Bemerkung 4.5: Für ein zeitdiskretes Kontrollsystem

$$x_{k+1} = f(x_k, u_k), \quad k = 0,1,2,\dots$$

(z.B. das Euler-Verfahren $f = \tilde{\varphi}_h$) und eine Kostenfunktion $g(x_k, u_k) > 0$ kann man analog zeigen, dass die optimale Wertefunktion

$$V(x) = \sup_{(u_k)} \sum_{k=0}^{\infty} \beta^k g(x_k, u_k)$$

($\beta < 1$, $x = x_0$) die Fixpunktgleichung (Optimalitätsprinzip)

$$\boxed{V(x) = \sup_{u \in U} \{ g(x,u) + \beta V(f(x,u)) \}}$$

erfüllt.

Satz 4.6: In jedem Punkt, in dem die optimale Wertefunktion differenzierbar ist, erfüllt sie die Hamilton-Jacobi-Bellman-Gleichung

$$\delta V(x) = \sup_{u \in U} \{ \nabla V(x) \cdot f(x,u) + g(x,u) \}$$

Beweis: Optimalitätsprinzip:

$$V(x) = \sup_{u \in \mathcal{U}} \left\{ \int_0^T e^{-\delta t} g(\varphi(t,x,u), u(t))\, dt + e^{-\delta T} V(\varphi(T,x,u)) \right\}$$

$$\Rightarrow \sup_{u \in \mathcal{U}} \left\{ \frac{1}{T} \int_0^T e^{-\delta t} g(\varphi(t,x,u), u(t))\, dt + \frac{e^{-\delta T} V(\varphi(T,x,u)) - V(x)}{T} \right\} = 0$$

Für $T \to 0$ wird daraus (falls V in x differenzierbar)

$$\sup_{u\in\mathcal{U}}\left\{ \frac{d}{d\tau}\Big|_{\tau=0} \int_0^\tau e^{-\delta t} g(\varphi(t,x,u),u(t))\,dt + \nabla V(x)\cdot \frac{d}{d\tau}\Big|_{\tau=0}\varphi(\tau,x,u) - \delta V(x)\right\} = 0$$

$\|$ ← Beschränkung auf konstante Kontrollfolgen

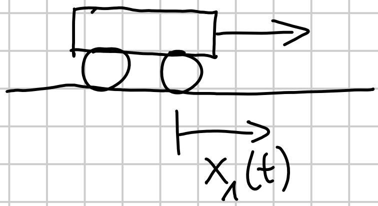
$$\sup_{u\equiv u_0}\left\{ \text{———} \;\;''\;\; \text{———} \right\} = 0$$

Für konstante Kontrollen $u \equiv u_0$ gilt

$$\frac{d}{d\tau}\varphi(\tau,x,u)\Big|_{\tau=0} = f(x,u_0) \quad \text{und} \quad \frac{d}{d\tau}\Big|_{\tau=0}\int_0^\tau e^{-\delta t} g(\varphi(t,x,u),u(t))\,dt = g(x,u_0)$$

$\blacksquare$

Beispiel:



Wagen mit Position x_1 und Geschwindigkeit $x_2 = \dot{x}_1$

Dynamik: $\ddot{x}_1(t) = a(t)$

System erster Ordnung

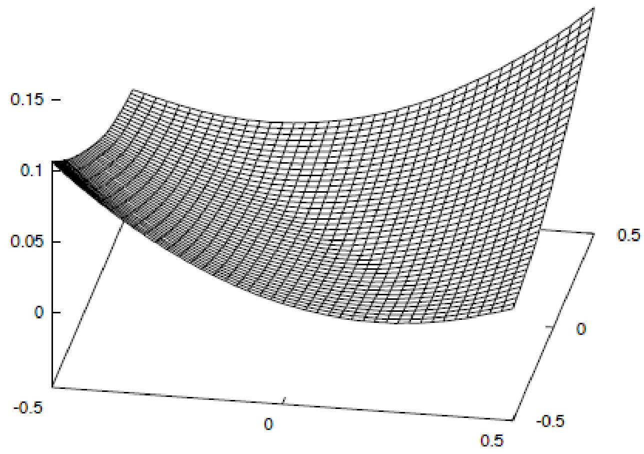
$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = a(t)$$

Setze: $a(t) = u_1(t) - (r + u_2(t))\,x_2(t)$

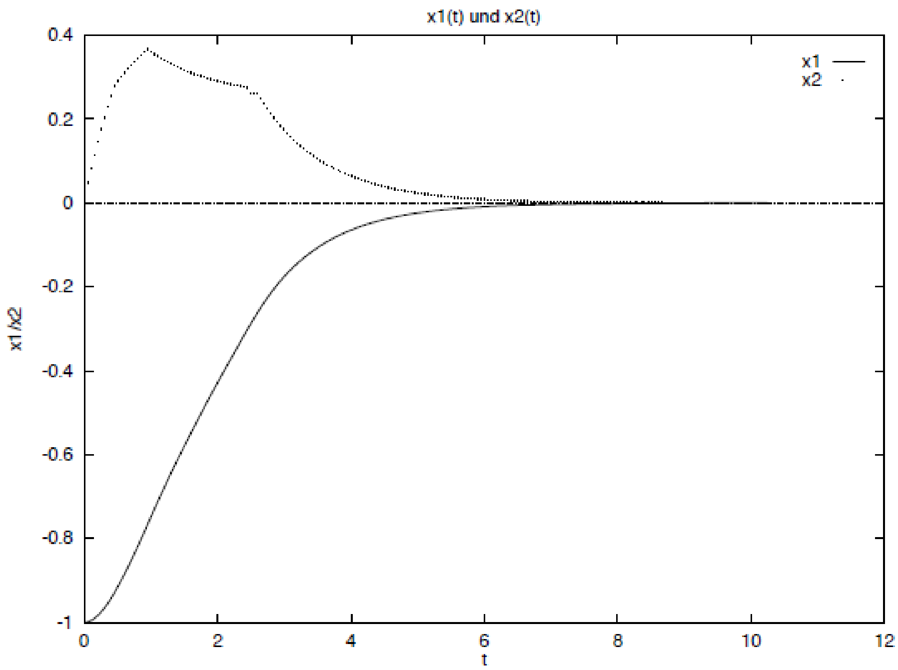
Motor (u_1), Reibung (r), Bremse (u_2)

Kostenfunktion: $g(x,u) = \|x\|_2^2 + u_1^2$, $u = (u_1,u_2)^T$

⇝ zur Minimierung von $J(x,u)$ steuere Wagen nach $(0,0)^T$.



optimale Wertefunktion
$u = (u_1, u_2)^T \in [-1,1] \times [0,1]$, $\delta = 0.1$, $r = 1$



optimale Trajektorie zum Anfangswert $(-1,0)^T$.

(4.4) Diskretisierung in der Zeit

Um die optimale Wertefunktion zu berechnen, diskretisieren wir jetzt zunächst das gegebene Problem $\dot{x} = f(x,u)$ (mit Kostenfunktion $g(x,u)$) in der Zeit:

Euler-Verfahren:

$$x_{k+1} = x_k + h f(x_k, u_k) =: \tilde{\varphi}_h(x_k, u_k) \quad , k = 0,1,2,\ldots$$

(zeit-diskretes Kontrollsystem)
zugehörige Kostenfunktion

$$\tilde{g}_h(x,u) := h\, g(x,u)$$

Das zugehörige zeitdiskrete diskontierte Kostenfunktional ist

$$J_h(x,(u_k)) = \sum_{k=0}^{\infty} \beta^k \tilde{g}_h(x_k, u_k) \qquad , \beta = 1 - \delta h$$

$(x = x_0)$ bzw. die zeitdiskrete optimale Wertefunktion

$$V_h(x) = \sup_{(u_k) \in U^{\mathbb{N}}} J(x,(u_k)).$$

Bemerkung: Für V_h gilt Satz 4.3 (Regularität der Wertefunktion) sinngemäß.

Frage: Fehler zwischen V_h und V?

Spezialfall: konvexes Optimalsteuerungsproblem in dem Sinne, dass die Mengen $f(x,U)$ und $g(x,U)$ für jedes $x \in \mathbb{R}^d$ konvex sind.

Satz 4.7: Falls das Kontrollsystem $\dot{x} = f(x,u)$ und die Kostenfunktion $g(x,u)$ die Voraussetzungen von Satz 3.1 erfüllen und konvex sind, gilt

$$|V(x) - V_h(x)| \leq K h^{\gamma}$$

mit der Stetigkeits-Konstanten γ aus Satz 4.3, einer Konstanten $K > 0$, $h \in [0, \frac{1}{\delta}]$ und alle $x \in \mathbb{R}^d$.

Beweis: zentral ist folgende Abschätzung: es gilt $\forall u \in \mathcal{U}$

$$|J(x,u) - J_h(x,(u_k))| \leq K(h^\gamma + h) \qquad (*)$$

für eine geeignet gewählte Kontrollfolge (u_k).

Sei (u_k) so gewählt, dass

$$f(x(hk),u_k) = \frac{1}{h}\int_{hk}^{h(k+1)} f(x(hk),u(t))\,dt \qquad (**)$$

$$g(x(hk),u_k) = \frac{1}{h}\int_{hk}^{h(k+1)} g(x(hk),u(t))\,dt \qquad (***)$$

wobei $x(t) = \varphi(t,x,u)$.

Es ist

$$J_h(x,(u_k)) = \sum_{k=0}^{\infty} \beta^k \tilde{g}_h(x_k,u_k)$$

$$= \sum_{k=0}^{\infty} \beta^k h g(x_k,u_k)$$

$$= \int_0^\infty \beta^{\lfloor t/h \rfloor} g(x_{\lfloor t/h \rfloor}, u_{\lfloor t/h \rfloor})\,dt$$

größte ganze Zahl $\leq t/h$

Also erhalten wir

$$|J(x,u) - J_h(x,(u_k))| \leq$$

(1) $\left| \int_0^\infty e^{-\delta t} g(x(t),u(t))\,dt - \int_0^\infty e^{-\delta t} g(x(h\lfloor t/h \rfloor),u(t))\,dt \right| +$

(2) $\left| \int_0^\infty e^{-\delta t} g(x(h\lfloor t/h \rfloor),u(t))\,dt - \int_0^\infty \beta^{\lfloor t/h \rfloor} g(x(h\lfloor t/h \rfloor),u(t))\,dt \right| +$

(3) $\left| \int_0^\infty \beta^{\lfloor t/h \rfloor} g(x(h\lfloor t/h \rfloor),u(t))\,dt - \int_0^\infty \beta^{\lfloor t/h \rfloor} g(x_{\lfloor t/h \rfloor}, u_{\lfloor t/h \rfloor})\,dt \right|$

Wir schätzen jetzt die Terme (1) bis (3) ab:

zu (1): Da $|f(x,u)| \le M$ folgt $\|x(t) - x(h\lfloor t/h \rfloor)\| \le Mh$,

also $(1) \le \int_0^\infty e^{-\delta t} L_g M h \, dt = K_1 h$

zu (2): Es ist $|\underbrace{\beta}_{=1-\delta h}^{\lfloor t/h \rfloor} - e^{-\delta t}| \le Che^{-\delta t}(\delta t + 1)$ $\boxed{\overset{\text{u}}{\text{Ü}}}$

damit folgt

$$(2) \le \int_0^\infty Che^{-\delta t}(\delta t + 1) M_g \, dt \le K_2 h$$

zu (3): aus (**) folgt

Lipschitz-K von f

$$\|x(h\lfloor t/h \rfloor) - x_{\lfloor t/h \rfloor}\| \le Che^{L_f t}$$

(vgl. Satz 3.2), aus (***) folgt

$$\int_0^\infty \beta^{\lfloor t/h \rfloor} g(x(h\lfloor t/h \rfloor), u(t)) \, dt = \int_0^\infty \beta^{\lfloor t/h \rfloor} g(x(h\lfloor t/h \rfloor), u_{\lfloor t/h \rfloor}) \, dt$$

und mit $\beta^{\lfloor t/h \rfloor} \le e^{1-\delta t}$ folgt

$$(3) \le \int_0^\infty e^{1-\delta t} L_g Che^{L_f t} \, dt \le K_3 h^\gamma$$

(analog zu Satz 4.3).

Daraus folgt (*) mit $K = \max\{K_1 + K_2, K_3\}$.
Wir zeigen jetzt noch, dass aus (*) die Aussage des Satzes folgt. Zunächst folgt aus der Konvexität von f und g, dass es eine Kontrollfolge gibt, die (**) und (***) erfüllt (vgl. den Beweis zu Satz 3.2). Aus (*) folgt nun

$$V_h(x) \geq J_h(x,(u_k)) \geq J(x,u) - K(h^\gamma + h)$$

und da u beliebig war, gilt die Ungleichung auch für das Supremum auf der rechten Seite – analog folgt die Ungleichung in die andere Richtung. ∎

(4.5) Werteiteration

Wir berechnen wir V_h?

Beobachtung: V_h ist Fixpunkt der Fixpunktgleichung

$$V(x) = \sup_{u \in U} \left\{ \tilde{g}_h(x,u) + \beta V(\tilde{\varphi}_h(x,u)) \right\}$$

Idee: Fixpunktiteration mit dem Bellman-Operator

$$T_h(v)(x) := \sup_{u \in U} \left\{ \tilde{g}_h(x,u) + \beta v(\tilde{\varphi}_h(x,u)) \right\}$$

d.h. konstruiere Folge (v_k) (von Funktionen!) mit $v_{k+1} = T_h(v_k)$. („Werteiteration")