbf{x}}(t), \dot{\hat{\mathbf{x}}}(t)) = \left( \frac{\partial}{\partial \dot{\hat{\mathbf{x}}}} l \right) (t, \hat{\mathbf{x}}(t), \dot{\hat{\mathbf{x}}}(t)) \dot{\hat{\mathbf{x}}}(t) - l(t, \hat{\mathbf{x}}(t), \dot{\hat{\mathbf{x}}}(t)) \quad (4.54)$$

die Bedingung

$$H(c, \hat{\mathbf{x}}^*(c), \dot{\hat{\mathbf{x}}}^*(c^-)) = H(c, \hat{\mathbf{x}}^*(c), \dot{\hat{\mathbf{x}}}^*(c^+)) \quad (4.55)$$

erfüllt ist, wobei  $\dot{\hat{\mathbf{x}}}^*(c^-)$  und  $\dot{\hat{\mathbf{x}}}^*(c^+)$  den links- bzw. rechtsseitigen Grenzwert von  $\dot{\hat{\mathbf{x}}}^*(t)$  an der Stelle  $t = c$  bezeichnen.

*Beispiel 4.4.* Gesucht ist ein (lokales) Minimum  $x^* \in \mathcal{X}_{\text{ad}}$  des Funktionals

$$J(x) = \int_{-1}^1 x^2(t)(1 - \dot{x}(t))^2 dt \quad (4.56)$$

in der zulässigen Menge  $\mathcal{X}_{\text{ad}} = \{x(t) \in C^1[-1, 1] : x(-1) = 0, x(1) = 1\}$ . Da die Lagrangesche Dichte nicht explizit von der Zeit  $t$  abhängt, ist die Hamiltonfunktion

$$H = \left( \frac{\partial}{\partial \dot{x}} l \right) \dot{x} - l = -2x^2(1 - \dot{x})\dot{x} - x^2(1 - \dot{x})^2 = x^2(\dot{x}^2 - 1) = -k_1 \quad (4.57)$$

für alle Zeiten  $t \in [-1, 1]$  konstant mit der Konstanten  $k_1$  und damit eine Invariante des Systems, siehe auch (4.24). Ersetzt man  $x^2(t) = z(t)$  und  $2x(t)\dot{x}(t) = \dot{z}(t)$  in (4.57), dann erhält man

$$z(t) - \frac{1}{4}\dot{z}^2(t) = k_1 \quad (4.58)$$

und die Lösung von (4.58) errechnet sich zu

$$z(t) = (t + k_2)^2 + k_1 \quad (4.59)$$

mit der Konstanten  $k_2$ . Mit  $x(-1) = 0$  und  $x(1) = 1$  sowie  $z(t) = x^2(t)$  folgen die Konstanten  $k_1$  und  $k_2$  zu  $k_1 = -\left(\frac{3}{4}\right)^2$  und  $k_2 = \frac{1}{4}$  und die mögliche stationäre Lösung  $\bar{x}(t)$  des Kostenfunktionals (4.56) lautet

$$\bar{x}(t) = \pm \sqrt{\left(t + \frac{1}{4}\right)^2 - \left(\frac{3}{4}\right)^2}. \quad (4.60)$$

Die Wurzel liefert nur für  $t \geq \frac{1}{2}$  und  $t < -1$  ein reellwertiges Ergebnis, weshalb  $\bar{x}(t)$  keine stationäre Lösung von (4.56) in der zulässigen Menge  $\mathcal{X}_{\text{ad}} = \{x(t) \in C^1[-1, 1] : x(-1) = 0, x(1) = 1\}$  darstellt.

Im nächsten Schritt soll das Kostenfunktional (4.56) in der zulässigen Menge  $\hat{\mathcal{X}}_{\text{ad}} = \{\hat{x}(t) \in \hat{C}^1[-1, 1] : \hat{x}(-1) = 0, \hat{x}(1) = 1\}$  minimiert werden. Die Weierstrass-Erdmann Bedingung (4.52) besagt nun, dass an einem Eckpunkt  $c \in (-1, 1)$  gilt

$$-2\hat{x}^2(c^-) [1 - \dot{\hat{x}}(c^-)] = -2\hat{x}^2(c^+) [1 - \dot{\hat{x}}(c^+)] \quad (4.61)$$

bzw. mit  $\hat{x}(c^-) = \hat{x}(c^+) = \hat{x}(c)$  folgt

$$\hat{x}^2(c) [\dot{\hat{x}}(c^+) - \dot{\hat{x}}(c^-)] = 0. \quad (4.62)$$

Da an einem Eckpunkt  $t = c$  gilt  $\dot{\hat{x}}(c^+) \neq \dot{\hat{x}}(c^-)$ , muss zur Erfüllung von (4.62) die Bedingung  $\hat{x}(c) = 0$  eingehalten werden. Den minimalen Wert des Kostenfunktionals (4.56) von Null erhält man offensichtlich für  $\hat{x}(t) = 0$  oder  $\dot{\hat{x}}(t) = 1$  für alle  $t$  in  $[-1, 1]$ . Da aber  $\hat{x}(1) = 1$  ist, muss man für ein möglichst großes Intervall  $(c, 1]$  mit  $\hat{x}(t) > 0$  die Bedingung  $\dot{\hat{x}}(t) = 1$  erfüllen. Eine Unstetigkeit in der Ableitung

der minimierenden Lösung kann nur an einer Stelle auftreten, bei der der Wert der Funktion selbst identisch Null ist. Man kann sich nun einfach davon überzeugen, dass die Funktion

$$\hat{x}^*(t) = \begin{cases} 0 & \text{für } -1 \leq t \leq 0 \\ t & \text{für } 0 < t \leq 1 \end{cases} \quad (4.63)$$

das eindeutige globale Minimum des Kostenfunktional (4.56) in der zulässigen Menge  $\hat{\mathcal{X}}_{\text{ad}} = \{\hat{x}(t) \in \hat{C}^1[-1, 1] : \hat{x}(-1) = 0, \hat{x}(1) = 1\}$  beschreibt.

## 4.2 Entwurf von Optimalsteuerungen

### 4.2.1 Problemformulierung

Die typische Aufgabe, die es zu lösen gilt, besteht darin, für ein dynamisches System beschrieben durch die Differenzialgleichungen

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (4.64)$$

mit der Zeit  $t \in \mathbb{R}$ , dem Zustand  $\mathbf{x} \in \mathbb{R}^n$  und dem Stelleingang  $\mathbf{u} \in \mathbb{R}^m$  eine geeignete Steuertrajektorie  $\mathbf{u}(t), t \in [t_0, t_1]$  so zu finden, dass ein Kostenfunktional der Form (siehe auch (4.2))

$$J(\mathbf{u}) = \varphi(t_0, \mathbf{x}(t_0), t_1, \mathbf{x}(t_1)) + \int_{t_0}^{t_1} l(t, \mathbf{x}(t), \mathbf{u}(t)) dt \quad (4.65)$$

bezüglich  $\mathbf{u}(t)$  minimiert wird und dabei gewisse Beschränkungen in  $\mathbf{x}(t)$  und  $\mathbf{u}(t)$  eingehalten werden. Beim Kostenfunktional unterscheidet man im Allgemeinen zwischen der Bolza-Form (4.65), der Lagrange-Form

$$J(\mathbf{u}) = \int_{t_0}^{t_1} l(t, \mathbf{x}(t), \mathbf{u}(t)) dt, \quad (4.66)$$

und der Mayer-Form

$$J(\mathbf{u}) = \varphi(t_0, \mathbf{x}(t_0), t_1, \mathbf{x}(t_1)). \quad (4.67)$$

*Aufgabe 4.2.* Zeigen Sie, dass die Lagrange-Form in die Mayer-Form überführt werden kann, indem man einen zusätzlichen Zustand

$$\dot{x}_{n+1} = l(t, \mathbf{x}, \mathbf{u}), \quad x_{n+1}(t_0) = 0 \quad (4.68)$$

eingeführt und das Kostenfunktional in der Form  $J(\mathbf{u}) = x_{n+1}(t_1)$  anschreibt.

Zeigen Sie, dass die Mayer-Form in die Lagrange-Form überführt werden kann, indem man einen zusätzlichen Zustand

$$\dot{x}_{n+1} = 0, \quad x_{n+1}(t_0) = \frac{1}{t_1 - t_0} \varphi(t_0, \mathbf{x}(t_0), t_1, \mathbf{x}(t_1)) \quad (4.69)$$

eingeführt und das Kostenfunktional in der Form  $J(\mathbf{u}) = \int_{t_0}^{t_1} x_{n+1}(t) dt$  anschreibt.

Zeigen Sie, wie man eine Bolza-Form in die Mayer- oder Lagrange-Form überführt.

Bezüglich der möglichen Beschränkungen unterscheidet man zwischen *Punktbeschränkungen*, beispielsweise *Endpunktbeschränkungen* der Form

$$\psi_1(t_1, \mathbf{x}(t_1)) \leq 0, \quad (4.70)$$

*Pfadbeschränkungen*

$$\psi(t, \mathbf{x}(t), \mathbf{u}(t)) \leq 0, \quad \forall t \in I \subset [t_0, t_1], \quad (4.71)$$

und *isoperimetrischen Beschränkungen*

$$\int_{t_0}^{t_1} \psi_k(t, \mathbf{x}(t), \mathbf{u}(t)) dt \leq C_k, \quad k = 1, \dots, \bar{n} < n. \quad (4.72)$$

Man beachte, dass Pfadbeschränkungen (4.71), die nur vom Zustand und nicht von der Stellgröße abhängen, wesentlich schwieriger zu handhaben sind.

#### 4.2.2 Existenz einer optimalen Lösung

Im Skript Regelungssysteme 2 (Satz 2.13) wurde gezeigt, dass wenn  $\mathbf{g}(t, \mathbf{x})$  stückweise stetig in  $t$  ist und der Lipschitz-Bedingung

$$\|\mathbf{g}(t, \mathbf{x}) - \mathbf{g}(t, \mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\|, \quad 0 < L < \infty \quad (4.73)$$

für alle  $\mathbf{x}, \mathbf{y} \in B_\gamma = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{x}_0\| \leq \gamma\}$  und alle  $t \in [t_0, t_0 + \tau]$  genügt, dann existiert ein  $\delta > 0$  so, dass das Anfangswertproblem

$$\dot{\mathbf{x}} = \mathbf{g}(t, \mathbf{x}), \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (4.74)$$

für  $t \in [t_0, t_0 + \delta]$  genau eine Lösung besitzt. Dadurch, dass nur stückweise Stetigkeit von  $\mathbf{g}(t, \mathbf{x})$  in  $t$  gefordert wird, darf man gemäß (4.64) mit  $\mathbf{g}(t, \mathbf{x}) := \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t))$  für die Stellgrößen  $\mathbf{u}(t)$  auch *stückweise stetige Funktionen* zulassen, d. h.  $\mathbf{u}(t) \in (\hat{C}[t_0, t_1])^m$ . Man nennt nun eine reellwertige Funktion  $u(t) \in \hat{C}[t_0, t_1]$  *stückweise stetig*, wenn eine *finite Partitionierung*  $t_0 = c_0 < c_1 < \dots < c_{N+1} = t_1$  so existiert, dass die Funktion  $u(t)$  in allen Intervallen  $(c_k, c_{k+1})$ ,  $k = 0, \dots, N$  stetig ist. Für stückweise stetige Stellgrößen  $\mathbf{u}(t)$  sind die zugehörigen Zustandsgrößen von (4.64) stückweise stetig differenzierbar, d. h.  $\mathbf{x}(t) \in (\hat{C}^1[t_0, t_1])^n$  mit den Eckpunkten an den Unstetigkeitsstellen der Stellgrößen. Zur Erinnerung sei angemerkt, dass gemäß Satz 2.14 vom Skript Regelungssysteme 2 die Stetigkeit von  $\mathbf{g}(t, \mathbf{x})$  und  $\left(\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right)(t, \mathbf{x})$  auf der Menge  $[t_0, t_0 + \delta] \times B_\gamma$  hinreichend dafür sind, dass  $\mathbf{g}(t, \mathbf{x})$  die Lipschitz-Bedingung (4.73) lokal erfüllt.

Bei den meisten praktischen Anwendungen unterliegen die Stellgrößen gewissen Beschränkungen, d. h.  $\mathbf{u}(t) \in U \subset \mathbb{R}^m$ . Typische Beschränkungen in diesem Zusammenhang sind sogenannte *box constraints* der Form

$$|u_i| \leq 1, \quad i = 1, \dots, m. \quad (4.75)$$

Eine stückweise stetige Stellgröße  $\mathbf{u}(t)$  im Intervall  $t_0 \leq t \leq t_1$  mit  $\mathbf{u}(t) \in U$  für alle  $t \in [t_0, t_1]$  bezeichnet man im Weiteren als *zulässige Stellgröße*. Für das Folgende sei angenommen, dass  $\bar{\mathbf{x}}(t; \mathbf{x}_0, \mathbf{u}(t))$  die Lösung von (4.64) zum Zeitpunkt  $t$  für den Anfangswert



$\mathbf{x}(t_0) = \mathbf{x}_0$  und die Stellgröße  $\mathbf{u}(\tau)$ ,  $t_0 \leq \tau \leq t$  bezeichnet. Dann nennt man eine zulässige Stellgröße *realisierbar*, wenn  $\bar{\mathbf{x}}(t; \mathbf{x}_0, \mathbf{u}(t))$  im gesamten Intervall  $t_0 \leq t \leq t_1$  definiert ist und sämtliche Beschränkungen einhält.

Das Problem bei der Existenz einer Lösung des Optimalsteuerungsproblems besteht häufig darin, dass die Menge der realisierbaren Lösungen *nicht kompakt* ist. Es könnte nämlich passieren, dass die Lösung von (4.64) innerhalb des Optimierungsintervalls  $[t_0, t_1]$  nach Unendlich strebt und somit das Kostenfunktional unendlich wird. Dieses Phänomen ist auch unter dem Namen *finite escape time* bekannt. Um dies zu verhindern, fordert man oft a priori, dass die Lösungen des dynamischen Systems (4.64) beschränkt sind, also dass gilt

$$\|\bar{\mathbf{x}}(t; \mathbf{x}_0, \mathbf{u}(t))\| \leq \alpha, \quad \forall t \geq t_0 \quad (4.76)$$

für ein finites  $\alpha > 0$ . Man beachte, dass bezüglich  $\mathbf{x}$  affine Systeme der Form

$$\dot{\mathbf{x}} = \mathbf{A}(t, \mathbf{u})\mathbf{x} + \mathbf{b}(t, \mathbf{u}), \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (4.77)$$

diese Eigenschaft erfüllen und in endlicher Zeit nicht nach Unendlich streben können.

Wenn das Optimierungsintervall selbst unendlich ist, dann ist die Menge der realisierbaren Stellgrößen unbeschränkt und nicht kompakt, weshalb stets ein kompaktes (finites) Zeitintervall  $[t_0, T]$  mit hinreichend großem  $T > t_1$  gewählt werden sollte. Dies wird anhand des nachfolgenden Beispiels näher erläutert.

*Beispiel 4.5.* Gegeben ist eine Punktmasse  $m$ , die über eine Kraft  $u(t)$  mit  $0 \leq u(t) \leq 1$  für alle  $t$  im Optimierungsintervall  $[t_0, t_1]$  beschleunigt wird. Die Aufgabe besteht nun darin, die Stellgröße  $u(t)$  so zu bestimmen, dass ausgehend von der Anfangsposition  $x(t_0) = x_0$  die Masse nach der Zeit  $t = t_1$  die Position  $x(t_1) = x_1$  erreicht und dabei das Kostenfunktional

$$J(u) = \int_{t_0}^{t_1} u^2(t) dt \quad (4.78)$$

minimiert. Man erkennt unmittelbar, dass für  $x_1 > x_0$  die Stellgröße  $u(t) \equiv 0$  nicht realisierbar ist und somit für jede realisierbare Stellgröße  $J(u) > 0$  gelten muss. Betrachtet man nun die Folge der realisierbaren konstanten Stellgrößen  $u_k(t) = \frac{1}{k}$ ,  $k \geq 1$  für alle  $t \geq t_0$ , so erhält man als Lösung des Differentialgleichungssystems

$$\dot{x}_k = v_k, \quad x_k(t_0) = x_0 \quad (4.79a)$$

$$m\dot{v}_k = u_k, \quad v_k(t_0) = 0 \quad (4.79b)$$

das Ergebnis

$$x_k(t) = x_0 + \frac{1}{2mk}(t - t_0)^2 \quad (4.80a)$$

$$v_k(t) = \frac{1}{mk}(t - t_0) . \quad (4.80b)$$

Die Zeit  $t_1$ , nach der der Zustand  $x_k(t_1) = x_1$  erreicht wird, errechnet sich direkt aus (4.80a) zu

$$t_{1,k} = t_0 + \sqrt{2mk} \sqrt{x_1 - x_0} . \quad (4.81)$$

Damit erhält man für  $u_k(t) = \frac{1}{k}$  im Optimierungsintervall  $[t_0, t_{1,k}]$  den Wert des Kostenfunktional zu

$$J(u_k) = \int_{t_0}^{t_{1,k}} \frac{1}{k^2} dt = \sqrt{\frac{2m}{k^3}} \sqrt{x_1 - x_0} \rightarrow 0 \quad (4.82)$$

für  $k \rightarrow \infty$  und damit  $t_{1,k} \rightarrow \infty$ . Man erkennt also, dass gilt  $\inf J(u_k) = 0$ , d. h. das Problem hat *kein Minimum*.

Damit die Existenz einer optimalen Steuerung auch tatsächlich gewährleistet ist, müssen weitere Einschränkungen der zulässigen Steuerungen vorgenommen werden. Zwei Möglichkeiten sollen im Folgenden kurz aufgezeigt werden. Einerseits besteht die Möglichkeit, zu fordern, dass die Stellgröße der *zusätzlichen Lipschitz-Bedingung*

$$\|\mathbf{u}(t) - \mathbf{u}(s)\| \leq L_u |t - s|, \quad 0 < L_u < \infty, \quad \forall s, t \in [t_0, t_1] \quad (4.83)$$

genügt und andererseits kann die Klasse der zulässigen Stellgrößen auf die *stückweise konstanten Stellgrößen* mit maximal  $r$  Unstetigkeitsstellen eingeschränkt werden.

### 4.2.3 Variationsformulierung

Im Folgenden werden die notwendigen Bedingungen erster Ordnung für ein Optimalsteuerungsproblem mit fester Endzeit und freiem Endwert formuliert und hergeleitet.

**Satz 4.8 (Steuerungsproblem: Endzeit fest/Endwert frei).** *Gesucht ist die Stellgröße  $\mathbf{u} \in (C[t_0, t_1])^m$  so, dass das Kostenfunktional (Lagrange-Form)*

$$J(\mathbf{u}) = \int_{t_0}^{t_1} l(t, \mathbf{x}(t), \mathbf{u}(t)) dt \quad (4.84)$$

*unter der Gleichungsbeschränkung (dynamisches System)*

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (4.85)$$

*für feste Zeiten  $t_0 < t_1$  minimiert wird. Dabei wird angenommen, dass  $l$  und  $\mathbf{f}$  stetig in  $(t, \mathbf{x}, \mathbf{u})$  und stetig differenzierbar bezüglich  $\mathbf{x}$  und  $\mathbf{u}$  für alle  $(t, \mathbf{x}, \mathbf{u}) \in [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m$  sind. Wenn  $\mathbf{u}^*(t) \in (C[t_0, t_1])^m$  die optimale Lösung des Optimierungsproblems bezeichnet und  $\mathbf{x}^*(t) \in (C^1[t_0, t_1])^n$  die zugehörige Lösung des Anfangswertproblems (4.85) ist, dann existiert ein  $\boldsymbol{\lambda}^*(t) \in (C^1[t_0, t_1])^n$  so, dass gilt*

$$\dot{\mathbf{x}}^* = \mathbf{f}(t, \mathbf{x}^*(t), \mathbf{u}^*(t)), \quad \mathbf{x}^*(t_0) = \mathbf{x}_0 \quad (4.86a)$$

$$\dot{\boldsymbol{\lambda}}^* = - \left( \frac{\partial}{\partial \mathbf{x}} l \right)^\top (t, \mathbf{x}^*(t), \mathbf{u}^*(t)) - \left( \frac{\partial}{\partial \mathbf{x}} \mathbf{f} \right)^\top (t, \mathbf{x}^*(t), \mathbf{u}^*(t)) \boldsymbol{\lambda}^*(t), \quad \boldsymbol{\lambda}^*(t_1) = \mathbf{0} \quad (4.86b)$$

$$\mathbf{0} = \left( \frac{\partial}{\partial \mathbf{u}} l \right)^\top (t, \mathbf{x}^*(t), \mathbf{u}^*(t)) + \left( \frac{\partial}{\partial \mathbf{u}} \mathbf{f} \right)^\top (t, \mathbf{x}^*(t), \mathbf{u}^*(t)) \boldsymbol{\lambda}^*(t) \quad (4.86c)$$

für  $t_0 \leq t \leq t_1$ . Die Gleichungen (4.86) werden als die Euler-Lagrange Gleichungen des Optimalsteuerungsproblems und  $\lambda^*(t)$  als der adjungierte Zustand oder der Kozustand bezeichnet.

*Beweis.* Man betrachte dazu die einparametrische Familie der zulässigen Stellgrößen  $\mathbf{v}(t; \eta) = \mathbf{u}^*(t) + \eta \boldsymbol{\xi}_u(t)$  mit  $\boldsymbol{\xi}_u(t) \in (C[t_0, t_1])^m$  und dem skalaren Parameter  $\eta$ . Aufgrund der Stetigkeits- und Differenzierbarkeitsannahmen für  $\mathbf{f}$  existiert ein  $\eta > 0$  so, dass die zu  $\mathbf{v}(t)$  zugehörige Lösung  $\mathbf{y}(t; \eta)$  des Anfangswertproblems (4.85) für alle  $t \in [t_0, t_1]$  eindeutig und bezüglich  $\eta$  differenzierbar ist. Für  $\eta = 0$  gilt offensichtlich  $\mathbf{y}(t; 0) = \mathbf{x}^*(t)$ ,  $t_0 \leq t \leq t_1$ . Das Kostenfunktional für  $\mathbf{v}(t; \eta)$  lautet

$$\begin{aligned} J(\mathbf{v}(t; \eta)) &= \int_{t_0}^{t_1} l(t, \mathbf{y}(t; \eta), \mathbf{v}(t; \eta)) + \boldsymbol{\lambda}^T(t) (\mathbf{f}(t, \mathbf{y}(t; \eta), \mathbf{v}(t; \eta)) - \dot{\mathbf{y}}) dt \\ &= \int_{t_0}^{t_1} l(t, \mathbf{y}(t; \eta), \mathbf{v}(t; \eta)) + \dot{\boldsymbol{\lambda}}^T \mathbf{y}(t; \eta) + \boldsymbol{\lambda}^T(t) \mathbf{f}(t, \mathbf{y}(t; \eta), \mathbf{v}(t; \eta)) dt - \left( \boldsymbol{\lambda}^T \mathbf{y} \right) \Big|_{t_0}^{t_1} \end{aligned} \quad (4.87)$$

für jedes  $\boldsymbol{\lambda}(t) \in (C^1[t_0, t_1])^n$ . Gemäß Satz 4.1 lautet die notwendige Bedingung für ein Minimum

$$\delta J(\mathbf{u}^*; \boldsymbol{\xi}_u) = \left. \frac{\partial J(\mathbf{v}(t; \eta))}{\partial \eta} \right|_{\eta=0} = 0 \quad (4.88)$$

bzw.

$$\begin{aligned} 0 &= \int_{t_0}^{t_1} \left( \frac{\partial}{\partial \mathbf{x}} l \right) (t, \mathbf{x}^*, \mathbf{u}^*) \boldsymbol{\xi}_y + \dot{\boldsymbol{\lambda}}^T(t) \boldsymbol{\xi}_y + \boldsymbol{\lambda}^T(t) \left( \frac{\partial}{\partial \mathbf{x}} \mathbf{f} \right) (t, \mathbf{x}^*, \mathbf{u}^*) \boldsymbol{\xi}_y dt + \\ & \int_{t_0}^{t_1} \left( \frac{\partial}{\partial \mathbf{u}} l \right) (t, \mathbf{x}^*, \mathbf{u}^*) \boldsymbol{\xi}_u + \boldsymbol{\lambda}^T(t) \left( \frac{\partial}{\partial \mathbf{u}} \mathbf{f} \right) (t, \mathbf{x}^*, \mathbf{u}^*) \boldsymbol{\xi}_u dt - \boldsymbol{\lambda}^T(t_1) \boldsymbol{\xi}_y(t_1) + \boldsymbol{\lambda}^T(t_0) \boldsymbol{\xi}_y(t_0) \end{aligned} \quad (4.89)$$

mit

$$\boldsymbol{\xi}_y(t) = \left( \frac{\partial}{\partial \eta} \mathbf{y} \right) (t; 0) . \quad (4.90)$$

Da die Auswirkung der Variation von  $\mathbf{u}^*(t)$  auf die Lösung in Form von  $\boldsymbol{\xi}_y(t)$  schwierig zu berechnen ist, wählt man  $\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}^*(t)$  so, dass gilt

$$\dot{\boldsymbol{\lambda}}^* = - \left( \frac{\partial}{\partial \mathbf{x}} \mathbf{f} \right)^T (t, \mathbf{x}^*(t), \mathbf{u}^*(t)) \boldsymbol{\lambda}^*(t) - \left( \frac{\partial}{\partial \mathbf{x}} l \right)^T (t, \mathbf{x}^*(t), \mathbf{u}^*(t)) . \quad (4.91)$$

Für  $\mathbf{y}(t; \eta)$  gilt näherungsweise

$$\mathbf{y}(t; \eta) \simeq \underbrace{\mathbf{y}(t; 0)}_{\mathbf{x}^*(t)} + \underbrace{\left( \frac{\partial}{\partial \eta} \mathbf{y} \right) (t; 0)}_{\boldsymbol{\xi}_y(t)} \eta \quad (4.92)$$

und mit  $\mathbf{x}^*(t_0) = \mathbf{y}(t_0; 0) = \mathbf{x}_0$  folgt  $\boldsymbol{\xi}_y(t_0) = \mathbf{0}$ . Der Punkt  $\mathbf{x}(t_1)$  hingegen ist frei, weshalb in (4.89) die Bedingung  $\boldsymbol{\lambda}^*(t_1) = \mathbf{0}$  gelten muss. Die adjungierte Differentialgleichung (4.91) mit der Endbedingung  $\boldsymbol{\lambda}^*(t_1) = \mathbf{0}$  ist *linear* und aufgrund der Differenzierbarkeitsannahmen für  $l$  und  $\mathbf{f}$  existiert die Lösung  $\boldsymbol{\lambda}^*(t)$  und ist im Intervall  $[t_0, t_1]$  eindeutig. Damit verbleibt in (4.89) der Ausdruck

$$\int_{t_0}^{t_1} \boldsymbol{\xi}_u^T \left[ \left( \frac{\partial}{\partial \mathbf{u}} l \right)^T (t, \mathbf{x}^*(t), \mathbf{u}^*(t)) + \left( \frac{\partial}{\partial \mathbf{u}} \mathbf{f} \right)^T (t, \mathbf{x}^*(t), \mathbf{u}^*(t)) \boldsymbol{\lambda}^*(t) \right] dt = 0, \quad (4.93)$$

welcher aufgrund des Fundamentallemmas der Variationsrechnung Lemma 4.2 die Bedingung

$$\left( \frac{\partial}{\partial \mathbf{u}} l \right)^T (t, \mathbf{x}^*(t), \mathbf{u}^*(t)) + \left( \frac{\partial}{\partial \mathbf{u}} \mathbf{f} \right)^T (t, \mathbf{x}^*(t), \mathbf{u}^*(t)) \boldsymbol{\lambda}^*(t) = \mathbf{0} \quad (4.94)$$

für alle  $t \in [t_0, t_1]$  impliziert.  $\square$

Wie man aus (4.86) erkennen kann, setzen sich die notwendigen Optimalitätsbedingungen für das Optimalsteuerungsproblem (4.84) und (4.85) aus  $2n$  Differentialgleichungen in  $\mathbf{x}^*$  und  $\boldsymbol{\lambda}^*$  und  $m$  algebraischen Gleichungen zusammen. Da für die Differentialgleichung in  $\mathbf{x}^*$  der Wert zum Anfangszeitpunkt  $t = t_0$  und für die Differentialgleichung in  $\boldsymbol{\lambda}^*$  der Wert zum Endzeitpunkt  $t = t_1$  gegeben ist, spricht man in diesem Zusammenhang auch von einem *Zweipunkttrandwertproblem*. Analog zu den Lagrange-Multiplikatoren von Abschnitt 3.1.2 lässt sich der adjungierte Zustand  $\boldsymbol{\lambda}(t)$  in der Form interpretieren, dass  $\boldsymbol{\lambda}(t_0)$  der *Sensitivität des Kostenfunctionals* (4.84) bezüglich einer Änderung des Anfangswertes  $\mathbf{x}_0$  entspricht.

Die Euler-Lagrange Gleichungen (4.86) lassen sich mit Hilfe der *Hamiltonfunktion*

$$H(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) = l(t, \mathbf{x}, \mathbf{u}) + \boldsymbol{\lambda}^T(t) \mathbf{f}(t, \mathbf{x}, \mathbf{u}) \quad (4.95)$$

auch in der Form

$$\dot{\mathbf{x}}^* = \left( \frac{\partial}{\partial \boldsymbol{\lambda}} H \right)^T (t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t)), \quad \mathbf{x}^*(t_0) = \mathbf{x}_0 \quad (4.96a)$$

$$\dot{\boldsymbol{\lambda}}^* = - \left( \frac{\partial}{\partial \mathbf{x}} H \right)^T (t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t)), \quad \boldsymbol{\lambda}^*(t_1) = \mathbf{0} \quad (4.96b)$$

$$\mathbf{0} = \left( \frac{\partial}{\partial \mathbf{u}} H \right)^T (t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t)) \quad (4.96c)$$

für  $t_0 \leq t \leq t_1$  anschreiben. Die letzte Bedingung (4.96c) zeigt, dass  $\mathbf{u}^*$  offensichtlich ein *stationärer Punkt* der Hamiltonfunktion  $H$  sein muss. Leitet man die Hamiltonfunktion entlang der optimalen Lösung  $(\mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t))$  nach der Zeit ab, so erhält man

$$\frac{d}{dt} H = \frac{\partial}{\partial t} H + \left( \frac{\partial}{\partial \mathbf{x}} H \right) \dot{\mathbf{x}}^* + \underbrace{\left( \frac{\partial}{\partial \mathbf{u}} H \right)}_{=\mathbf{0}} \dot{\mathbf{u}}^* + \left( \frac{\partial}{\partial \boldsymbol{\lambda}} H \right) \dot{\boldsymbol{\lambda}}^* = \frac{\partial}{\partial t} H - (\dot{\boldsymbol{\lambda}}^*)^T \mathbf{f} + (\dot{\mathbf{x}}^*)^T \dot{\boldsymbol{\lambda}}^* = \frac{\partial}{\partial t} H. \quad (4.97)$$

Wenn weder  $\mathbf{f}$  noch  $l$  explizit von der Zeit  $t$  abhängen, dann ist die Hamiltonfunktion  $H$  offensichtlich eine *Invariante* des Zweipunkttrandwertproblems (4.96). Im Weiteren muss analog zu Satz 4.3 für ein Minimum des Kostenfunktionals  $J$  die *notwendige Bedingung zweiter Ordnung* (*Legendre-Clebsch Bedingung*)

$$\left( \frac{\partial^2}{\partial \mathbf{u}^2} H \right) (t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t)) \text{ ist positiv semi-definit} \quad (4.98)$$

für alle Zeiten  $t_0 \leq t \leq t_1$  gelten.

In Satz 4.8 wurde angenommen, dass die optimale Stellgröße  $\mathbf{u}^*$  stetig ist, d. h.  $\mathbf{u}^*(t) \in (C[t_0, t_1])^m$ . Für manche Beispiele findet man keine Lösung der Euler-Lagrange Gleichungen (4.86) in der Klasse der stetigen Stellgrößen. Aus diesem Grund sucht man Extremale in der erweiterten Klasse der stückweise stetigen Stellgrößen  $(\hat{C}[t_0, t_1])^m$ . Wie bereits im Abschnitt 4.2.2 diskutiert, sind für stückweise stetige Stellgrößen  $\mathbf{u}(t)$  die zugehörigen Zustandsgrößen  $\mathbf{x}(t)$  von (4.85) stückweise stetig differenzierbar, d. h.  $\mathbf{x}(t) \in (\hat{C}^1[t_0, t_1])^n$  mit den Eckpunkten an den Unstetigkeitsstellen der Stellgrößen. Bezeichnet man mit  $\hat{\mathbf{u}}^*(t) \in (\hat{C}[t_0, t_1])^m$  die optimale Stellgröße und mit  $\hat{\mathbf{x}}^*(t)$  und  $\hat{\boldsymbol{\lambda}}^*(t)$  den zugehörigen Zustand und den adjungierten Zustand des Optimalsteuerungsproblems (4.84), (4.85), dann gelten für jeden Eckpunkt  $c \in (t_0, t_1)$  die Bedingungen

$$\hat{\mathbf{x}}^*(c^-) = \hat{\mathbf{x}}^*(c^+) \quad (4.99a)$$

$$\hat{\boldsymbol{\lambda}}^*(c^-) = \hat{\boldsymbol{\lambda}}^*(c^+) \quad (4.99b)$$

$$H(c^-, \hat{\mathbf{x}}^*(c), \hat{\mathbf{u}}^*(c^-), \hat{\boldsymbol{\lambda}}^*(c)) = H(c^+, \hat{\mathbf{x}}^*(c), \hat{\mathbf{u}}^*(c^+), \hat{\boldsymbol{\lambda}}^*(c)) , \quad (4.99c)$$

wobei  $c^-$  bzw.  $c^+$  den jeweiligen links- bzw. rechtsseitigen Grenzwert angeben. Man beachte, dass (4.99b) und (4.99c) der ersten und zweiten Weierstrass-Erdmann Bedingung von Abschnitt 4.1.3 entsprechen.

Im Folgenden werden die notwendigen Bedingungen erster Ordnung für ein Optimalsteuerungsproblem mit freier Endzeit und allgemeinen Endbeschränkungen formuliert und hergeleitet.

**Satz 4.9 (Steuerungsproblem: Endzeit frei/Endbeschränkung).** *Gesucht ist die Stellgröße  $\mathbf{u} \in (C[t_0, t_1])^m$  so, dass das Kostenfunktional (Bolza-Form)*

$$J(\mathbf{u}) = \varphi(t_1, \mathbf{x}(t_1)) + \int_{t_0}^{t_1} l(t, \mathbf{x}(t), \mathbf{u}(t)) dt \quad (4.100)$$

*unter den Gleichungsbeschränkungen*

$$\dot{\mathbf{x}} - \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t)) = \mathbf{0}, \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (4.101a)$$

$$G_k(\mathbf{u}, t_1) = \psi_k(t_1, \mathbf{x}(t_1)) = 0, \quad k = 1, \dots, p \quad (4.101b)$$

*mit fester Anfangszeit  $t_0$  und freier Endzeit  $t_1 \ll T$  minimiert wird. Dabei wird angenommen, dass  $l$  und  $\mathbf{f}$  stetig in  $t$ ,  $\mathbf{x}$  und  $\mathbf{u}$  und stetig differenzierbar bezüglich  $\mathbf{x}$  und  $\mathbf{u}$  für alle  $(t, \mathbf{x}, \mathbf{u}) \in [t_0, T] \times \mathbb{R}^n \times \mathbb{R}^m$  sind sowie die Funktionen  $\varphi$  und  $\psi_k$ ,  $k = 1, \dots, p$  stetig und stetig differenzierbar bezüglich  $t_1$  und  $\mathbf{x}_1$  für alle  $(t_1, \mathbf{x}_1) \in [t_0, T] \times \mathbb{R}^n$  sind.*

Weiters sei  $(\mathbf{u}^*, t_1^*) \in (C[t_0, t_1])^m \times [t_0, T]$  die optimale Lösung des Optimierungsproblems und  $\mathbf{x}^* \in (C^1[t_0, T])^n$  die zugehörige Lösung des Anfangswertproblems (4.101a). Darüber hinaus wird angenommen, dass für  $p$  unabhängige zulässige Richtungen  $(\boldsymbol{\xi}_k, \tau_k) \in (C[t_0, t_1])^m \times [t_0, T)$ ,  $k = 1, \dots, p$  die folgende Regularitätsbedingung

$$\det \left( \begin{bmatrix} \delta G_1(\mathbf{u}^*, t_1^*; \boldsymbol{\xi}_1, \tau_1) & \cdots & \delta G_1(\mathbf{u}^*, t_1^*; \boldsymbol{\xi}_p, \tau_p) \\ \vdots & \ddots & \vdots \\ \delta G_p(\mathbf{u}^*, t_1^*; \boldsymbol{\xi}_1, \tau_1) & \cdots & \delta G_p(\mathbf{u}^*, t_1^*; \boldsymbol{\xi}_p, \tau_p) \end{bmatrix} \right) \neq 0 \quad (4.102)$$

gilt. Dann existiert ein  $\boldsymbol{\lambda}^* \in (C^1[t_0, t_1^*])^n$  und ein  $\boldsymbol{\mu}^* \in \mathbb{R}^p$  so, dass die Beziehungen

$$\dot{\mathbf{x}}^* = \left( \frac{\partial}{\partial \boldsymbol{\lambda}} H \right)^T (t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t)), \quad \mathbf{x}^*(t_0) = \mathbf{x}_0 \quad (4.103a)$$

$$\dot{\boldsymbol{\lambda}}^* = - \left( \frac{\partial}{\partial \mathbf{x}} H \right)^T (t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t)), \quad \boldsymbol{\lambda}^*(t_1^*) = \left( \frac{\partial}{\partial \mathbf{x}_1} \Phi \right)^T (t_1^*, \mathbf{x}^*(t_1^*), \boldsymbol{\mu}^*) \quad (4.103b)$$

$$\mathbf{0} = \left( \frac{\partial}{\partial \mathbf{u}} H \right)^T (t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t)) \quad (4.103c)$$

für  $t_0 \leq t \leq t_1$  mit den Transversalitätsbedingungen

$$\boldsymbol{\psi}(t_1^*, \mathbf{x}^*(t_1^*)) = \mathbf{0} \quad (4.104a)$$

$$\left( \frac{\partial}{\partial t_1} \Phi \right) (t_1^*, \mathbf{x}^*(t_1^*), \boldsymbol{\mu}^*) + H(t_1^*, \mathbf{x}^*(t_1^*), \mathbf{u}^*(t_1^*), \boldsymbol{\lambda}^*(t_1^*)) = 0, \quad (4.104b)$$

und der Hamiltonfunktion  $H = l + \boldsymbol{\lambda}^T \mathbf{f}$  sowie  $\Phi = \varphi + \boldsymbol{\mu}^T \boldsymbol{\psi}$ ,  $\boldsymbol{\psi}^T = [\psi_1 \ \psi_2 \ \dots \ \psi_p]$  erfüllt sind.

*Beweis.* Man betrachte dazu wiederum die einparametrische Familie der zulässigen Stellgrößen  $\mathbf{v}(t; \eta) = \mathbf{u}^*(t) + \eta \boldsymbol{\xi}_u(t)$  mit  $\boldsymbol{\xi}_u \in (C[t_0, t_1])^m$  und dem skalaren Parameter  $\eta$ . Aufgrund der Stetigkeits- und Differenzierbarkeitsannahmen für  $\mathbf{f}$  existiert ein  $\eta > 0$  so, dass die zu  $\mathbf{v}(t)$  zugehörige Lösung  $\mathbf{y}(t; \eta)$  des Anfangswertproblems (4.101a) für alle  $t \in [t_0, T]$  eindeutig und bezüglich  $\eta$  differenzierbar ist. Für  $\eta = 0$  gilt offensichtlich  $\mathbf{v}(t; 0) = \mathbf{u}^*(t)$ ,  $t_0 \leq t \leq t_1^*$ , und  $\mathbf{y}(t; 0) = \mathbf{x}^*(t)$ ,  $t_0 \leq t \leq T$ . Da der Anfangswert  $\mathbf{x}(t_0) = \mathbf{x}_0$  fest ist, muss die Beziehung  $\mathbf{y}(t_0; \eta) = \mathbf{x}_0$  sowie  $(\frac{\partial}{\partial \eta} \mathbf{y})(t_0; 0) = \boldsymbol{\xi}_y(t_0) = \mathbf{0}$ , siehe auch (4.90), gelten. Das um die Gleichungsbeschränkungen (4.101b) erweiterte

Kostenfunktional (4.100) für  $\mathbf{v}(t; \eta)$  und der Endzeit  $\bar{t}_1 = t_1^* + \eta\tau$  lautet dann

$$\begin{aligned} \bar{J}(\mathbf{v}(t; \eta), \bar{t}_1) &= \int_{t_0}^{\bar{t}_1} l(t, \mathbf{y}(t; \eta), \mathbf{v}(t; \eta)) + \boldsymbol{\lambda}^T(t)(\mathbf{f}(t, \mathbf{y}(t; \eta), \mathbf{v}(t; \eta)) - \dot{\mathbf{y}}) dt \\ &\quad + \varphi(\bar{t}_1, \mathbf{y}(\bar{t}_1; \eta)) + \boldsymbol{\mu}^T \boldsymbol{\psi}(\bar{t}_1, \mathbf{y}(\bar{t}_1; \eta)) \\ &= \int_{t_0}^{\bar{t}_1} l(t, \mathbf{y}(t; \eta), \mathbf{v}(t; \eta)) + \boldsymbol{\lambda}^T(t)\mathbf{f}(t, \mathbf{y}(t; \eta), \mathbf{v}(t; \eta)) + \dot{\boldsymbol{\lambda}}^T(t)\mathbf{y}(t; \eta) dt \\ &\quad - \boldsymbol{\lambda}^T(\bar{t}_1)\mathbf{y}(\bar{t}_1; \eta) + \boldsymbol{\lambda}^T(t_0)\mathbf{y}(t_0; \eta) + \varphi(\bar{t}_1, \mathbf{y}(\bar{t}_1; \eta)) + \boldsymbol{\mu}^T \boldsymbol{\psi}(\bar{t}_1, \mathbf{y}(\bar{t}_1; \eta)) . \end{aligned} \quad (4.105)$$

Gemäß Satz 4.1 berechnet sich die notwendige Bedingung für ein Minimum aus

$$\delta \bar{J}(\mathbf{u}^*, t_1^*; \boldsymbol{\xi}_u, \tau) = \left. \frac{\partial \bar{J}(\mathbf{v}(t; \eta), \bar{t}_1)}{\partial \eta} \right|_{\eta=0} = 0 \quad (4.106)$$

zu

$$\begin{aligned} 0 &= \int_{t_0}^{t_1^*} \left( \frac{\partial}{\partial \mathbf{x}} l \right) (t, \mathbf{x}^*, \mathbf{u}^*) \boldsymbol{\xi}_y + \left( \frac{\partial}{\partial \mathbf{u}} l \right) (t, \mathbf{x}^*, \mathbf{u}^*) \boldsymbol{\xi}_u + \boldsymbol{\lambda}^T(t) \left( \frac{\partial}{\partial \mathbf{x}} \mathbf{f} \right) (t, \mathbf{x}^*, \mathbf{u}^*) \boldsymbol{\xi}_y dt \\ &\quad + \int_{t_0}^{t_1^*} \boldsymbol{\lambda}^T(t) \left( \frac{\partial}{\partial \mathbf{u}} \mathbf{f} \right) (t, \mathbf{x}^*, \mathbf{u}^*) \boldsymbol{\xi}_u + \dot{\boldsymbol{\lambda}}^T(t) \boldsymbol{\xi}_y dt + l(t_1^*, \mathbf{y}(t_1^*; 0), \mathbf{v}(t_1^*; 0))\tau \\ &\quad + \boldsymbol{\lambda}^T(t_1^*) \mathbf{f}(t_1^*, \mathbf{y}(t_1^*; 0), \mathbf{v}(t_1^*; 0))\tau + \dot{\boldsymbol{\lambda}}^T(t_1^*) \mathbf{y}(t_1^*; 0)\tau - \boldsymbol{\lambda}^T(t_1^*) \boldsymbol{\xi}_y(t_1^*) + \boldsymbol{\lambda}^T(t_0) \boldsymbol{\xi}_y(t_0) \\ &\quad - \dot{\boldsymbol{\lambda}}^T(t_1^*) \mathbf{y}(t_1^*; 0)\tau - \boldsymbol{\lambda}^T(t_1^*) \dot{\mathbf{y}}(t_1^*; 0)\tau + \left( \frac{\partial}{\partial \mathbf{x}_1} \varphi \right) (t_1^*, \mathbf{y}(t_1^*; 0)) \boldsymbol{\xi}_y(t_1^*) \\ &\quad + \left( \frac{\partial}{\partial t_1} \varphi \right) (t_1^*, \mathbf{y}(t_1^*; 0))\tau + \left( \frac{\partial}{\partial \mathbf{x}_1} \varphi \right) (t_1^*, \mathbf{y}(t_1^*; 0)) \dot{\mathbf{y}}(t_1^*; 0)\tau \\ &\quad + \boldsymbol{\mu}^T \left( \frac{\partial}{\partial \mathbf{x}_1} \boldsymbol{\psi} \right) (t_1^*, \mathbf{y}(t_1^*; 0)) \boldsymbol{\xi}_y(t_1^*) + \boldsymbol{\mu}^T \left( \frac{\partial}{\partial t_1} \boldsymbol{\psi} \right) (t_1^*, \mathbf{y}(t_1^*; 0))\tau \\ &\quad + \boldsymbol{\mu}^T \left( \frac{\partial}{\partial \mathbf{x}_1} \boldsymbol{\psi} \right) (t_1^*, \mathbf{y}(t_1^*; 0)) \dot{\mathbf{y}}(t_1^*; 0)\tau \end{aligned} \quad (4.107)$$

und mit  $\mathbf{y}(t_1^*; 0) = \mathbf{x}^*(t_1^*)$ ,  $\mathbf{v}(t_1^*; 0) = \mathbf{u}^*(t_1^*)$  folgt

$$\begin{aligned} 0 &= \int_{t_0}^{t_1^*} \left\{ \left( \frac{\partial}{\partial \mathbf{x}} l \right) (t, \mathbf{x}^*, \mathbf{u}^*) + \boldsymbol{\lambda}^T(t) \left( \frac{\partial}{\partial \mathbf{x}} \mathbf{f} \right) (t, \mathbf{x}^*, \mathbf{u}^*) + \dot{\boldsymbol{\lambda}}^T(t) \right\} \boldsymbol{\xi}_y dt \\ &\quad + \int_{t_0}^{t_1^*} \left\{ \left( \frac{\partial}{\partial \mathbf{u}} l \right) (t, \mathbf{x}^*, \mathbf{u}^*) + \boldsymbol{\lambda}^T(t) \left( \frac{\partial}{\partial \mathbf{u}} \mathbf{f} \right) (t, \mathbf{x}^*, \mathbf{u}^*) \right\} \boldsymbol{\xi}_u dt + \\ &\quad + \left\{ \left( \frac{\partial}{\partial \mathbf{x}_1} \varphi + \boldsymbol{\mu}^T \frac{\partial}{\partial \mathbf{x}_1} \boldsymbol{\psi} \right) (t_1^*, \mathbf{x}^*(t_1^*)) - \boldsymbol{\lambda}^T(t_1^*) \right\} (\boldsymbol{\xi}_y(t_1^*) + \dot{\mathbf{x}}(t_1^*)\tau) \\ &\quad + \left\{ \left( \frac{\partial}{\partial t_1} \varphi + \boldsymbol{\mu}^T \frac{\partial}{\partial t_1} \boldsymbol{\psi} \right) (t_1^*, \mathbf{x}^*(t_1^*)) + (l + \boldsymbol{\lambda}^T \mathbf{f}) (t_1^*, \mathbf{x}^*(t_1^*), \mathbf{u}^*(t_1^*)) \right\} \tau . \end{aligned} \quad (4.108)$$

Da die Auswirkung der Variation von  $\mathbf{u}^*(t)$  auf die Lösung in Form von  $\xi_y(t)$  schwierig zu berechnen ist, wählt man  $\boldsymbol{\lambda}^*(t)$  so, dass die erste Zeile in (4.108) identisch verschwindet, d. h.

$$\dot{\boldsymbol{\lambda}}^* = -\left(\frac{\partial}{\partial \mathbf{x}} \mathbf{f}\right)^T(t, \mathbf{x}^*(t), \mathbf{u}^*(t)) \boldsymbol{\lambda}^*(t) - \left(\frac{\partial}{\partial \mathbf{x}} l\right)^T(t, \mathbf{x}^*(t), \mathbf{u}^*(t)) \quad (4.109)$$

mit der zugehörigen Endwertbedingung so, dass die dritte Zeile in (4.108) zu Null wird, also

$$\boldsymbol{\lambda}^*(t_1^*) = \left(\frac{\partial}{\partial \mathbf{x}_1} \varphi + (\boldsymbol{\mu}^*)^T \frac{\partial}{\partial \mathbf{x}_1} \boldsymbol{\psi}\right)^T(t_1^*, \mathbf{x}^*(t_1^*)). \quad (4.110)$$

Die Transversalitätsbedingung (4.104b) folgt unmittelbar aus der letzten Zeile von (4.108) und die Extremalbedingung für die Hamiltonfunktion in (4.103) ergibt sich wiederum direkt aufgrund des Fundamentallemmas der Variationsrechnung angewandt auf die zweite Zeile von (4.108).  $\square$

Für eine Zusammenfassung der Ergebnisse von Satz 4.9 werden in weiterer Folge nur so genannte partielle Endbedingungen der Form

$$\psi_j = x_k(t_1) - \bar{x}_k, \quad j = 1, \dots, p \quad (4.111)$$

mit  $\bar{x}_k = \text{konst.}$  als fixem Endwert der Komponente  $x_k$  von  $\mathbf{x}$  betrachtet. Für diesen vereinfachenden Spezialfall kann die Endbedingung für  $\boldsymbol{\lambda}^*(t_1^*)$  von (4.103b) ersetzt werden und unter Berücksichtigung von (4.108) gilt Folgendes:

- (a) Wenn die *Endzeit fest ist*, dann gilt  $t_1 = t_1^*$  und damit  $\tau = 0$ , womit automatisch der zugehörige Eintrag in der vierten Zeile von (4.108) verschwindet. Es liegt somit keine Transversalitätsbedingung gemäß (4.104b) vor.
- (i) Wenn für eine Komponente  $x_k$  von  $\mathbf{x}$  gilt, dass *der Endwert fest ist*, dann gilt  $y_k(t_1^*; \eta) = x_k(t_1^*) = x_k^*(t_1^*) = \bar{x}_k$ . Daraus folgt  $\xi_{y,k}(t_1^*) = 0$  womit automatisch der zugehörige Eintrag in der dritten Zeile von (4.108) verschwindet. Damit liegt für diese Komponente keine Endbedingung für den zugehörigen adjungierten Zustand  $\lambda_k^*(t_1^*)$  vor.
- (ii) Wenn für eine Komponente  $x_k$  von  $\mathbf{x}$  gilt, dass *der Endwert frei ist*, dann muss, wie man aus der dritten Zeile von (4.108) erkennen kann, die Komponente des zugehörigen adjungierten Zustands  $\lambda_k^*(t)$  folgende Endbedingung

$$\lambda_k^*(t_1^*) = \left(\frac{\partial}{\partial x_{1,k}} \varphi\right)(t_1^*, \mathbf{x}^*(t_1^*)) \quad (4.112)$$

erfüllen.

- (b) Wenn die *Endzeit frei ist*, dann muss die Transversalitätsbedingung

$$\left(\frac{\partial}{\partial t} \Phi\right)(t_1^*, \mathbf{x}^*(t_1^*), \boldsymbol{\mu}^*) + H(t_1^*, \mathbf{x}^*(t_1^*), \mathbf{u}^*(t_1^*), \boldsymbol{\lambda}^*(t_1^*)) = 0, \quad H = l + \boldsymbol{\lambda}^T \mathbf{f} \quad (4.113)$$



gelten. In Abhängigkeit davon, ob der Endwert einer Komponente  $x_k$  von  $\mathbf{x}$  fest oder frei ist, können die Unterpunkte (i) und (ii) vom Fall (a) auch direkt hier angewandt werden.

Wenn in Satz 4.9 die Gleichungsbeschränkungen (4.101b) durch *Ungleichungsbeschränkungen* der Form

$$G_k(\mathbf{u}, t_1) = \psi_k(t_1, \mathbf{x}(t_1)) \leq 0, \quad k = 1, \dots, p \quad (4.114)$$

ersetzt werden, so ändert sich lediglich (4.104a) zu

$$\psi_k(t_1^*, \mathbf{x}^*(t_1^*)) \leq 0 \quad (4.115a)$$

$$\boldsymbol{\mu}^* \geq \mathbf{0} \quad (4.115b)$$

$$(\boldsymbol{\mu}^*)^T \boldsymbol{\psi}(t_1^*, \mathbf{x}^*(t_1^*)) = 0. \quad (4.115c)$$

**Aufgabe 4.3.** Gesucht ist eine Lösung des Optimierungsproblems

$$\min_{u(\cdot)} \int_0^1 \frac{1}{2} u^2 + \frac{a}{2} x^2 dt, \quad a > 0 \quad (4.116a)$$

$$\text{u.B.v. } \dot{x} = u, \quad x(0) = 1, \quad x(1) = 0. \quad (4.116b)$$

Zeigen Sie, dass die Lösung durch

$$x^*(t) = \frac{1}{1 - e^{2\sqrt{a}}} (e^{\sqrt{a}t} - e^{\sqrt{a}(2-t)}), \quad u^*(t) = \frac{\sqrt{a}}{1 - e^{2\sqrt{a}}} (e^{\sqrt{a}t} + e^{\sqrt{a}(2-t)}) \quad (4.117)$$

gegeben ist und interpretieren Sie die Ergebnisse in Abbildung 4.4, die für verschiedene Parameterwerte  $a$  dargestellt sind.

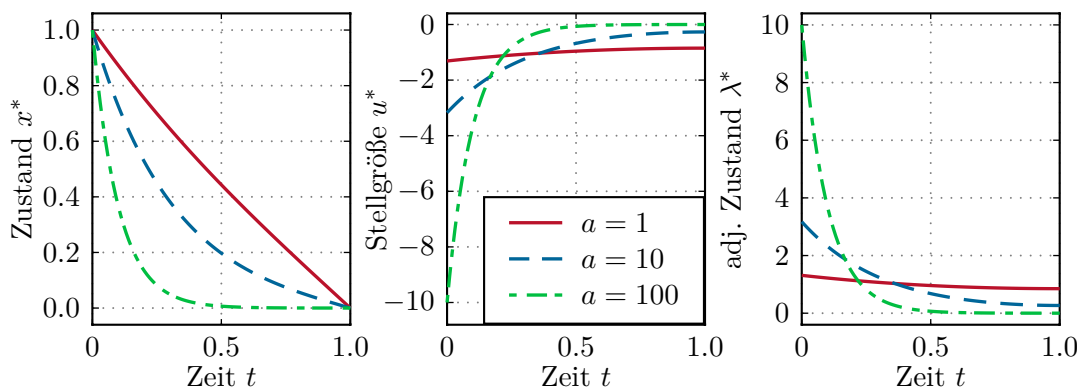


Abbildung 4.4: Optimale Trajektorien in Aufgabe 4.3.

**Aufgabe 4.4.** Gesucht ist eine Lösung des Optimierungsproblems

$$\min_{u(\cdot)} \quad \frac{a}{2} x_2^2(1) + \int_0^1 \frac{1}{2} u^2 dt, \quad a \geq 0 \quad (4.118a)$$

$$\text{u.B.v.} \quad \dot{x}_1 = x_2, \quad x_1(0) = 1, \quad x_1(1) = 0 \quad (4.118b)$$

$$\dot{x}_2 = u, \quad x_2(0) = 0. \quad (4.118c)$$

Zeigen Sie, dass sich für den (freien) Endzustand  $x_2^*(1) = -6/(4+a)$  in Abhängigkeit des Parameters  $a \geq 0$  ergibt und dass die optimale Lösung durch

$$x_1^*(t) = \frac{2(1+a)}{4+a} t^3 - \frac{3(2+a)}{a+4} t^2 + 1, \quad u^*(t) = \frac{12(1+a)}{4+a} t - \frac{6(2+a)}{a+4} \quad (4.119)$$

gegeben ist. Interpretieren Sie die Ergebnisse in Abbildung 4.5.

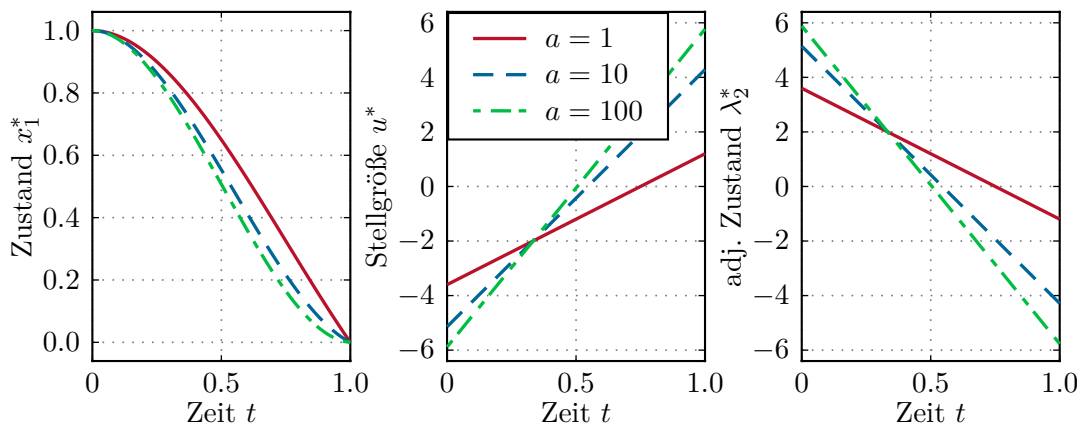


Abbildung 4.5: Optimal Trajektorien in Aufgabe 4.4.

**Beispiel 4.6.** Betrachtet wird eine Punktmasse der Masse  $m$  in der  $(x, y)$ -Ebene, auf die eine konstante Kraft  $F = ma$  wirkt. Die Stellgröße  $u$  des Problems ist der Winkel zwischen der Schubrichtung und der  $x$ -Achse, siehe Abbildung 4.6. Ziel ist es, die Punktmasse in *minimaler Zeit*  $[t_0 = 0, t_1^*]$  zu einem *fest vorgegebenen Zielpunkt*  $(\bar{x}_1, \bar{y}_1)$  zu steuern. Unter der Annahme, dass außer dem Schub keine weiteren Kräfte auftreten, kann das Optimalsteuerungsproblem wie folgt formuliert werden

$$\min_{u(\cdot)} \quad t_1 \quad (4.120a)$$

$$\text{u.B.v.} \quad \dot{x} = v, \quad x(0) = x_0, \quad x(t_1) = \bar{x}_1 \quad (4.120b)$$

$$\dot{v} = a \cos(u), \quad v(0) = v_0 \quad (4.120c)$$

$$\dot{y} = w, \quad y(0) = y_0, \quad y(t_1) = \bar{y}_1 \quad (4.120d)$$

$$\dot{w} = a \sin(u), \quad w(0) = w_0. \quad (4.120e)$$

Man beachte, dass der Endzustand nur für die Position  $(x, y)$  aber nicht für die Geschwindigkeiten  $(v, w)$  vorgegeben ist.

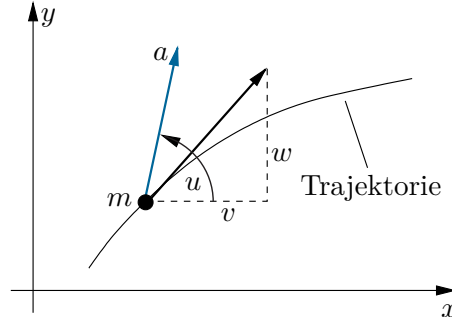


Abbildung 4.6: Bewegung einer Punktmasse der Masse  $m$  in der  $(x, y)$ -Ebene.

Die beiden fest vorgegebenen Endwerte für  $x$  und  $y$  können als Gleichungsbeschränkungen gemäß (4.101b) in der Form

$$\psi_1(t_1, \mathbf{x}(t_1)) = x(t_1) - \bar{x}_1, \quad \psi_2(t_1, \mathbf{x}(t_1)) = y(t_1) - \bar{y}_1 \quad (4.121)$$

formuliert werden. Die Hamiltonfunktion  $H$  und die Funktion  $\Phi$  gemäß Satz 4.9 lauten dann für das vorliegende Optimierungsproblem

$$H(\mathbf{x}, u, \boldsymbol{\lambda}) = \lambda_x v + \lambda_v a \cos(u) + \lambda_y w + \lambda_w a \sin(u) \quad (4.122a)$$

$$\Phi(t_1, \mathbf{x}(t_1), \boldsymbol{\mu}) = \varphi + \mu_x \psi_1 + \mu_y \psi_2 = t_1 + \mu_x (x(t_1) - \bar{x}_1) + \mu_y (y(t_1) - \bar{y}_1) \quad (4.122b)$$

mit  $\mathbf{x} = [x \quad v \quad y \quad w]^T$ , den adjungierten Zuständen  $\boldsymbol{\lambda} = [\lambda_x \quad \lambda_v \quad \lambda_y \quad \lambda_w]^T$  und den konstanten Lagrange-Multiplikatoren  $\boldsymbol{\mu} = [\mu_x \quad \mu_y]^T$ . Die Randbedingungen für den adjungierten Zustand errechnen sich gemäß (4.103) zu

$$\lambda_x^*(t_1^*) = \left( \frac{\partial}{\partial x_1} \Phi \right) (t_1^*, \mathbf{x}^*(t_1^*), \boldsymbol{\mu}^*) = \mu_x^*, \quad \lambda_v^*(t_1^*) = \left( \frac{\partial}{\partial v_1} \Phi \right) (t_1^*, \mathbf{x}^*(t_1^*), \boldsymbol{\mu}^*) = 0 \quad (4.123a)$$

$$\lambda_y^*(t_1^*) = \left( \frac{\partial}{\partial y_1} \Phi \right) (t_1^*, \mathbf{x}^*(t_1^*), \boldsymbol{\mu}^*) = \mu_y^*, \quad \lambda_w^*(t_1^*) = \left( \frac{\partial}{\partial w_1} \Phi \right) (t_1^*, \mathbf{x}^*(t_1^*), \boldsymbol{\mu}^*) = 0 \quad (4.123b)$$

mit  $\mathbf{x}(t_1^*) = [x(t_1^*) \quad v(t_1^*) \quad y(t_1^*) \quad w(t_1^*)]^T = \mathbf{x}_1 = [x_1 \quad v_1 \quad y_1 \quad w_1]^T$  und damit

lautet das adjungierte System

$$\dot{\lambda}_x^* = -\left(\frac{\partial H}{\partial x}\right)(\mathbf{x}^*, \mathbf{u}^*, \boldsymbol{\lambda}^*) = 0, \quad \lambda_x^*(t_1^*) = \mu_x^* \quad (4.124a)$$

$$\dot{\lambda}_v^* = -\left(\frac{\partial H}{\partial v}\right)(\mathbf{x}^*, \mathbf{u}^*, \boldsymbol{\lambda}^*) = -\lambda_x^*, \quad \lambda_v^*(t_1^*) = 0 \quad (4.124b)$$

$$\dot{\lambda}_y^* = -\left(\frac{\partial H}{\partial y}\right)(\mathbf{x}^*, \mathbf{u}^*, \boldsymbol{\lambda}^*) = 0, \quad \lambda_y^*(t_1^*) = \mu_y^* \quad (4.124c)$$

$$\dot{\lambda}_w^* = -\left(\frac{\partial H}{\partial w}\right)(\mathbf{x}^*, \mathbf{u}^*, \boldsymbol{\lambda}^*) = -\lambda_y^*, \quad \lambda_w^*(t_1^*) = 0, \quad (4.124d)$$

woraus direkt

$$\lambda_x^* = \mu_x^*, \quad \lambda_v^* = \mu_x^*(t_1^* - t), \quad \lambda_y^* = \mu_y^*, \quad \lambda_w^* = \mu_y^*(t_1^* - t) \quad (4.125)$$

folgt. Des Weiteren muss die Hamiltonfunktion  $H$  gemäß (4.103) extremal sein, weshalb die folgende Bedingung

$$\left(\frac{\partial H}{\partial u}\right)(\mathbf{x}^*, \mathbf{u}^*, \boldsymbol{\lambda}^*) = -\lambda_v^* a \sin(u^*) + \lambda_w^* a \cos(u^*) = 0, \quad (4.126)$$

erfüllt sein muss und  $u^*$  sich in der Form

$$\begin{aligned} \tan(u^*) &= \frac{\lambda_w^*}{\lambda_v^*} \stackrel{(4.125)}{=} \frac{\mu_y^*(t_1^* - t)}{\mu_x^*(t_1^* - t)} = \frac{\mu_y^*}{\mu_x^*} = \text{konst.} \\ \Rightarrow u^* &= \arctan\left(\frac{\mu_y^*}{\mu_x^*}\right) \quad \text{mit} \quad -\frac{\pi}{2} < u^* < \frac{\pi}{2} \end{aligned} \quad (4.127)$$

berechnen lässt. Die optimale Steuerung  $u^*$  ist also auf dem gesamten Zeitintervall  $[t_0, t_1^*]$  konstant und die zugehörigen optimalen Zustandstrajektorien  $\mathbf{x}^*(t)$  können durch Lösen der Differentialgleichungen (4.120) und Einsetzen der Anfangsbedingungen in der Form

$$x^*(t) = g_x(\mu_x^*, \mu_y^*, t) = x_0 + v_0 t + \frac{1}{2} a \cos(u^*) t^2, \quad \cos(u^*) = \frac{1}{\sqrt{1 + (\mu_y^*/\mu_x^*)^2}} \quad (4.128a)$$

$$v^*(t) = g_v(\mu_x^*, \mu_y^*, t) = v_0 + a \cos(u^*) t \quad (4.128b)$$

$$y^*(t) = g_y(\mu_x^*, \mu_y^*, t) = y_0 + w_0 t + \frac{1}{2} a \sin(u^*) t^2, \quad \sin(u^*) = \frac{\mu_y^*}{\mu_x^* \sqrt{1 + (\mu_y^*/\mu_x^*)^2}} \quad (4.128c)$$

$$w^*(t) = g_w(\mu_x^*, \mu_y^*, t) = w_0 + a \sin(u^*) t. \quad (4.128d)$$

bestimmt werden. Dabei wurden die trigonometrischen Beziehungen

$$\sin(\arctan(b)) = \frac{b}{\sqrt{1+b^2}}, \quad \cos(\arctan(b)) = \frac{1}{\sqrt{1+b^2}} \quad (4.129)$$

verwendet. Da die Endzeit  $t_1$  frei ist, muss zusätzlich die Transversalitätsbedingung (4.104) gelten, wobei sich die Hamiltonfunktion (4.122a) aufgrund der Endbedingungen  $\lambda_v^*(t_1^*) = \lambda_w^*(t_1^*) = 0$  entsprechend vereinfacht

$$0 = \left( \frac{\partial}{\partial t_1} \Phi \right) (t_1^*, \mathbf{x}^*(t_1^*), \boldsymbol{\mu}^*) + H(\mathbf{x}^*(t_1^*), \mathbf{u}^*(t_1^*), \boldsymbol{\lambda}^*(t_1^*)) \quad (4.130a)$$

$$= 1 + \mu_x^* g_v(\mu_x^*, \mu_y^*, t_1^*) + \mu_y^* g_w(\mu_x^*, \mu_y^*, t_1^*) . \quad (4.130b)$$

Mit Hilfe der zwei Gleichungsbeschränkungen (4.121) und der Transversalitätsbedingung (4.130) lässt sich ein Gleichungssystem für die verbleibenden drei Unbekannten  $\mu_x^*$ ,  $\mu_y^*$  und  $t_1^*$  in der Form

$$\begin{bmatrix} g_x(\mu_x^*, \mu_y^*, t_1^*) \\ g_y(\mu_x^*, \mu_y^*, t_1^*) \\ \mu_x^* g_v(\mu_x^*, \mu_y^*, t_1^*) + \mu_y^* g_w(\mu_x^*, \mu_y^*, t_1^*) \end{bmatrix} - \begin{bmatrix} \bar{x}_1 \\ \bar{y}_1 \\ -1 \end{bmatrix} = \mathbf{0} \quad (4.131)$$

formulieren, welches auf *numerischem Wege* gelöst werden kann. Eine geeignete MATLAB-Funktion zur Lösung von nichtlinearen Gleichungen ist mit dem Befehl `fsolve` aus der Optimization Toolbox gegeben. Die Funktion `fsolve` verwendet standardmäßig die Methode der Vertrauensbereiche (siehe Abschnitt 2.4), um ein Gleichungssystem in Residuenform  $\mathbf{F}(\mathbf{x}) = \mathbf{0}$  als Minimierungsproblem in  $\mathbf{x}$  zu lösen. Als Beispiel ist in der Code-Auflistung 4.1 der MATLAB-Code dargestellt, wie `fsolve` zur Lösung von (4.131) verwendet werden kann. Der gewünschte Endpunkt  $(\bar{x}_1, \bar{y}_1)$  wird beim Aufruf der Funktion `punktmasse(x1,y1)` übergeben, wobei angenommen wird, dass die Punktmasse am Punkt  $(x_0, y_0) = (0, 0)$  mit der Geschwindigkeit  $(v_0, w_0) = (0, 1)$  in vertikale Richtung startet. Abbildung 4.7 stellt die optimalen Bahnen  $x^*(t)$ ,  $y^*(t)$  der Punktmasse in der  $(x, y)$ -Ebene für verschiedene Endpunkte  $(\bar{x}_1, \bar{y}_1)$  dar. Die Pfeile zeigen die (konstante) Richtung  $u^* = \arctan(\mu_y^*/\mu_x^*)$  der angreifenden Kraft  $ma$  an.

Listing 4.1: MATLAB-Code für das Punktmasse-Problem unter Verwendung von `fsolve`.

```
function [t,x,y,p] = punktmasse(x1,y1)
% -----
% (x1,y1): gewünschter Endpunkt
% (t,x,y): Trajektorien der Punktmasse
% p:      Parameterstruktur

p.a = 1; % Parameter
p.x0=0; p.v0=0; p.y0=0; p.w0=1; % Anfangsbedingungen
p.x1=x1; p.y1=y1; % Endbedingungen (Übergabe aus Funktionsaufruf)

opt = optimoptions('fsolve','Display','iter'); % Optionen
X0 = [-1,0,1]; % Startwert
Xopt = fsolve(@eqns,X0,opt,p); % Numerische Lösung mit fsolve
p.mux=Xopt(1); p.muy=Xopt(2); p.t1=Xopt(3); % Lösung
```

```

t = linspace(0,p.t1,100);          % Trajektorien
x = xfct(p.mux,p.muy,t,p);
y = yfct(p.mux,p.muy,t,p);
% -----
function res = eqns(X,p)           % Gleichungen in Residuenform
mux=X(1); muy=X(2); t1=X(3);
res = [ xfct(mux,muy,t1,p) - p.x1;
        yfct(mux,muy,t1,p) - p.y1;
        mux*vfct(mux,muy,t1,p) + muy*wfct(mux,muy,t1,p) + 1 ];
% -----
function x = xfct(mux,muy,t,p)     % Funktionen für x und v
cosu = 1/sqrt(1+(muy/mux)^2);
x = p.x0 + p.v0*t + p.a/2*cosu*t.^2; % 't.^2' steht für komponentenweise Auswertung
function v = vfct(mux,muy,t,p)
cosu = 1/sqrt(1+(muy/mux)^2);
v = p.v0 + p.a*cosu*t;
% -----
function y = yfct(mux,muy,t,p)     % Funktionen für y und w
sinu = muy/(mux*sqrt(1+(muy/mux)^2));
y = p.y0 + p.w0*t + p.a/2*sinu*t.^2; % 't.^2' steht für komponentenweise Auswertung
function w = wfct(mux,muy,t,p)
sinu = muy/(mux*sqrt(1+(muy/mux)^2));
w = p.w0 + p.a*sinu*t;

```

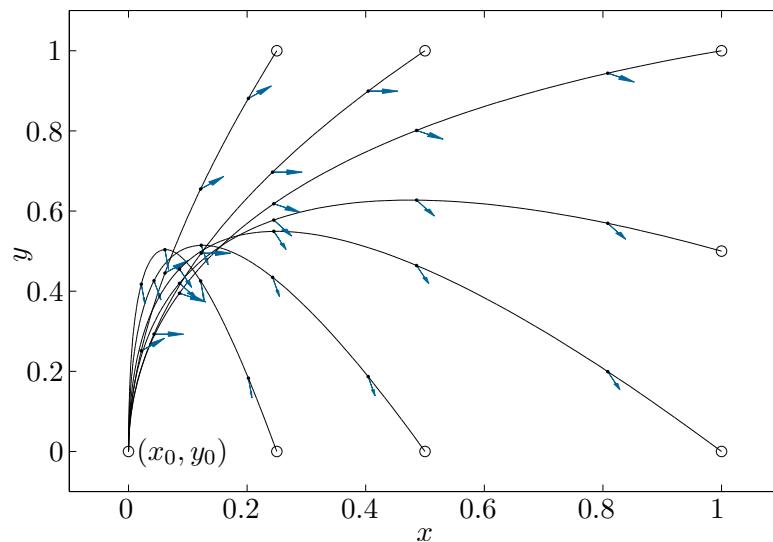


Abbildung 4.7: Zeitoptimale Steuerung einer Punktmasse zu verschiedenen Endpunkten.

**Aufgabe 4.5.** Gegeben ist ein lineares zeitvariantes Mehrgrößensystem der Form

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{u}, \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (4.132)$$

mit dem Zustand  $\mathbf{x} \in \mathbb{R}^n$  und dem Stelleingang  $\mathbf{u} \in \mathbb{R}^m$ . Zeigen Sie, dass das zeitvariante Zustandsregelgesetz

$$\mathbf{u}^*(t) = -\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{S}(t)\mathbf{x}^*(t) \quad (4.133)$$

mit  $\mathbf{S}(t)$  als Lösung der *Matrix-Riccati-Differentialgleichung*

$$\dot{\mathbf{S}} = -\mathbf{S}\mathbf{A} - \mathbf{A}^T\mathbf{S} + \mathbf{S}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T\mathbf{S} - \mathbf{Q}, \quad \mathbf{S}(t_1) = \mathbf{S}_1, \quad t_0 \leq t \leq t_1 \quad (4.134)$$

das Kostenfunktional

$$J(\mathbf{u}) = \frac{1}{2} \int_{t_0}^{t_1} \mathbf{x}^T(t) \mathbf{Q}(t) \mathbf{x}(t) + \mathbf{u}^T(t) \mathbf{R}(t) \mathbf{u}(t) dt + \frac{1}{2} \mathbf{x}^T(t_1) \mathbf{S}_1 \mathbf{x}(t_1) \quad (4.135)$$

mit der für alle Zeiten  $t_0 \leq t \leq t_1$  positiv definiten Matrix  $\mathbf{R}(t)$ , der für alle Zeiten  $t_0 \leq t \leq t_1$  positiv semidefiniten Matrix  $\mathbf{Q}(t)$  und der positiv semidefiniten Matrix  $\mathbf{S}_1$  minimiert. Dieses Problem ist auch unter dem Namen *LQR (Linear Quadratic Regulator) Problem* bekannt, vergleiche dazu auch Kapitel 3 im Skript Regelungssysteme 1.

#### 4.2.4 Minimumsprinzip von Pontryagin

Für das Weitere betrachte man im ersten Schritt die Minimierung des Kostenfunktionals

$$J(\mathbf{u}) = \int_{t_0}^{t_1} l(\mathbf{x}(t), \mathbf{u}(t)) dt \quad (4.136)$$

mit der freien Endzeit  $t_1$  und einem festen Endzustand  $\mathbf{x}(t_1) = \mathbf{x}_1$  unter der Gleichungsbeschränkung des dynamischen Systems

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (4.137)$$

für die zulässigen Stellgrößen

$$\mathbf{u} \in (\hat{C}_U[t_0, T])^m := \left\{ \mathbf{u} \in (\hat{C}[t_0, T])^m : \mathbf{u}(t) \in U, \forall t_0 \leq t \leq t_1 \right\} \quad (4.138)$$

mit einer hinreichend großen Zeit  $T \gg t_1$  und der nichtleeren Menge der Stellgrößenbeschränkungen  $U$ .

Man kann nun das Optimierungsproblem (4.136), (4.137) durch Erweiterung des Zustandsvektors in der Form  $\bar{\mathbf{x}}^T = [\mathbf{x}^T \quad x_{n+1}]$  mit

$$x_{n+1}(t) := \int_{t_0}^t l(\mathbf{x}(\tau), \mathbf{u}(\tau)) d\tau \quad (4.139)$$

wie folgt umformulieren: Gesucht wird eine zulässige Stellgröße  $\mathbf{u}(t) \in (\hat{C}_U[t_0, T])^m$  und eine Endzeit  $t_1$  so, dass die Lösung des erweiterten Systems

$$\underbrace{\begin{bmatrix} \dot{\mathbf{x}} \\ \dot{x}_{n+1} \end{bmatrix}}_{\dot{\bar{\mathbf{x}}}} = \underbrace{\begin{bmatrix} \mathbf{f}(\mathbf{x}, \mathbf{u}) \\ l(\mathbf{x}, \mathbf{u}) \end{bmatrix}}_{\bar{\mathbf{f}}(\bar{\mathbf{x}}, \mathbf{u})}, \quad \bar{\mathbf{x}}(t_0) = \begin{bmatrix} \mathbf{x}_0 \\ 0 \end{bmatrix} \quad (4.140)$$

beim Punkt  $\bar{\mathbf{x}}^T(t_1) = [\mathbf{x}_1^T \quad x_{n+1}(t_1)]$  terminiert und dabei  $x_{n+1}(t_1)$  möglichst klein gemacht wird. Abbildung 4.8 veranschaulicht diesen Sachverhalt.

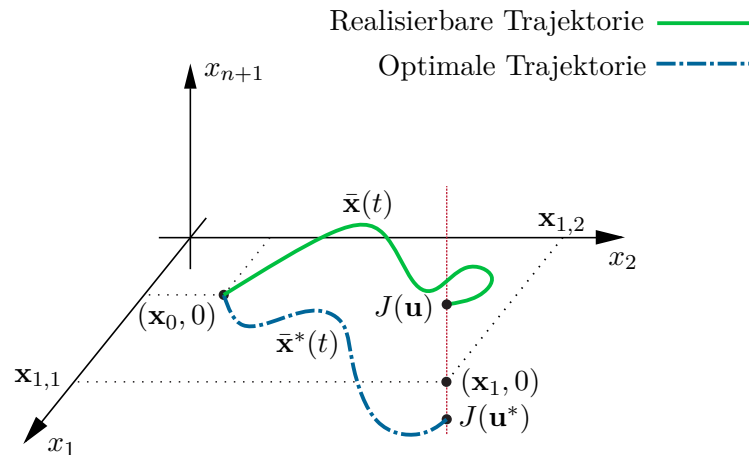


Abbildung 4.8: Zum Minimumsprinzip von Pontryagin.

Die Linie durch den Punkt  $(\mathbf{x}_1, 0)$  parallel zur  $x_{n+1}$ -Achse beschreibt alle Punkte einer Familie von Trajektorien  $\bar{\mathbf{x}}(t)$  des erweiterten Systems (4.140), die die Bedingung  $\mathbf{x}(t_1) = \mathbf{x}_1$  erfüllen und unterschiedliche Werte des Kostenfunktional  $x_{n+1}(t_1)$  ergeben. Keine andere Trajektorie kann diese vertikale Linie an einem kleineren Wert für  $x_{n+1}(t_1)$  schneiden als  $x_{n+1}^*(t_1^*)$ , der durch die optimale Stellgröße  $\mathbf{u}^*(t)$  erzielt wird. Diese geometrischen Überlegungen sind auch der Ausgangspunkt für die Herleitung des Minimumsprinzips von Pontryagin. An dieser Stelle wird jedoch auf einen Beweis verzichtet und auf die am Ende des Kapitels angegebene Literatur verwiesen.

**Satz 4.10 (Minimumsprinzip von Pontryagin, zeitinvarianter Fall).** *Gesucht ist die Stellgröße  $\mathbf{u} \in (\hat{C}_U[t_0, t_1])^m$  so, dass das Kostenfunktional (Lagrange-Form)*

$$J(\mathbf{u}) = \int_{t_0}^{t_1} l(\mathbf{x}(t), \mathbf{u}(t)) dt \quad (4.141)$$

*unter den Gleichungsbeschränkungen (dynamisches System)*

$$\dot{\mathbf{x}} - \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) = \mathbf{0}, \quad \mathbf{x}(t_0) = \mathbf{x}_0, \quad \mathbf{x}(t_1) = \mathbf{x}_1 \quad (4.142)$$

*mit fester Anfangszeit  $t_0$  und freier Endzeit  $t_1 \ll T$  minimiert wird. Dabei wird angenommen, dass  $l$  und  $\mathbf{f}$  stetig in  $\mathbf{x}$  und  $\mathbf{u}$  und stetig differenzierbar bezüglich  $\mathbf{x}$  für alle  $(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^n \times \mathbb{R}^m$  sind. Weiters sei  $(\mathbf{u}^*, t_1^*) \in (\hat{C}_U[t_0, t_1^*])^m \times [t_0, T)$  die optimale Lösung des Optimierungsproblems und  $\mathbf{x}^*$  die zugehörige Lösung von (4.142). Dann existiert ein  $\bar{\boldsymbol{\lambda}}^* \in (\hat{C}^1[t_0, t_1^*])^{n+1}$ ,  $\bar{\boldsymbol{\lambda}}^* = [\bar{\lambda}_1^* \ \dots \ \bar{\lambda}_{n+1}^*]^\top \neq [0 \ \dots \ 0]^\top$  so, dass die Beziehung*

$$\dot{\bar{\boldsymbol{\lambda}}}^* = - \left( \frac{\partial}{\partial \bar{\mathbf{x}}} H \right)^\top (\mathbf{x}^*(t), \mathbf{u}^*(t), \bar{\boldsymbol{\lambda}}^*(t)) \quad (4.143)$$



für  $t_0 \leq t \leq t_1$  mit  $H(\mathbf{x}^*(t), \mathbf{u}^*(t), \bar{\boldsymbol{\lambda}}^*(t)) = (\bar{\boldsymbol{\lambda}}^*)^T \bar{\mathbf{f}}(\mathbf{x}^*(t), \mathbf{u}^*(t))$  erfüllt ist, wobei  $\bar{\boldsymbol{\lambda}}$  und  $\bar{\mathbf{f}}$  gemäß (4.140) definiert sind, und folgende Eigenschaften gelten:

- (a) Die optimale Lösung  $\mathbf{u}^*(t)$  minimiert die Funktion  $H(\mathbf{x}^*(t), \mathbf{u}(t), \bar{\boldsymbol{\lambda}}^*(t))$  für alle Zeiten  $t_0 \leq t \leq t_1^*$  in der Menge der Stellgrößenbeschränkungen  $U$ , d. h.

$$H(\mathbf{x}^*(t), \mathbf{v}, \bar{\boldsymbol{\lambda}}^*(t)) \geq H(\mathbf{x}^*(t), \mathbf{u}^*(t), \bar{\boldsymbol{\lambda}}^*(t)), \quad \forall \mathbf{v} \in U. \quad (4.144)$$

- (b) Es gilt für alle Zeiten  $t_0 \leq t \leq t_1^*$

$$\bar{\lambda}_{n+1}^*(t) = \text{konst.} \geq 0 \quad (4.145a)$$

$$H(\mathbf{x}^*(t), \mathbf{u}^*(t), \bar{\boldsymbol{\lambda}}^*(t)) = \text{konst.} \geq 0. \quad (4.145b)$$

- (c) Es gilt die folgende Transversalitätsbedingung (für  $t_1$  frei)

$$H(\mathbf{x}^*(t_1^*), \mathbf{u}^*(t_1^*), \bar{\boldsymbol{\lambda}}^*(t_1^*)) = 0. \quad (4.146)$$

Zur Berechnung der  $2n + m + 3$  Unbekannten  $(\mathbf{u}^*, \bar{\mathbf{x}}^*, \bar{\boldsymbol{\lambda}}^*, t_1^*)$  stehen  $2n + m + 3$  Bedingungen zur Verfügung, nämlich  $m$  algebraische Bedingungen aus der Forderung, dass gemäß (4.144) die Hamiltonfunktion  $H$  zu jedem Zeitpunkt  $t_0 \leq t \leq t_1^*$  am Punkt  $\mathbf{u}^*(t)$  in der Menge  $U$  ein Minimum aufweisen muss, eine algebraische Gleichung in Form der Transversalitätsbedingung (4.146) und  $2n + 2$  Differentialgleichungen für den erweiterten Zustand  $\bar{\mathbf{x}}^*$  gemäß (4.140) und den erweiterten adjungierten Zustand  $\bar{\boldsymbol{\lambda}}^*$  gemäß (4.143) mit den  $n + 1$  zugehörigen Anfangsbedingungen  $\bar{\mathbf{x}}^*(t_0) = [\mathbf{x}_0^T \ 0]^T$ , den  $n$  Endbedingungen  $\mathbf{x}^*(t_1^*) = \mathbf{x}_1$  sowie der Bedingung  $\bar{\lambda}_{n+1}^*(t_1^*) \geq 0$ . Für  $\bar{\lambda}_{n+1}^* = 0$  liegt ein *abnormaler Fall* vor, da dann wegen

$$H(\mathbf{x}^*, \mathbf{u}^*, \bar{\boldsymbol{\lambda}}^*) = (\bar{\boldsymbol{\lambda}}^*)^T \bar{\mathbf{f}}(\mathbf{x}^*, \mathbf{u}^*) = (\boldsymbol{\lambda}^*)^T \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*) + \underbrace{\bar{\lambda}_{n+1}^*}_{=0} l(\mathbf{x}^*, \mathbf{u}^*) \quad (4.147)$$

die Hamiltonfunktion  $H$  und damit die Minimierungsaufgabe (4.144) unabhängig vom Kostenfunktional  $l$  ist. In diesem Fall ist die Optimierungsaufgabe *nicht sinnvoll gestellt*. Für den *normalen Fall* ist  $\bar{\lambda}_{n+1}^* > 0$  und  $\boldsymbol{\lambda}^*$  ist bis auf einen multiplikativen Faktor definiert. In der Praxis verwendet man üblicherweise den Wert  $\bar{\lambda}_{n+1}^* = 1$  und erhält damit die Hamiltonfunktion wie sie bereits bei der Variationsrechnung verwendet wurde, siehe (4.95).

Die notwendigen Bedingungen dafür, dass die Hamiltonfunktion  $H(\mathbf{x}, \mathbf{u}, \bar{\boldsymbol{\lambda}})$  für  $\bar{\lambda}_{n+1}^* = 1$  bezüglich  $\mathbf{u}$  minimal ist, wie in (4.144) gefordert, entsprechen den notwendigen Bedingungen erster und zweiter Ordnung (4.96c) und (4.98)

$$\left( \frac{\partial}{\partial \mathbf{u}} H \right) (\mathbf{x}^*, \mathbf{u}^*, \boldsymbol{\lambda}^*) = \mathbf{0} \quad \text{und} \quad \left( \frac{\partial^2}{\partial \mathbf{u}^2} H \right) (\mathbf{x}^*, \mathbf{u}^*, \boldsymbol{\lambda}^*) \text{ ist positiv semi-definit} \quad (4.148)$$

für alle Zeiten  $t_0 \leq t \leq t_1^*$ , die im Rahmen der Variationsrechnung hergeleitet wurden. Trotz dieser Analogie ist das Minimumsprinzip von Pontryagin nach Satz 4.10 wesentlich

allgemeiner, da die Bedingung  $\frac{\partial}{\partial \mathbf{u}} H = \mathbf{0}$  im Allgemeinen dann nicht mehr gültig ist, wenn das Minimum von  $H$  am Rand der Menge  $U$  der Stellgrößenbeschränkungen liegt. Im Weiteren fordert man beim Minimumsprinzip von Pontryagin lediglich die Stetigkeit von  $l$  und  $\mathbf{f}$  bezüglich  $\mathbf{u}$ , wohingegen bei der Herleitung der Euler-Lagrange Gleichungen die stetige Differenzierbarkeit bezüglich  $\mathbf{u}$  gefordert wurde, siehe Satz 4.8.

*Beispiel 4.7 (Abnormaler Fall).* Gegeben ist das Optimalsteuerungsproblem

$$\min_{u(\cdot)} \int_0^1 l(x, u) dt \quad (4.149a)$$

$$\text{u.B.v. } \dot{x} = u, \quad x(0) = 0, \quad x(1) = 1 \quad (4.149b)$$

$$u(t) \in [0, 1]. \quad (4.149c)$$

Es existiert nur eine Steuerung  $u^*(t) = 1$ , die den Zustand  $x^*(t) = t$  von  $x^*(0) = 0$  nach  $x^*(1) = 1$  überführt. Somit ist die optimale Lösung unabhängig von der Wahl der Kostenfunktion  $l(x, u)$  und es liegt ein abnormaler Fall vor.

*Beispiel 4.8.* Gesucht ist das Minimum des Kostenfunktional

$$J(u) = \frac{1}{2} \int_0^1 u^2(t) dt \quad (4.150)$$

für das dynamische System

$$\dot{x} = -x + u, \quad x(0) = 1, \quad x(1) = 0 \quad (4.151)$$

unter Berücksichtigung der Stellgrößenbeschränkung  $-0.6 \leq u(t) \leq 0$  für alle  $0 \leq t \leq 1$ . Die Hamiltonfunktion  $H$  von Satz 4.10 für dieses Beispiel lautet

$$H(x, u, \bar{\lambda}) = \bar{\lambda}_1(-x + u) + \bar{\lambda}_2 \frac{1}{2} u^2 \quad (4.152)$$

und die adjungierten Zustände  $\bar{\lambda}$  erfüllen gemäß (4.143) und (4.145a) die Gleichungen

$$\frac{d}{dt} \bar{\lambda}_1^* = - \left( \frac{\partial}{\partial x} H \right) (x^*, u^*, \bar{\lambda}^*) = \bar{\lambda}_1^* \quad (4.153a)$$

$$\frac{d}{dt} \bar{\lambda}_2^* = 0. \quad (4.153b)$$

Aus (4.153) erhält man als Lösung

$$\bar{\lambda}_1^*(t) = C_1 e^t \quad \text{und} \quad \bar{\lambda}_2^*(t) = C_2 \quad (4.154)$$

für geeignete Konstanten  $C_1$  und  $C_2 \geq 0$ . Im Weiteren setzt man  $C_2 = 1$ , da ein *normaler Fall* vorliegt. Die optimale Lösung  $u^*$  mit  $-0.6 \leq u^* \leq 0$  muss gemäß

(4.144) der Ungleichung

$$H(x^*(t), v, \bar{\lambda}^*(t)) \geq H(x^*(t), u^*(t), \bar{\lambda}^*(t)), \quad \forall v \in [-0.6, 0], \quad \forall t \in [0, 1] \quad (4.155)$$

genügen. Aus

$$\left( \frac{\partial}{\partial u} H \right) (x^*, u^*, \bar{\lambda}^*) = \bar{\lambda}_1^* + \bar{\lambda}_2^* u^* = \bar{\lambda}_1^* + u^* \quad (4.156)$$

folgt für die optimale Stellgröße unter Berücksichtigung der Stellgrößenbeschränkung

$$u^*(t) = \begin{cases} 0 & \text{für } \bar{\lambda}_1^* \leq 0 \\ -\bar{\lambda}_1^* = -C_1 e^t & \text{für } 0 < \bar{\lambda}_1^* < 0.6 \\ -0.6 & \text{für } \bar{\lambda}_1^* \geq 0.6 . \end{cases} \quad (4.157)$$

Man erkennt sofort, dass  $C_1 > 0$  gelten muss, denn für  $C_1 \leq 0$  ist  $\bar{\lambda}_1^* \leq 0$  und damit  $u^*(t) = 0$  für  $0 \leq t \leq 1$ , woraus aber wegen  $x^*(1) = e^{-1} \neq 0$  keine zulässige Lösung resultiert. Mit  $C_1 > 0$  muss deshalb die optimale Stellgröße  $u^*(t)$  zwischen  $-C_1 e^t$  und  $-0.6$  umschalten. Da  $\bar{\lambda}_1^*(t) = C_1 e^t$ ,  $C_1 > 0$  streng monoton steigend in  $t$  ist, setzt man eine stückweise stetige Steuerung, nämlich für das Zeitintervall  $0 \leq t \leq c^*$  die Steuerung  $u^*(t) = u_{(1)}^*(t) = -C_1 e^t$  und für das Zeitintervall  $c^* < t \leq 1$  die Steuerung  $u^*(t) = u_{(2)}^*(t) = -0.6$ , mit einem Umschaltpunkt (Eckpunkt) zur Zeit  $t = c^*$  an. Für das erste Zeitintervall errechnet sich die Lösung von

$$\dot{x}_{(1)}^* = -x_{(1)}^* + u_{(1)}^*, \quad x_{(1)}^*(0) = 1 \quad (4.158)$$

zu

$$x_{(1)}^*(t) = -\frac{1}{2} C_1 e^t + e^{-t} \left( \frac{1}{2} C_1 + 1 \right) \quad (4.159)$$

und für das zweite Zeitintervall folgt aus

$$\dot{x}_{(2)}^* = -x_{(2)}^* + u_{(2)}^*, \quad x_{(2)}^*(1) = 0 \quad (4.160)$$

die Lösung zu

$$x_{(2)}^*(t) = \frac{3}{5} (e^{1-t} - 1) . \quad (4.161)$$

Nach (4.145b) muss die Hamiltonfunktion  $H(x^*(t), u^*(t), \bar{\lambda}^*(t))$  im gesamten Zeitintervall konstant sein, weshalb gilt

$$\bar{\lambda}_1^* (-x_{(1)}^* + u_{(1)}^*) + \frac{1}{2} (u_{(1)}^*)^2 = \bar{\lambda}_1^* (-x_{(2)}^* + u_{(2)}^*) + \frac{1}{2} (u_{(2)}^*)^2 \quad (4.162a)$$

$$-C_1 \left( 1 + \frac{1}{2} C_1 \right) = -1.631 C_1 + 0.18 \quad (4.162b)$$

und daraus errechnen sich zwei mögliche Lösungen, nämlich  $C_{1,1} = 0.436$  und  $C_{1,2} = 0.826$ . Der Zeitpunkt der Umschaltung (Eckpunkt)  $t = c^*$  folgt aus der *Stetigkeitsbedingung der Zustandsgröße*

$$x_{(1)}^*(c^*) = x_{(2)}^*(c^*) \quad (4.163)$$

zu  $c_1^* = 0.32$  für  $C_{1,1}$  und  $c_1^* = -0.32$  für  $C_{1,2}$ , weshalb für das betrachtete Zeitintervall  $0 \leq t \leq 1$  nur die Lösung für  $C_1 = C_{1,1} = 0.436$  in Frage kommt.

Im nächsten Schritt wird gezeigt, wie sich das Minimumsprinzip von Pontryagin nach Satz 4.10 ändert, wenn die Endbedingung  $\mathbf{x}(t_1) = \mathbf{x}_1$  durch eine *Endbeschränkung* der Form  $\mathbf{x}(t_1) \in \mathcal{X}_1$  mit der glatten Mannigfaltigkeit  $\mathcal{X}_1$  der Dimension  $n - p \leq n$  ersetzt wird. An dieser Stelle sei erwähnt (siehe Abschnitt 3.1.1), dass eine  $(n - p)$ -dimensionale Mannigfaltigkeit meist durch  $p$  Gleichungen der Form

$$\psi_k(\mathbf{x}) = 0, \quad k = 1, \dots, p \quad (4.164)$$

beschrieben wird. Der Tangentialraum  $\mathcal{T}_{\check{\mathbf{x}}}\mathcal{X}_1$  an der Stelle  $\mathbf{x} = \check{\mathbf{x}}$  ist dann in der Form

$$\mathcal{T}_{\check{\mathbf{x}}}\mathcal{X}_1 = \left\{ \mathbf{d} : \left( \frac{\partial}{\partial \mathbf{x}} \psi_k(\check{\mathbf{x}}) \right) \mathbf{d} = 0, k = 1, \dots, p \right\} \quad (4.165)$$

definiert.

**Satz 4.11 (Minimumsprinzip von Pontryagin, Transversalitätsbedingungen).** *Gesucht ist die Stellgröße  $\mathbf{u} \in (\hat{C}_U[t_0, t_1])^m$  so, dass das Kostenfunktional (Lagrange-Form)*

$$J(\mathbf{u}) = \int_{t_0}^{t_1} l(\mathbf{x}(t), \mathbf{u}(t)) dt \quad (4.166)$$

*unter den Gleichungsbeschränkungen (dynamisches System)*

$$\dot{\mathbf{x}} - \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) = \mathbf{0}, \quad \mathbf{x}(t_0) = \mathbf{x}_0, \quad \mathbf{x}(t_1) \in \mathcal{X}_1 \quad (4.167)$$

*mit fester Anfangszeit  $t_0$ , freier Endzeit  $t_1 \ll T$  und der glatten  $(n - p)$ -dimensionalen Mannigfaltigkeit  $\mathcal{X}_1$  minimiert wird. Dabei wird angenommen, dass  $l$  und  $\mathbf{f}$  stetig in  $\mathbf{x}$  und  $\mathbf{u}$  und stetig differenzierbar bezüglich  $\mathbf{x}$  für alle  $(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^n \times \mathbb{R}^m$  sind. Weiters sei  $(\mathbf{u}^*, t_1^*) \in (\hat{C}_U[t_0, t_1^*])^m \times [t_0, T]$  die optimale Lösung des Optimierungsproblems und  $\mathbf{x}^*$  die zugehörige Lösung von (4.167). Dann existiert ein  $\bar{\boldsymbol{\lambda}}^* \in (\hat{C}^1[t_0, t_1^*])^{n+1}$ ,  $\bar{\boldsymbol{\lambda}}^* = [(\boldsymbol{\lambda}^*)^T \quad \bar{\lambda}_{n+1}^*]^T = [\bar{\lambda}_1^* \quad \dots \quad \bar{\lambda}_{n+1}^*]^T \neq [0 \quad \dots \quad 0]^T$  so, dass die Beziehungen (4.143)–(4.146) von Satz 4.10 erfüllt sind. Im Weiteren ist  $\boldsymbol{\lambda}^*(t_1^*) = [\lambda_1^*(t_1^*) \quad \dots \quad \lambda_n^*(t_1^*)]^T$  orthogonal zum Tangentialraum  $\mathcal{T}_{\mathbf{x}^*(t_1^*)}\mathcal{X}_1$ , d. h. es gelten die Transversalitätsbedingungen*

$$(\boldsymbol{\lambda}^*)^T(t_1^*) \mathbf{d} = 0, \quad \forall \mathbf{d} \in \mathcal{T}_{\mathbf{x}^*(t_1^*)}\mathcal{X}_1. \quad (4.168)$$

Nach Satz 4.11 und (4.165) muss  $\boldsymbol{\lambda}^*(t_1^*)$  sich also als Linearkombination von  $\left( \frac{\partial}{\partial \mathbf{x}} \psi_k \right)(\mathbf{x}_1^*)$ ,

$k = 1, \dots, p$  in der Form

$$\boldsymbol{\lambda}^*(t_1^*) = \sum_{k=1}^p \mu_k \left( \frac{\partial}{\partial \mathbf{x}} \psi_k \right)^T (\mathbf{x}_1^*), \quad \mathbf{x}_1^* = \mathbf{x}^*(t_1^*) \quad (4.169)$$

mit dem Lagrange-Multiplikator  $\boldsymbol{\mu} = [\mu_1 \ \dots \ \mu_p]^T \in \mathbb{R}^p$  darstellen lassen. Die Bedingung

$$\text{rang} \left( \frac{\partial}{\partial \mathbf{x}} \boldsymbol{\psi} \right) (\mathbf{x}_1^*) = p, \quad \boldsymbol{\psi} = [\psi_1 \ \dots \ \psi_p]^T \quad (4.170)$$

entspricht der CQ (constraint qualification) Bedingung der statischen Optimierung mit Beschränkungen, siehe auch (3.4).

Für *zeitvariante nichtlineare Systeme*

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (4.171)$$

führt man eine weitere Zustandsgröße der Form  $x_{n+1} = t$  ein und entwirft für das erweiterte System

$$\frac{d}{dt} \underbrace{\begin{bmatrix} \mathbf{x} \\ x_{n+1} \end{bmatrix}}_{\mathbf{x}_e} = \underbrace{\begin{bmatrix} \mathbf{f}(x_{n+1}, \mathbf{x}, \mathbf{u}) \\ 1 \end{bmatrix}}_{\mathbf{f}_e(\mathbf{x}_e, \mathbf{u})} \quad (4.172)$$

die optimale Steuerung gemäß Satz 4.10 und Satz 4.11. Dabei wird vorausgesetzt, dass  $\mathbf{f}$  und  $l$  stetig differenzierbar in  $t$  sind.

### 4.2.5 Minimumsprinzip für eingangsaffine Systeme

Den weiteren Betrachtungen liege das *eingangsaffine System*

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) = \mathbf{f}_0(\mathbf{x}) + \sum_{i=1}^m \mathbf{f}_i(\mathbf{x}) u_i \quad (4.173)$$

mit den Stellgrößenbeschränkungen der Form

$$\mathbf{u} \in U = [\mathbf{u}^-, \mathbf{u}^+] \quad \text{bzw.} \quad u_i \in [u_i^-, u_i^+], \quad i = 1, \dots, m \quad (4.174)$$

zugrunde.

#### 4.2.5.1 Kostenfunktional mit verbrauchsoptimalem Anteil

In der Literatur findet man häufig im Zusammenhang mit dem Entwurf von *verbrauchsoptimalen Steuerungen* Kostenfunktionale der Form

$$J(\mathbf{u}) = \int_{t_0}^{t_1} l_0(\mathbf{x}) + \sum_{i=1}^m r_i |u_i| \, dt, \quad r_i > 0. \quad (4.175)$$

Die zu (4.173) und (4.175) zugehörige Hamiltonfunktion lautet (mit  $\bar{\lambda}_{n+1} = 1$ )

$$H(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) = l_0(\mathbf{x}) + \sum_{i=1}^m r_i |u_i| + \boldsymbol{\lambda}^T \left( \mathbf{f}_0(\mathbf{x}) + \sum_{i=1}^m \mathbf{f}_i(\mathbf{x}) u_i \right). \quad (4.176)$$

Da der Anteil  $l_0(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{f}_0(\mathbf{x})$  in  $H$  unabhängig von  $\mathbf{u}$  ist, kann er im Minimierungsproblem (4.144) vernachlässigt werden. Die Minimierung von  $H$

$$\min_{u_i \in [u_i^-, u_i^+]} H_i(u_i) = r_i |u_i| + q_i(\mathbf{x}, \boldsymbol{\lambda}) u_i, \quad q_i(\mathbf{x}, \boldsymbol{\lambda}) = \boldsymbol{\lambda}^T \mathbf{f}_i(\mathbf{x}) \quad (4.177)$$

kann nun für jedes  $u_i$ ,  $i = 1, \dots, m$ , separat durchgeführt werden. Der Term  $q_i(\mathbf{x}, \boldsymbol{\lambda})$  spielt eine maßgebliche Rolle bei der Lösung dieses Problems. Abbildung 4.9 illustriert die unterschiedlichen Fälle a)–d), mit denen die optimale Stellgröße  $\mathbf{u}^*(t)$

$$u_i^* = \begin{cases} u_i^- & \text{falls } q_i(\mathbf{x}, \boldsymbol{\lambda}) > r_i \\ 0 & \text{falls } q_i(\mathbf{x}, \boldsymbol{\lambda}) \in (-r_i, r_i), \\ u_i^+ & \text{falls } q_i(\mathbf{x}, \boldsymbol{\lambda}) < -r_i \end{cases}, \quad i = 1, \dots, m \quad (4.178)$$

komponentenweise bestimmt ist. Ein kritischer Fall liegt vor, falls auf einem Subintervall  $I_s \subset [t_0, t_1]$  die Bedingung  $q_i(\mathbf{x}(t), \boldsymbol{\lambda}(t)) = \pm r_i$  identisch erfüllt ist. Die optimale Stellgröße  $u_i^*$  ist dann nicht mehr eindeutig aus der Minimierungsbedingung (4.177) bestimmbar. Dieser Fall wird als *singulär* bezeichnet und im Abschnitt 4.2.6 näher erläutert.

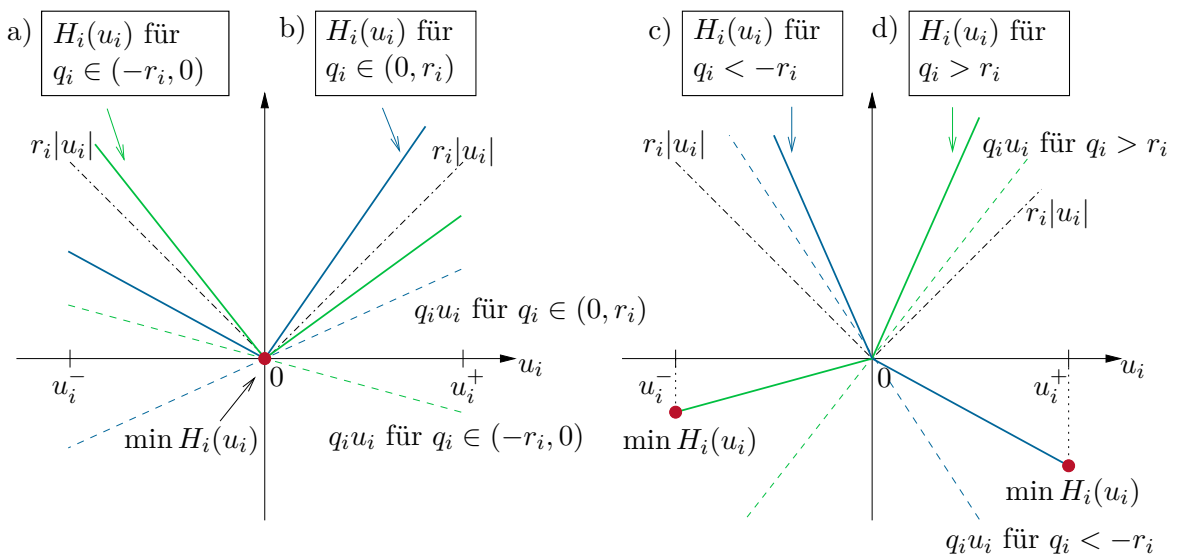


Abbildung 4.9: Verbrauchsoptimaler Fall: Grafische Veranschaulichung von (4.177).

#### 4.2.5.2 Kostenfunktional mit energieoptimalem Anteil

Unter dem Begriff *energieoptimale Steuerung* wird häufig die Minimierung eines Kostenfunktional der Form

$$J(\mathbf{u}) = \int_{t_0}^{t_1} l_0(\mathbf{x}) + \frac{1}{2} \sum_{i=1}^m r_i u_i^2 dt, \quad r_i > 0 \quad (4.179)$$

verstanden. Analog zum vorherigen Fall kann die Minimierung der Hamiltonfunktion  $H$

$$\min_{u_i \in [u_i^-, u_i^+]} H_i(u_i) = \frac{1}{2} r_i u_i^2 + q_i(\mathbf{x}, \boldsymbol{\lambda}) u_i, \quad q_i(\mathbf{x}, \boldsymbol{\lambda}) = \boldsymbol{\lambda}^T \mathbf{f}_i(\mathbf{x}) \quad (4.180)$$

wiederum für jedes  $u_i$ ,  $i = 1, \dots, m$ , separat erfolgen. Durch den quadratischen Term mit  $r_i > 0$  hat die Funktion  $H_i(u_i)$  stets ein *unbeschränktes Minimum* an der Stelle

$$u_i^0 = -\frac{1}{r_i} q_i(\mathbf{x}, \boldsymbol{\lambda}). \quad (4.181)$$

Falls  $u_i^0$  innerhalb des zulässigen Intervalls  $[u_i^-, u_i^+]$  liegt, ist die optimale Lösung von (4.173), (4.174) und (4.179) durch  $u_i^* = u_i^0$  gegeben. Falls  $u_i^0$  außerhalb von  $[u_i^-, u_i^+]$  liegt, so befindet sich das Minimum von  $H_i(u_i)$  stets an den Grenzen von  $U$ , also bei  $u_i^-$  oder  $u_i^+$ , da  $H_i(u_i)$  für  $u_i^0 < u_i^-$  (bzw.  $u_i^0 > u_i^+$ ) im Intervall  $[u_i^-, u_i^+]$  *streng monoton steigend (fallend)* ist, siehe Abbildung 4.10. Somit ist die optimale Stellgröße  $\mathbf{u}^*(t)$  komponentenweise wie folgt definiert

$$u_i^* = \begin{cases} u_i^- & \text{falls } u_i^0 \leq u_i^- \\ u_i^0 & \text{falls } u_i^0 \in (u_i^-, u_i^+), \\ u_i^+ & \text{falls } u_i^0 \geq u_i^+ \end{cases}, \quad i = 1, \dots, m. \quad (4.182)$$

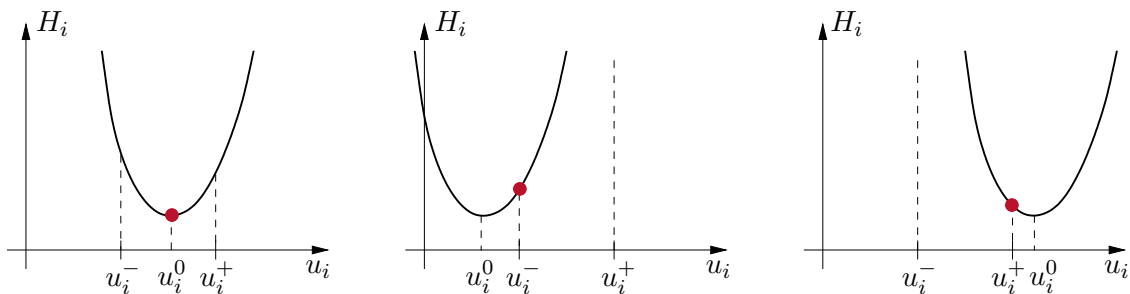


Abbildung 4.10: Energieoptimaler Fall: Grafische Veranschaulichung von (4.180).

#### 4.2.5.3 Zeitoptimales Kostenfunktional

Für zeitoptimale Probleme lautet das Kostenfunktional

$$J(\mathbf{u}) = \int_{t_0}^{t_1} 1 \, dt = t_1 - t_0 \quad (4.183)$$

und die Hamiltonfunktion lässt sich in der Form (mit  $\bar{\lambda}_{n+1} = 1$ )

$$H(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) = 1 + \boldsymbol{\lambda}^T \left( \mathbf{f}_0(\mathbf{x}) + \sum_{i=1}^m \mathbf{f}_i(\mathbf{x}) u_i \right) \quad (4.184)$$

anschreiben. Minimiert man die Hamiltonfunktion  $H$  wiederum für jedes  $u_i$ ,  $i = 1, \dots, m$  separat

$$\min_{u_i \in [u_i^-, u_i^+]} H_i(u_i) = q_i(\mathbf{x}, \boldsymbol{\lambda})u_i, \quad q_i(\mathbf{x}, \boldsymbol{\lambda}) = \boldsymbol{\lambda}^T \mathbf{f}_i(\mathbf{x}), \quad (4.185)$$

so erhält man die optimale Stellgröße  $\mathbf{u}^*$  direkt in Abhängigkeit des Vorzeichens von  $q_i(\mathbf{x}, \boldsymbol{\lambda})$ ,  $i = 1, \dots, m$  in der Form

$$u_i^* = \begin{cases} u_i^- & \text{falls } q_i(\mathbf{x}, \boldsymbol{\lambda}) > 0 \\ u_i^+ & \text{falls } q_i(\mathbf{x}, \boldsymbol{\lambda}) < 0 \end{cases}, \quad i = 1, \dots, m. \quad (4.186)$$

Diese Steuerung wird in der Literatur häufig auch als *Bang-Bang-Steuerung* bezeichnet, da lediglich zwischen den Maximal- und Minimalwerten des Stellgrößenbereiches hin- und hergeschaltet wird. Ein singulärer Fall liegt vor, falls  $q_i(\mathbf{x}(t), \boldsymbol{\lambda}(t))$  auf einem Subintervall  $I_s \subset [t_0, t_1]$  identisch Null wird. Die Hamiltonfunktion  $H$  ist dann *unabhängig von  $u_i$* , sodass  $H$  für jeden beliebigen Wert von  $u_i$  trivialerweise minimal ist. Die Minimumsforderung (4.185) ist damit zwar erfüllt, liefert aber keine Informationen über die Wahl von  $u_i$ .

In der Praxis wird der singuläre Fall oft durch einen zusätzlichen *Regularisierungsterm*

$$J(\mathbf{u}) = \int_{t_0}^{t_1} 1 + \frac{1}{2} \sum_{i=1}^m r_i u_i^2 dt \quad (4.187)$$

vermieden, wobei  $r_i > 0$  hinreichend klein gewählt wird, um annähernd Zeitoptimalität zu erzielen. Durch den Regularisierungsterm entspricht das Kostenfunktional dem Fall (4.179), so dass die optimale Stellgröße  $\mathbf{u}^*$  gemäß (4.182) berechnet werden kann.

**Beispiel 4.9 (Doppelintegrator).** Ein bekanntes Beispiel zur Veranschaulichung des Minimumsprinzips von Pontryagin ist die *zeitminimale* Überführung eines Doppelintegrators in den Ursprung  $\mathbf{x}(t_1) = \mathbf{0}$ . Das Optimalsteuerungsproblem mit dem Zustand  $\mathbf{x} = [x_1 \quad x_2]^T$  kann wie folgt formuliert werden

$$\min_{u(\cdot)} \int_{t_0=0}^{t_1} dt = t_1 \quad (4.188a)$$

$$\text{u.B.v. } \dot{x}_1 = x_2, \quad \dot{x}_2 = u \quad (4.188b)$$

$$\mathbf{x}(0) = \mathbf{x}_0, \quad \mathbf{x}(t_1) = \mathbf{0} \quad (4.188c)$$

$$|u(t)| \leq 1 \quad \forall t \in [0, t_1]. \quad (4.188d)$$

Mit der Hamiltonfunktion ( $\bar{\lambda}_{n+1} = 1$ )  $H(\mathbf{x}, u, \boldsymbol{\lambda}) = 1 + \lambda_1 x_2 + \lambda_2 u$  ergeben sich die adjungierten Zustände  $\boldsymbol{\lambda}^* = [\lambda_1^* \quad \lambda_2^*]^T$  aus (4.143) zu

$$\dot{\lambda}_1^* = 0 \quad \Rightarrow \quad \lambda_1^*(t) = C_1 \quad (4.189a)$$

$$\dot{\lambda}_2^* = -\lambda_1^* \quad \Rightarrow \quad \lambda_2^*(t) = -C_1 t + C_2, \quad (4.189b)$$



wobei  $C_1$  und  $C_2$  Integrationskonstanten darstellen. Die Minimierungsbedingung (4.144) für die Hamiltonfunktion führt auf die optimale Stellgröße

$$u^*(t) = \begin{cases} +1 & \text{falls } \lambda_2^* < 0 \\ -1 & \text{falls } \lambda_2^* > 0. \end{cases} \quad (4.190)$$

Der *singuläre Fall*, d. h.  $\lambda_2^*(t) = 0$  auf einem nicht verschwindenden Subintervall  $t \in I_s$ , kann nicht auftreten, da dann aufgrund von (4.189)  $\lambda_2^*(t) = 0$  und  $\lambda_1^*(t) = 0$  auf dem Gesamtintervall  $[0, t_1]$  gelten müsste. Dies widerspricht aber der Transversalitätsbedingung (4.146) für die *freie Endzeit*  $t_1$

$$H(\mathbf{x}^*, u^*, \boldsymbol{\lambda}^*)|_{t=t_1^*} = 1 + \lambda_1^*(t_1^*)x_2^*(t_1^*) + \lambda_2^*(t_1^*)u^*(t_1^*) = 0. \quad (4.191)$$

Der Fall  $\lambda_2^*(t) = 0$  kann also nur zu einem diskreten Zeitpunkt auftreten, was einem Umschalten von  $u^*(t)$  zwischen  $-1$  und  $+1$  entspricht. Aufgrund der Affinität von  $\lambda_2^*(t)$  in  $t$  gibt es *maximal einen Umschaltzeitpunkt*  $t = t_s$  im Zeitintervall  $[0, t_1^*]$ , an dem  $u^*(t)$  zwischen  $-1$  und  $+1$  wechselt. Somit existieren lediglich *vier mögliche Schaltsequenzen*  $\{+1\}$ ,  $\{-1\}$ ,  $\{+1, -1\}$ ,  $\{-1, +1\}$ , die für eine optimale Lösung in Frage kommen. Da  $u$  stets auf einem Zeitintervall konstant ist, stellen die Trajektorien in der  $(x_1, x_2)$ -Ebene *Parabeln* dar, wie in der folgenden Aufgabe gezeigt werden soll.

**Aufgabe 4.6.** Zeigen Sie, dass die Lösung von (4.188b) für  $u(t) = \pm 1 = \text{konst.}$  die folgende Parabelgleichung erfüllt

$$x_1 = \frac{x_2^2}{2u} + c, \quad u = \pm 1. \quad (4.192)$$

Das Optimierungsziel ist das *schnellstmögliche Erreichen des Ursprungs*  $\mathbf{x}(t_1) = \mathbf{0}$  ausgehend von einem beliebigen Anfangspunkt  $\mathbf{x}(0) = \mathbf{x}_0$ . Der Ursprung ist aber nur über die *Schaltkurve*  $x_1 = x_2^2/2$  für  $u = 1$  bzw.  $x_1 = -x_2^2/2$  für  $u = -1$  erreichbar, siehe Abbildung 4.11. Diese beiden Fälle können in der Funktion

$$S(x_2) = -\frac{1}{2}x_2|x_2| \quad (4.193)$$

zusammengefasst werden. Angesichts der Tatsache, dass lediglich die Schaltsequenzen  $\{+1\}$ ,  $\{-1\}$ ,  $\{+1, -1\}$ ,  $\{-1, +1\}$  für  $u$  in Frage kommen, gibt es nur eine Möglichkeit, um  $\mathbf{x}(0) = \mathbf{x}_0$  schnellstmöglich zum Ursprung  $\mathbf{x}(t_1) = \mathbf{0}$  zu bringen:

- Falls  $x_{1,0} = S(x_{2,0})$  gilt, ist keine Umschaltung notwendig, und  $\mathbf{x}(t_1) = \mathbf{0}$  wird direkt über die Schaltkurve mit  $u(t) = 1$  oder  $u(t) = -1$  erreicht, siehe Abbildung 4.11.
- Falls  $\mathbf{x}_0$  nicht auf der Schaltkurve liegt, d. h.  $x_{1,0} < S(x_{2,0})$  oder  $x_{1,0} > S(x_{2,0})$ , ist genau eine Umschaltung notwendig, um die Schaltkurve zu erreichen und anschließend zum Ursprung zu laufen, siehe Abbildung 4.12.

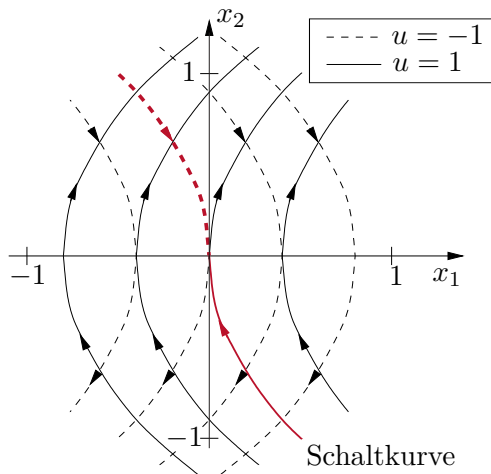


Abb. 4.11: Mögliche Trajektorien des Doppelintegrators in der  $(x_1, x_2)$ -Ebene.

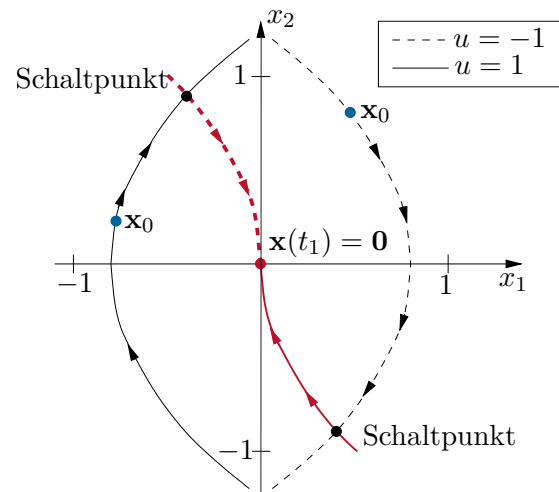


Abb. 4.12: Optimale Umschaltung für verschiedene Anfangspunkte  $\mathbf{x}_0$ .

Das optimale Stellgesetz lautet also

$$u^*(t) = \begin{cases} +1 & \text{falls } x_1 < S(x_2) \\ +1 & \text{falls } x_1 = S(x_2) \text{ und } x_1 > 0 \\ -1 & \text{falls } x_1 > S(x_2) \\ -1 & \text{falls } x_1 = S(x_2) \text{ und } x_1 < 0, \end{cases} \quad (4.194)$$

mit dessen Hilfe der Schaltzeitpunkt  $t_s$  und die minimale Endzeit  $t_1^*$  berechnet werden können.

**Aufgabe 4.7.** Verifizieren Sie, dass der optimale Umschaltzeitpunkt  $t_s$  und die minimale Endzeit  $t_1^*$  wie folgt definiert sind

$$t_s = \begin{cases} x_{2,0} + \sqrt{\frac{1}{2}x_{2,0}^2 + x_{1,0}} & \text{falls } x_{1,0} > S(x_{2,0}) \\ -x_{2,0} + \sqrt{\frac{1}{2}x_{2,0}^2 - x_{1,0}} & \text{falls } x_{1,0} < S(x_{2,0}) \end{cases} \quad (4.195)$$

$$t_1^* = \begin{cases} x_{2,0} + \sqrt{2x_{2,0}^2 + 4x_{1,0}} & \text{falls } x_{1,0} > S(x_{2,0}) \\ -x_{2,0} + \sqrt{2x_{2,0}^2 - 4x_{1,0}} & \text{falls } x_{1,0} < S(x_{2,0}) \\ |x_{2,0}| & \text{falls } x_{1,0} = S(x_{2,0}). \end{cases} \quad (4.196)$$

Abbildung 4.13 stellt die zeitoptimalen Trajektorien für den Doppelintegrator für verschiedene Anfangswerte  $\mathbf{x}_0 = [x_{1,0} \quad x_{2,0}]^T$  dar.

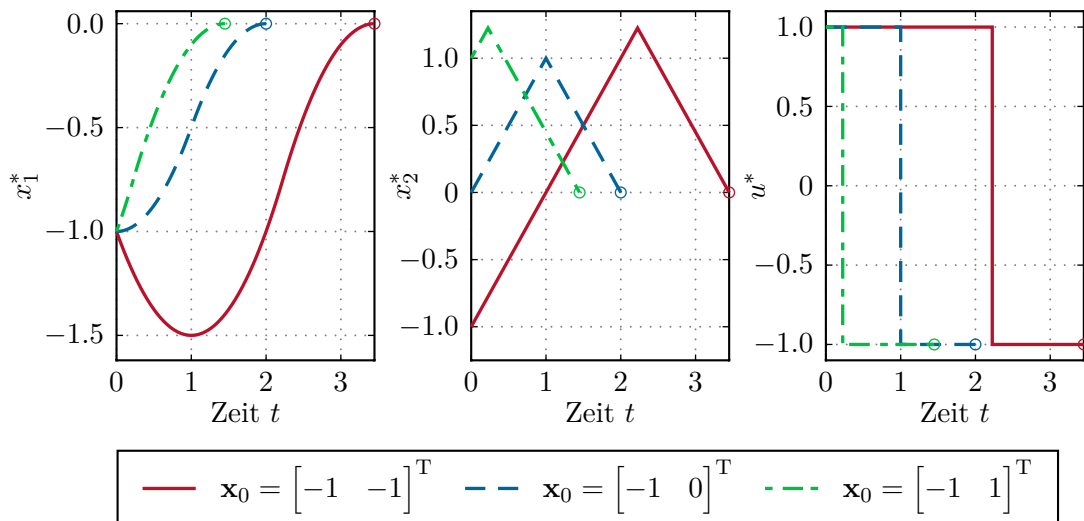


Abbildung 4.13: Zeitoptimale Trajektorien des Doppelintegrators für verschiedene Anfangswerte  $\mathbf{x}_0$ .

**Aufgabe 4.8.** Implementieren Sie das doppelintegrierende System (4.188b) mit dem optimalen Stellgesetz (4.194) in MATLAB/SIMULINK und verifizieren Sie die Ergebnisse in Abbildung 4.13 für verschiedene Anfangswerte  $\mathbf{x}_0$ . Verwenden Sie  $t_1^*$  gemäß (4.196) als Zeithorizont für die Simulation.

#### 4.2.6 Der singuläre Fall

Wie bereits in den Abschnitten 4.2.5.1 und 4.2.5.3 diskutiert, kann es passieren, dass auf einem endlichen Subintervall  $I_s \subset [t_0, t_1]$  die optimale Stellgröße  $\mathbf{u}^*$  nicht aus der Minimierungsbedingung (4.144) bestimmt werden kann. Zur Verdeutlichung dieser Problematik soll im Weiteren das Optimalsteuerungsproblem

$$\min_{u(\cdot)} J(u) = \int_{t_0}^{t_1} l_0(\mathbf{x}) + l_1(\mathbf{x})u \, dt \quad (4.197a)$$

$$\text{u.B.v.} \quad \dot{\mathbf{x}} = \mathbf{f}_0(\mathbf{x}) + \mathbf{f}_1(\mathbf{x})u, \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (4.197b)$$

$$u \in \hat{C}_U[t_0, t_1] \quad (4.197c)$$

mit skalarer und affin auftretender Stellgröße  $u(t)$  betrachtet werden. Die grundsätzliche Vorgehensweise ist aber auch auf allgemeinere Optimalsteuerungsprobleme anwendbar. Die Hamiltonfunktion

$$H(t, \mathbf{x}, u, \boldsymbol{\lambda}) = l_0(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{f}_0(\mathbf{x}) + (l_1(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{f}_1(\mathbf{x}))u \quad (4.198)$$

ist auf Grund der speziellen Struktur von (4.197) ebenfalls affin in  $u$ . Die Funktion

$$\zeta(t) = \left( \frac{\partial}{\partial u} H \right) (t, \mathbf{x}, u, \boldsymbol{\lambda}) = l_1(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{f}_1(\mathbf{x}) \quad (4.199)$$

wird als *Schaltfunktion* bezeichnet, wobei  $\zeta^*(t) = \zeta(t)|_{\mathbf{x}=\mathbf{x}^*(t), \boldsymbol{\lambda}=\boldsymbol{\lambda}^*(t)}$ . Wenn  $\zeta^*(t)$  die Schaltfunktion entlang der optimalen Lösung  $\mathbf{x}^*(t)$ ,  $\boldsymbol{\lambda}^*(t)$  auf einem endlichen Zeitintervall  $I_s \subset [t_0, t_1]$  identisch Null ist, liefert die Minimierungsbedingung (4.144) keine Aussage für die optimale Stellgröße  $u^*(t)$ ,  $t \in I_s$ . Man spricht in diesem Zusammenhang von einem *singulären Pfad*. Entlang des singulären Pfades ist die Minimierungsbedingung (4.144) für alle zulässigen Stellgrößen erfüllt. Um dennoch eine Stellgröße  $u^*(t)$  ermitteln zu können, verfolgt man die Idee, dass im Zeitintervall  $I_s$  auch sämtliche Ableitungen  $\frac{d^k}{dt^k} \zeta^*(t)$ ,  $k = 1, 2, \dots$  identisch verschwinden müssen. Man sucht nun die *kleinste positive natürliche Zahl*  $\bar{k}$  so, dass gilt

$$\frac{d^{\bar{k}}}{dt^{\bar{k}}}(\zeta^*(t)) = 0 \quad (4.200a)$$

$$\frac{\partial}{\partial u} \left[ \frac{d^{\bar{k}}}{dt^{\bar{k}}}(\zeta^*(t)) \right] \neq 0. \quad (4.200b)$$

Man kann zeigen, dass diese Zahl  $\bar{k}$  eine *gerade Zahl* sein muss und nennt  $p = \bar{k}/2$  die *Ordnung des singulären Pfades*. Entlang eines singulären Pfades müssen die Zustandsgrößen  $\mathbf{x}^*(t)$  und die adjungierten Zustände  $\boldsymbol{\lambda}^*(t)$  auf einer Mannigfaltigkeit definiert durch die Gleichungen

$$\zeta^*(t) = 0 \quad (4.201a)$$

$$\frac{d}{dt}(\zeta^*(t)) = 0 \quad (4.201b)$$

$$\vdots \quad (4.201c)$$

$$\frac{d^{2p-1}}{dt^{2p-1}}(\zeta^*(t)) = 0 \quad (4.201d)$$

zu liegen kommen. Analog zur Legendre-Clebsch Bedingung (4.98) muss entlang eines singulären Pfades für alle Zeiten  $t \in I_s$  die sogenannte *generalisierte Legendre-Clebsch Bedingung*

$$(-1)^p \frac{\partial}{\partial u} \left[ \frac{d^{2p}}{dt^{2p}}(\zeta^*(t)) \right] \geq 0 \quad (4.202)$$

gelten.

*Beispiel 4.10.* Gesucht ist das Minimum des Kostenfunktional

$$J(u) = \frac{1}{2} \int_0^2 x_1^2(t) dt \quad (4.203)$$

für das dynamische System

$$\dot{x}_1 = x_2 + u \quad x_1(0) = 1 \quad x_1(2) = 0 \quad (4.204a)$$

$$\dot{x}_2 = -u \quad x_2(0) = 1 \quad x_2(2) = 0 \quad (4.204b)$$

unter Berücksichtigung der Stellgrößenbeschränkung  $-10 \leq u(t) \leq 10$  für alle  $0 \leq t \leq 2$ . Die Hamiltonfunktion  $H$  für dieses Beispiel lautet (mit  $\bar{\lambda}_{n+1} = \bar{\lambda}_3 = 1$ )

$$H(\mathbf{x}, u, \bar{\boldsymbol{\lambda}}) = \bar{\lambda}_1(x_2 + u) - \bar{\lambda}_2 u + \frac{1}{2}x_1^2 \quad (4.205)$$

und die adjungierten Zustände  $\bar{\boldsymbol{\lambda}}$  erfüllen gemäß (4.143) die Gleichungen

$$\frac{d}{dt}\bar{\lambda}_1^* = -\left(\frac{\partial}{\partial x_1}H\right)(\mathbf{x}^*, u^*, \bar{\boldsymbol{\lambda}}^*) = -x_1^* \quad (4.206a)$$

$$\frac{d}{dt}\bar{\lambda}_2^* = -\left(\frac{\partial}{\partial x_2}H\right)(\mathbf{x}^*, u^*, \bar{\boldsymbol{\lambda}}^*) = -\bar{\lambda}_1^* . \quad (4.206b)$$

Die optimale Lösung  $u^*$  mit  $-10 \leq u^* \leq 10$  muss der Ungleichung (siehe (4.144))

$$H(\mathbf{x}^*(t), v, \bar{\boldsymbol{\lambda}}^*(t)) \geq H(\mathbf{x}^*(t), u^*(t), \bar{\boldsymbol{\lambda}}^*(t)), \quad \forall v \in [-10, 10] \quad (4.207)$$

genügen und damit folgt für die optimale Stellgröße unter Berücksichtigung der Stellgrößenbeschränkung

$$u^*(t) = \begin{cases} 10 & \text{für } \bar{\lambda}_1^* < \bar{\lambda}_2^* \\ -10 & \text{für } \bar{\lambda}_1^* > \bar{\lambda}_2^* \\ \text{unbestimmt} & \text{für } \bar{\lambda}_1^* = \bar{\lambda}_2^* . \end{cases} \quad (4.208)$$

Für  $\bar{\lambda}_1^* = \bar{\lambda}_2^*$  tritt somit ein *singulärer Pfad* auf. Im nächsten Schritt wird nun gemäß (4.200) die Ordnung  $p$  des singulären Pfades über die Beziehungen

$$\left(\frac{\partial}{\partial u}H\right)(t, \mathbf{x}^*, u, \bar{\boldsymbol{\lambda}}^*) = \bar{\lambda}_1^* - \bar{\lambda}_2^* = 0 \quad (4.209a)$$

$$\frac{d}{dt}\left(\frac{\partial}{\partial u}H\right)(t, \mathbf{x}^*, u, \bar{\boldsymbol{\lambda}}^*) = \frac{d}{dt}\bar{\lambda}_1^* - \frac{d}{dt}\bar{\lambda}_2^* = -x_1^* + \bar{\lambda}_1^* = 0 \quad (4.209b)$$

$$\frac{d^2}{dt^2}\left(\frac{\partial}{\partial u}H\right)(t, \mathbf{x}^*, u, \bar{\boldsymbol{\lambda}}^*) = -x_2^* - u^* - x_1^* = 0 \quad (4.209c)$$

zu  $p = 1$  bestimmt. Aus (4.209c) folgt

$$u^* = -x_2^* - x_1^* \quad (4.210)$$

und aus

$$(-1)^1 \frac{\partial}{\partial u} \left[ \frac{d^2}{dt^2} \left( \frac{\partial}{\partial u} H \right) (t, \mathbf{x}^*, u, \bar{\boldsymbol{\lambda}}^*) \right] = 1 > 0 \quad (4.211)$$

erkennt man, dass die generalisierte Legendre-Clebsch Bedingung (4.202) für ein Minimum entlang des singulären Pfades erfüllt ist. Nach (4.145b) muss die Hamiltonfunktion  $H(\mathbf{x}^*(t), u^*(t), \bar{\boldsymbol{\lambda}}^*(t))$  im gesamten Zeitintervall konstant sein, weshalb gilt

$$\bar{\lambda}_1^* x_2^* + \frac{1}{2}(x_1^*)^2 + (\bar{\lambda}_1^* - \bar{\lambda}_2^*)u^* = C \quad (4.212)$$

mit einer Konstanten  $C$ . Entlang des singulären Pfades müssen die Zustandsgrößen  $\mathbf{x}^*(t)$  und die adjungierten Zustände  $\bar{\boldsymbol{\lambda}}^*(t)$  auf der durch (4.209a) und (4.209b) definierten Mannigfaltigkeit  $\bar{\lambda}_1^* = \bar{\lambda}_2^* = x_1^*$  liegen (siehe auch (4.201)), weshalb sich für diesen Fall (4.212) zu

$$x_1^* x_2^* + \frac{1}{2}(x_1^*)^2 = C \quad (4.213)$$

vereinfacht.

*Aufgabe 4.9.* Zeigen Sie, dass die optimale Stellgröße durch

$$u^*(t) = \begin{cases} 10 & \text{für } 0 \leq t \leq 0.299 \\ -x_2^* - x_1^* & \text{für } 0.299 < t < 1.927 \\ -10 & \text{für } 1.927 \leq t \leq 2 \end{cases} \quad (4.214)$$

gegeben ist.

## 4.3 Literatur

- [4.1] A. E. Bryson, Jr. und Y.-C. Ho, *Applied Optimal Control: Optimization, Estimation, and Control*. John Wiley & Sons, 1975.
- [4.2] R. F. Hartl, S. P. Sethi und R. G. Vickson, „A survey of the maximum principles for optimal control problems with state constraints“, *SIAM Review*, Bd. 37, Nr. 2, S. 181–218, 1995.
- [4.3] M. Papageorgiou, M. Leibold und M. Buss, *Optimierung: Statische, dynamische, stochastische Verfahren für die Anwendung*, 3. Aufl. Springer, 2012.
- [4.4] B. C. Chachuat, „Nonlinear and Dynamic Optimization: From Theory to Practice“, <http://infoscience.epfl.ch/record/111939>, besucht am 28.09.2015, 2007.
- [4.5] M. Athans und P. L. Falb, *Optimal Control: An Introduction to the Theory and Its Applications*, Ser. Lincoln Laboratory Publications. McGraw-Hill, 1966.
- [4.6] B. van Brunt, *The Calculus of Variations*, Ser. Universitext. Springer, 2004.
- [4.7] M. I. Kamien und N. L. Schwartz, *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*, 2. Aufl. North-Holland, 1992.
- [4.8] O. Föllinger, *Optimale Regelung und Steuerung*, Ser. Methoden der Regelungs- und Automatisierungstechnik. R. Oldenbourg Verlag, 1994.
- [4.9] D. S. Naidu, *Optimal Control Systems*, Ser. Electrical Engineering Series. CRC Press, 2003.
- [4.10] D. E. Kirk, *Optimal Control Theory: An Introduction*. Dover Publications, 2004.