Vorlesung Variationsrechnung und optimale Steuerung

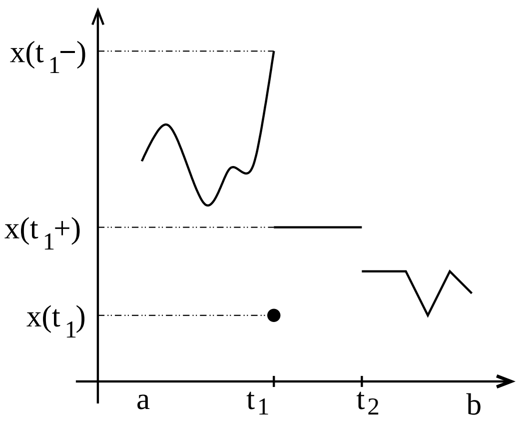
1. Foliensatz, Version vom 1.7.2014

Michael Karow

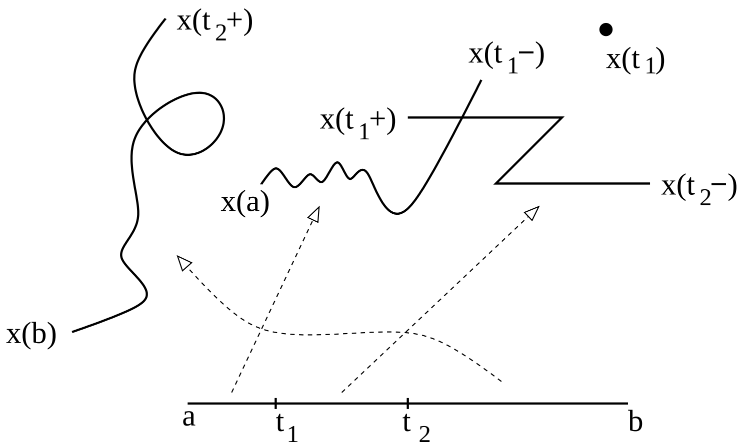
Stückweise stetige Funktionen, $\mathcal{PC}^1([a,b],\mathbb{R}^n)$

$x : [a,b] \to \mathbb{R}^n$ an höchstens endlich vielen Stellen t_j unstetig oder nicht definiert.
An diesen Stellen existieren die einseitigen Grenzwerte $x(t_j+) \in \mathbb{R}^n$ und $x(t_j-) \in \mathbb{R}^n$.

$x : [a,b] \to \mathbb{R}$

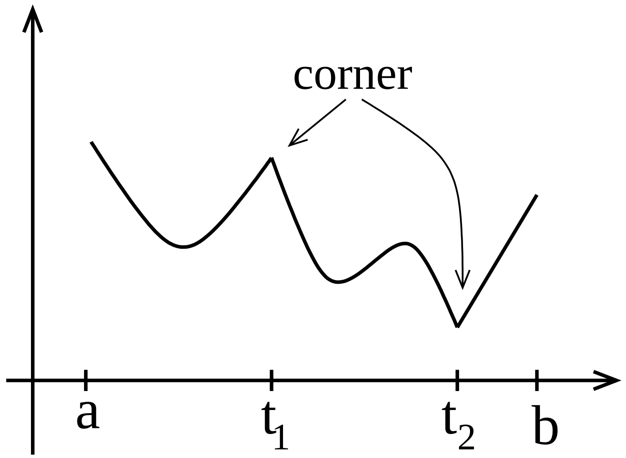


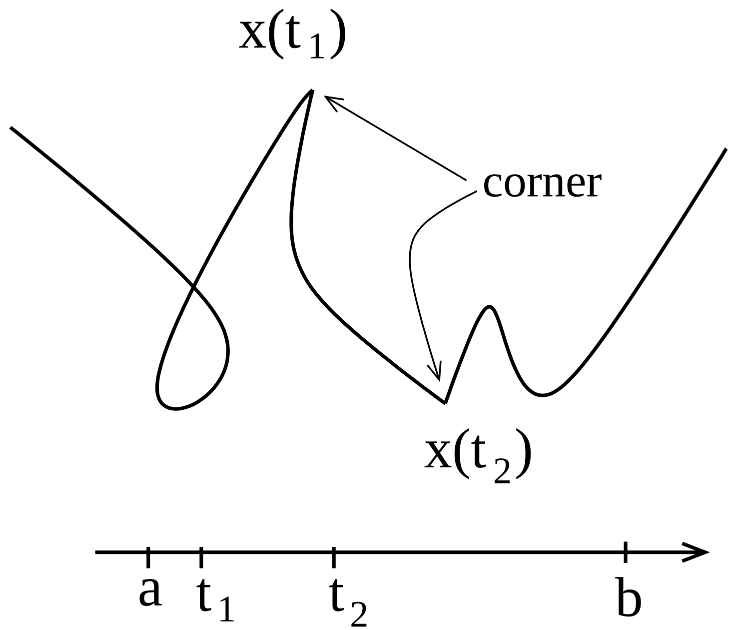
$x : [a,b] \to \mathbb{R}^2$



Stückweise stetig differenzierbare Funktionen, $\mathcal{PC}^1([a,b],\mathbb{R}^n)$

$x : [a,b] \to \mathbb{R}^n$ stetig.
Ableitung $\dot{x} : [a,b] \to \mathbb{R}^n$ an höchstens endlich vielen Stellen t_j unstetig oder nicht definiert.
An diesen Stellen existieren die einseitigen Grenzwerte $\dot{x}(t_j+) \in \mathbb{R}^n$ und $\dot{x}(t_j-) \in \mathbb{R}^n$. Ecken: $(t_j, x(t_j))$





Normen und Abstandsmessung

Euklidische Vektornorm: $\|v\|_2 = \sqrt{\sum_i v_i^2}$

sup-norm auf $\mathcal{PC}([a,b],\mathbb{R}^n)$: $\quad \|x\|_\infty = \sup_{t\in[a,b],t\neq t_j} \|x(t)\|_2$

sup-norm auf $\mathcal{PC}^1([a,b],\mathbb{R}^n)$: $\quad \|x\|_{\infty,1} = \|x\|_\infty + \|\dot{x}\|_\infty$

Abstände:

$$\|x_1 - x_2\|_\infty, \qquad \|x_1 - x_2\|_{\infty,1} = \|x_1 - x_2\|_\infty + \|\dot{x}_1 - \dot{x}_2\|_\infty$$

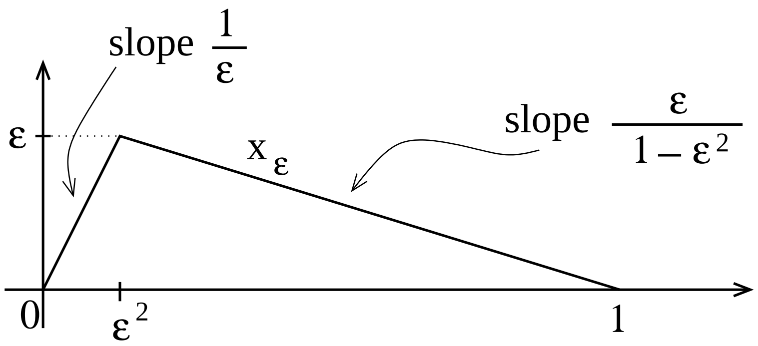
$\|x_1 - x_2\|_\infty$ kann klein sein, während $\|x_1 - x_2\|_{\infty,1}$ groß ist.
Siehe nächstes Beispiel.

Konvergenz von Funktionenfamilien

Familie $x_\epsilon \in \mathcal{PC}^1([a,b],\mathbb{R}^n)$, $\epsilon \in E \subseteq \mathbb{R}$, konvergiert gegen x bzgl. $\|\cdot\|$ wenn

$$\lim_{\epsilon\to 0} \|x_\epsilon - x\| = 0.$$

Konvergenz hängt von der Norm ab.
Beispiel: Sei $x_\epsilon \in \mathcal{PC}^1([0,1],\mathbb{R})$ wie im Bild definiert.



Dann gilt

$$\|x_\epsilon - 0\|_\infty \to 0, \qquad \|x_\epsilon - 0\|_{\infty,1} = \|x_\epsilon - 0\|_\infty + \|\dot{x}_\epsilon - \dot{0}\|_\infty \to \infty$$

für $\epsilon \to 0$.

Grundaufgabe der Variationsrechnung:

Minimiere das Integral

$$J(x) = \int_a^b f(t, x(t), \dot{x}(t))\, dt, \qquad x \in \mathcal{Z} \subseteq \mathcal{PC}^1([a,b], \mathbb{R}^n).$$

Dabei ist $\mathcal{Z}$ die Menge der zulässigen Funktionen.
Einfachstes Beispiel:

$$\mathcal{Z} = \{x \in \mathcal{PC}^1([a,b], \mathbb{R}^n) \mid \ x(a) = A, x(b) = B \ \}.$$

globale, schwache und starke lokale Minimierer

Problem: Minimiere $J : \mathcal{Z} \to \mathbb{R}$

$x_0 \in \mathcal{Z}$ heißt

- **globaler** Minimierer, wenn

$$J(x_0) \leq J(x) \text{ für alle } x \in \mathcal{Z}.$$

- **starker** lokaler Minimierer, wenn

$$J(x_0) \leq J(x) \text{ für alle } x \in \mathcal{Z} \text{ mit } \|x - x_0\|_\infty < \epsilon_0.$$

- **schwacher** lokaler Minimierer, wenn

$$J(x_0) \leq J(x) \text{ für alle } x \in \mathcal{Z} \text{ mit } \|x - x_0\|_\infty + \|\dot{x} - \dot{x}_0\|_\infty < \epsilon_0.$$

Es gilt:

x_0 globaler Minimierer $\Rightarrow$ starker Minimierer $\Rightarrow$ schwacher Minimierer.

Die Umkehrung gilt im allgemeinen nicht. (nächstes Beispiel)

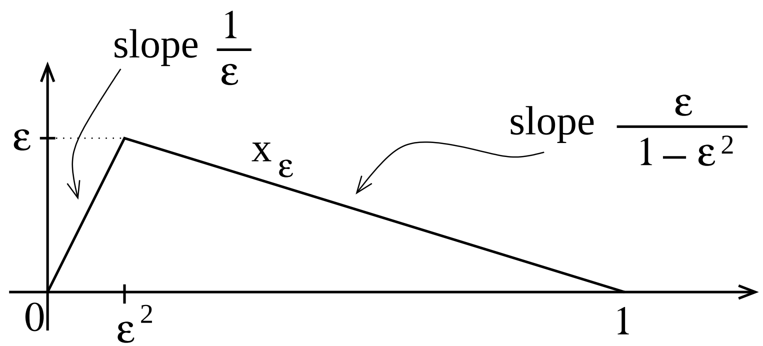
1. Beispiel zu Minimierern

Betrachte das Funktional

$$J : \mathcal{PC}^1([0,1],\mathbb{R}) \to \mathbb{R}, \qquad J(x) = \int_0^1 \dot{x}(t)^2 - \dot{x}(t)^3\,dt$$

Die Nullfunktion $x_0 \equiv 0$ ist ein schwacher Minimierer, denn $\|x\|_\infty + \|\dot{x}\|_\infty < 1$ impliziert $|\dot{x}(t)| < 1$, also $J(x) \geq 0 = J(x_0)$.

x_0 ist aber kein starker Minimierer, denn die Familie



konvergiert gegen x_0 bzgl. $\|\cdot\|_\infty$, und für $\epsilon \to 0$,

$$J(x_\epsilon) = \epsilon^2\left(\frac{1}{\epsilon^2} - \frac{1}{\epsilon^3}\right) + (1-\epsilon^2)\left((\frac{\epsilon}{1-\epsilon^2})^2 - (\frac{\epsilon}{1-\epsilon^2})^3\right) \quad \to \quad -\infty$$

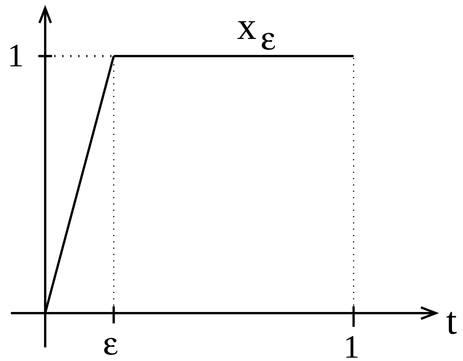
Einen globalen Minimierer gibt es also auch nicht.

2. Beispiel zu Minimierern

Betrachte das Funktional $J : \mathcal{Z} \to \mathbb{R}$ mit

$$J(x) = \int_0^1 t^2 \dot{x}(t)^2 \, dt, \qquad \mathcal{Z} = \{ x \in \mathcal{PC}^1([0,1],\mathbb{R}) \mid x(0) = 0, x(1) = 1 \}.$$

Offensichtlich ist $J(x) \geq 0$ für alle x. Es gibt eine Funktionenfamilie für die $J(x_\epsilon)$ dem Wert 0 beliebig nahe kommt, nämlich diese:



$$J(x_\epsilon) = \int_0^\epsilon t^2 \left(\frac{1}{\epsilon}\right)^2 \, dt = \frac{\epsilon}{3}.$$

Angenommen, es gibt einen globalen Minimierer x_*. Dann gilt

$$0 \leq J(x_*) \leq J(x_\epsilon) \qquad \text{für alle } \epsilon > 0.$$

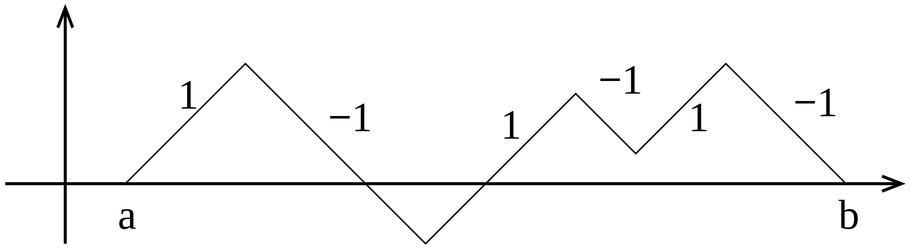
Also $J(x_*) = 0$. Der Wert $J(x_*) = 0$ impliziert aber, dass $\dot{x}_*(t) = 0$, also dass x_* konstant ist, also $x_* \notin \mathcal{Z}$.
Somit gibt es keinen globalen Minimierer.

3. Beispiel zu Minimierern

Für das Funktional $J : \mathcal{Z} \to \mathbb{R}$ mit

$$J(x) = \int_0^1 (\dot{x}(t)^2 - 1)^2 \, dt, \qquad \mathcal{Z} = \{x \in \mathcal{PC}^1([0,1], \mathbb{R}) \mid x(0) = x(1) = 0\}$$

gibt es unendlich viele globale Minimierer, nämlich diese:



Grund: Es ist $J(x) \geq 0$ für jedes x. $J(x) = 0$ gilt genau dann, wenn $\dot{x}(t)^2 = 1$ für alle t, für die $\dot{x}(t)$ stetig ist, also $\dot{x}(t) = \pm 1$.

Notwendige Bedingungen für einen Minimierer 1

Grundidee: Sei $x_0 \in \mathcal{Z}$ ein (starker oder schwacher) lokaler Minimierer für $J : \mathcal{Z} \to \mathbb{R}$, und sei x_ϵ eine Funktionenfamilie die (bzgl. $\|\cdot\|_{\infty,1}$ oder $\|\cdot\|_\infty$) gegen x_0 konvergiert. Dann gilt für hinreichend kleine ϵ,

$$J(x_0) \leq J(x_\epsilon).$$

Die Funktion $\epsilon \mapsto J(x_\epsilon)$ nimmt an der Stelle 0 ein Minimum an.
Bei Differenzierbarkeit dieser Funktion folgt

- falls x_ϵ für positive und negative ϵ definiert ist:

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} J(x_\epsilon) = 0, \qquad \frac{d^2}{d\epsilon^2}\bigg|_{\epsilon=0} J(x_\epsilon) \geq 0$$

- falls x_ϵ nur für nichtnegative ϵ definiert ist:

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} J(x_\epsilon) \geq 0.$$

Notwendige Bedingungen für einen Minimierer 2

Die 3 klassischen notwendigen Bedingungen 1. Ordnung für einen Minimierer von $\int_a^b f(t, x(t), \dot{x}(t))\, dt$ wenn

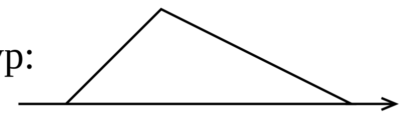
$$\mathcal{Z} \supseteq \{x \in \mathcal{PC}^1([a,b], \mathbb{R}^n) | x(a) = A, x(b) = B\}$$

sind folgende.

1. Jeder schwache lokale Minimierer erfüllt die **Euler-Lagrange-Gleichung**.
   Beweis mit vertikalen Variationen vom 'Dachtyp' (Fundamentallemma).
2. Jeder starke lokale Minimierer erfüllt die **'zweite' Euler-Lagrange-Gleichung**.
   Beweis mit horizontalen Variationen vom 'Dachtyp'(Fundamentallemma).
3. Für jeden starken lokalen Minimierer nimmt die **Weierstraß-Exzess-Funktion** nur nichtnegative Werte an.
   Beweis mit vertikalen Variationen vom 'Hakentyp'.

Dachtyp:

Hakentyp:

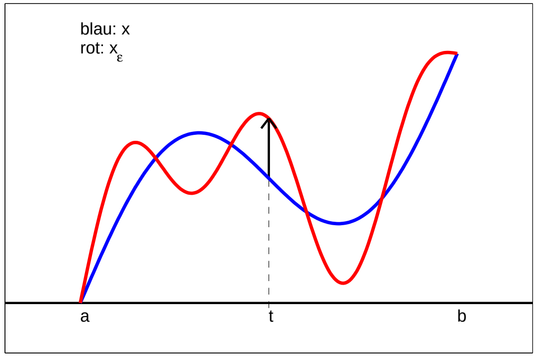


Die erste Variation (Richtungsableitung)

Vertikale Variationen:

$$x_\epsilon(t) = x(t) + \epsilon\, h(t), \qquad \epsilon \in \mathbb{R},\ h \in \mathcal{PC}^1([a,b],\mathbb{R}^n) \text{ so dass } x_\epsilon \in \mathcal{Z}$$

Wenn $\mathcal{Z} = \{x \in \mathcal{PC}^1([a,b],\mathbb{R}^n) \mid x(a) = A, x(b) = B\}$ dann $h(a) = h(b) = 0$.



Erste Variation (Richtungsableitung) von J:

$$\delta J(x,h) := \frac{d}{d\epsilon}\Big|_{\epsilon=0} J(x_\epsilon) = \frac{d}{d\epsilon}\Big|_{\epsilon=0} J(x + \epsilon\, h).$$

Häufige Notation in der Literatur: $h = \delta x$ (virtuelle Verschiebung)

Berechnung der 1. Variation:

$$\delta J(x,h) = \frac{d}{d\epsilon}\Big|_{\epsilon=0} \int_a^b f(\,t,\ x(t)+\epsilon\, h(t)\ ,\ \dot{x}(t)+\epsilon\, \dot{h}(t)\ )\,dt$$

$$= \int_a^b \frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0} f(\,t,\ x(t)+\epsilon\, h(t)\ ,\ \dot{x}(t)+\epsilon\, \dot{h}(t)\ )\,dt$$

$$= \int_a^b f_x(t,x(t),\dot{x}(t))\ h(t)\ +\ f_{\dot{x}}(t,x(t),\dot{x}(t))\ \dot{h}(t)\,dt$$

$$\underset{(*)}{=} \int_a^b \left[f_x(t,x(t),\dot{x}(t)) - \frac{d}{dt} f_{\dot{x}}(t,x(t),\dot{x}(t)) \right] h(t)\,dt \qquad (*)$$

$$+\ \underbrace{f_{\dot{x}}(t,x(t),\dot{x}(t))\ h(t)\ \Big|_{t=a}^{t=b}}_{=0 \text{ falls } h(a)=h(b)=0},$$

mit den Zeilenvektoren

$$f_x = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right), \qquad f_{\dot{x}} = \left(\frac{\partial f}{\partial \dot{x}_1}, \ldots, \frac{\partial f}{\partial \dot{x}_n} \right).$$

$(*)$ Partielle Integration. Geht nur, falls $f_{\dot{x}}(t,x(t),\dot{x}(t)) \in \mathcal{PC}^1$. Offensichtlich ist $\delta J(x,h) = 0$, wenn der Klammerausdruck 0 ist, x also die Euler-Lagrange-Gleichung löst. (→ nächste Seite).

Die Euler-Lagrange-Gleichung

Satz. Sei x_0 ein schwacher lokaler Minimierer für das Funktional

$$J : \mathcal{Z} \to \mathbb{R}, \qquad J(x) = \int_a^b f(t, x(t), \dot{x}(t))\, dt.$$

Angenommen für jedes $h \in \mathcal{PC}_0^1([a,b], \mathbb{R}^n)$ (d.h. $h(a) = h(b) = 0$) enthält $\mathcal{Z}$ die Funktionen

$$x_\epsilon := x_0 + \epsilon\, h, \qquad \epsilon \in [-\epsilon_0, \epsilon_0] \text{ für ein } \epsilon_0 > 0.$$

Dann erfüllt $x := x_0$ an allen Stetigkeitspunkten von $\dot{x}$ die **Euler-Lagrange-Gleichung**

in Integralform:

$$f_{\dot{x}}(t, x(t), \dot{x}(t)) = \int_a^t f_x(s, x(s), \dot{x}(s))\, ds + c. \qquad (*)$$

in differentieller Form:

$$\frac{d}{dt} f_{\dot{x}}(t, x(t), \dot{x}(t)) = f_x(t, x(t), \dot{x}(t)).$$

Bei 2-maliger stetiger Differenzierbarkeit von f und x an der Stelle t:

$$f_{\dot{x}\dot{x}}(t, x(t), \dot{x}(t))\, \ddot{x}(t) + f_{\dot{x}x}(t, x(t), \dot{x}(t))\, \dot{x}(t) + f_{\dot{x}t}(t, x(t), \dot{x}(t)) = f_x(t, x(t), \dot{x}(t))$$

Die Euler-Lagrange-Gleichung liefert also ein **notwendiges** Kriterium für einen schwachen (und damit auch einen starken oder globalen) Minimierer.
$(*)$ Die untere Integrationsgrenze kann auch zwischen a und t liegen.

1. Erläuterung zur Euler-Lagrange-Gleichung

Bei 2-maliger stetiger Differenzierbarkeit von f und x an der Stelle t
lautet die Euler-Lagrange-Gleichung:

$$f_{\dot{x}\dot{x}}(t,x(t),\dot{x}(t))\,\ddot{x}(t)+f_{\dot{x}x}(t,x(t),\dot{x}(t))\,\dot{x}(t)+f_{\dot{x}t}(t,x(t),\dot{x}(t))=f_x(t,x(t),\dot{x}(t))$$

In Matrix-Vektor-Notation (mit wegelassenen Argumenten):

$$\begin{bmatrix} f_{\dot{x}_1\dot{x}_1} & \cdots & f_{\dot{x}_1\dot{x}_n} \\ \vdots & & \vdots \\ f_{\dot{x}_n\dot{x}_1} & \cdots & f_{\dot{x}_n\dot{x}_n} \end{bmatrix} \begin{bmatrix} \ddot{x}_1 \\ \vdots \\ \ddot{x}_n \end{bmatrix} + \begin{bmatrix} f_{\dot{x}_1 x_1} & \cdots & f_{\dot{x}_1 x_n} \\ \vdots & & \vdots \\ f_{\dot{x}_n x_1} & \cdots & f_{\dot{x}_n x_n} \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{bmatrix} + \begin{bmatrix} f_{\dot{x}_1 t} \\ \vdots \\ f_{\dot{x}_n t} \end{bmatrix} = \begin{bmatrix} f_{x_1} \\ \vdots \\ f_{x_n} \end{bmatrix}$$

mit

$$f_{\dot{x}_i\dot{x}_k}=\frac{\partial^2 f}{\partial\dot{x}_i\,\partial\dot{x}_k},\qquad f_{\dot{x}_i x_k}=\frac{\partial^2 f}{\partial x_i\,\partial\dot{x}_k},\qquad f_{\dot{x}_i t}=\frac{\partial^2 f}{\partial t\,\partial\dot{x}_i},\qquad f_{x_i}=\frac{\partial f}{\partial x_i}.$$

Wenn die symmetrische Matrix $f_{\dot{x}\dot{x}}$ invertierbar ist, handelt es sich um eine Differentialgleichung 2. Ordnung, anderenfalls möglicherweise um eine differential-algebraische Gleichung. Für eine eindeutige Lösung braucht man 2 vektorwertige Randbedingungen, die im Fall

$$\mathcal{Z}=\{x\in\mathcal{PC}^1([a,b],\mathbb{R}^n)\mid x(a)=A,x(b)=B\}$$

unmittelbar gegeben sind.

Terminologie: Die Lösungen der Euler-Lagrange-Gleichung heißen **Extremalen** (bei Ecken: gebrochene Extemalen) unabhängig davon, ob sie Minimierer sind oder nicht.

2. Erläuterung zur Euler-Lagrange-Gleichung

Von der Integralform

$$f_{\dot{x}}(t, x(t), \dot{x}(t)) = \int_a^t f_x(s, x(s), \dot{x}(s))\, ds + c. \qquad (*)$$

zur differentiellen Form

$$\frac{d}{dt} f_{\dot{x}}(t, x(t), \dot{x}(t)) = f_x(t, x(t), \dot{x}(t))$$

gelangt man über den **Hauptsatz der Differential- und Integralrechnung.**

Für $\varphi \in \mathcal{PC}^1$ lautet er: Die Funktion

$$\Phi(t) = \int_a^t \varphi(s)\, ds$$

ist stetig und an allen Stetigkeitsstellen t von φ differenzierbar mit

$$\dot{\Phi}(t) = \varphi(t).$$

Φ ist also eine Stammfunktion von φ. Alle anderen Stammfunktionen sind $\Phi(t) + c$.

Der Hauptsatz wiederum ist eine einfache Konsequenz aus dem Schrankensatz und dem Mittelwertsatz der Integralrechnung, siehe nächste Seite.

Erinnerung: Schrankensatz und Mittelwertsatz der Integralrechnung

Schrankensatz:

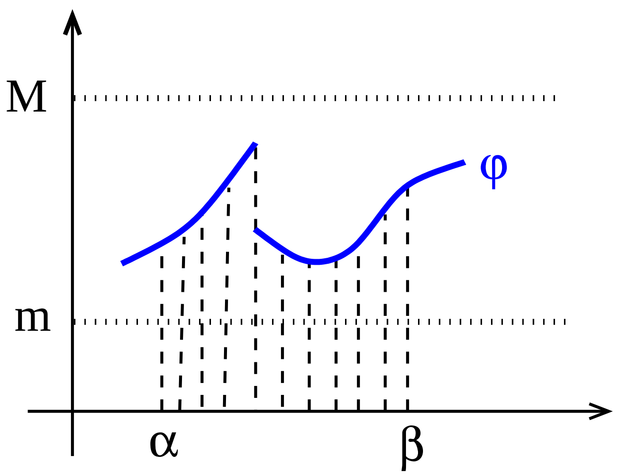
Wenn $m \leq \varphi(t) \leq M$ für $t \in [\alpha, \beta]$, dann $m \leq \dfrac{1}{\beta - \alpha} \displaystyle\int_\alpha^\beta \varphi(s)\, ds \leq M.$

Mittelwertsatz:

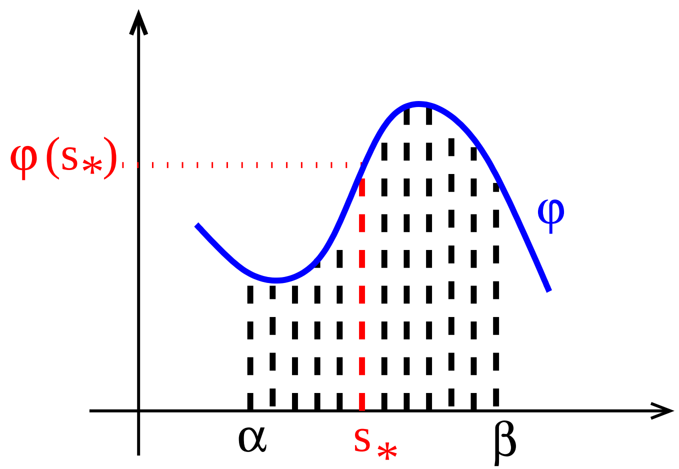
Wenn φ im Intervall $[\alpha, \beta]$ stetig ist, dann $\dfrac{1}{\beta - \alpha} \displaystyle\int_\alpha^\beta \varphi(s)\, ds = \varphi(s_*)$ für ein $s_* \in [\alpha, \beta]$.

Der Hauptsatz der Differential- und Integralrechnung folgt diesen Sätzen, indem man $(\alpha, \beta) = (t, t + \epsilon)$ und $(\alpha, \beta) = (t - \epsilon, t)$ setzt und einen Grenzübergang $\epsilon \downarrow 0$ macht.

Zum Schrankensatz:



Zum Mittelwertsatz:



Herleitung der Euler-Lagrange-Gleichung mit dem Fundamentallemma der Variationsrechnung

Bisher wissen wir: Wenn x_0 schwacher lokaler Minimierer ist,
dann ist für jedes $h \in \mathcal{PC}^1([a,b],\mathbb{R}^n)$ mit $h(a) = h(b) = 0$,

$$0 = \delta J(x_0, h) = \int_a^b f_x(t, x(t), \dot{x}(t))\, h(t) + f_{\dot{x}}(t, x(t), \dot{x}(t))\, \dot{h}(t)\, dt.$$

Hieraus gelangen wir sofort zu Euler-Lagrange-Gleichung durch das

Fundamentallemma der Variationsrechnung (Lemma on Du Bois-Reymond):

Seien $u, v \in \mathcal{PC}([a,b],\mathbb{R}^{n\times 1})$ (Zeilenvektoren). Folgende Aussagen sind äquivalent.

(1) Für alle $h \in \mathcal{PC}^1([a,b],\mathbb{R}^n)$ mit $h(a) = h(b) = 0$,

$$0 = \int_a^b u(t)\, h(t) + v(t)\, \dot{h}(t)\, dt.$$

(2) Es gibt eine Konstante $c \in \mathbb{R}$, so dass an jeder Stetigkeitsstelle t von u und v,

$$v(t) = \int_a^t u(s)\, ds + c.$$

Zum Beweis von (1) $\Rightarrow$ (2) (nächste Seite) reichen Funktionen h vom Dachtyp aus. Man könnte die Voraussetzung des Lemmas also auf solche h einschränken. Beachte auch, dass keine Differenzierbarkeitsvoraussetzungen an u und v gestellt werden. Die stückweise Differenzierbarkeit von v ist eine Konsequenz des Lemmas.

Beweis des Fundamentallemmas

(2) $\Rightarrow$ (1) folgt so: $\quad v(t) = \int_a^t u(s)\,ds + c \;\Rightarrow\; \dot v(t) = u(t) \;\Rightarrow$

$$\int_a^b u(t)\,h(t) + v(t)\,\dot h(t)\,dt = \int_a^b \dot v(t)\,h(t) + v(t)\,\dot h(t)\,dt = \int_a^b \frac{d}{dt}(v(t)h(t))\,dt = v(t)h(t)|_{t=a}^{t=b} = 0$$

für alle $h \in \mathcal{PC}^1$ mit $h(a) = h(b) = 0$.

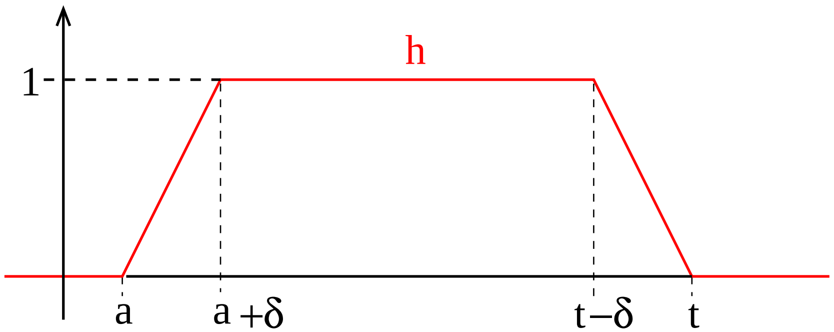
Beweis von (1) $\Rightarrow$ (2):

Wir betrachten nur skalaren Fall $u, v, h \in \mathbb{R}$.

Vektorieller Fall folgt dann leicht (wie?).

Fixiere ein $\delta > 0$,

und wähle h wie im Bild.



(1) $\Rightarrow 0 = \int_a^b u(t)\,h(t) + v(t)\,\dot h(t)\,dt = I_1(\delta) + I_2(\delta)$, wobei

$$I_1(\delta) \;:=\; \int_a^t u(t)\,h(t)\,dt = \int_a^{a+\delta} u(t)\,h(t)\,dt + \int_{a+\delta}^{t-\delta} u(t)\,dt + \int_t^{t-\delta} u(t)\,h(t)\,dt$$

$$I_2(\delta) \;:=\; \int_a^t v(t)\,\dot h(t)\,dt = \frac{1}{\delta}\int_a^{a+\delta} v(t)\,dt + \left(-\frac{1}{\delta}\right)\int_{t-\delta}^t v(t)\,dt = v(t_*) - v(t_{**})$$

mit $t_* \in [a, a+\delta]$, $t_{**} \in [t-\delta, t]$ wegen Mittelwertsatz.

Grenzübergang: $0 = \lim_{\delta\downarrow 0}(I_1(\delta) + I_2(\delta)) = \int_a^t u(t)\,dt + v(a) - v(t)$ $\square$

Folgerungen aus dem Fundamentallemma

Notation: $\mathcal{PC}_0^1 := \{h \in \mathcal{PC}^1([a,b],\mathbb{R}^n) \mid h(a) = h(b) = 0\}$

(a) Wenn

$$\int_a^b u(t)\, h(t)\, dt = 0 \qquad \text{für alle} \quad h \in \mathcal{PC}_0^1,$$

dann ist $u(t) = 0$ an jeder Stetigkeitsstelle von u.

(b) Wenn

$$\int_a^b v(t)\, \dot{h}(t)\, dt = 0 \qquad \text{für alle} \quad h \in \mathcal{PC}_0^1,$$

dann gibt es eine Konstante c, so dass $v(t) = c$ an jeder Stetigkeitsstelle t von v.

Beweis: Setze im Fundamentallemma $v = 0$ oder $u = 0$.

Die Aussagen (a) und (b) werden in der Literatur oft auch als Fundamentallemma bezeichnet.

Übung: ein erweitertes Funktional

Wir leiten notwendige Bedingungen für einen schwachen lokalen Minimierer des folgenden Funktionals her:

$$J : \mathcal{PC}^1([a,b],\mathbb{R}) \to \mathbb{R}, \qquad J(x) = \int_a^b f(t,x(t),\dot{x}(t))\,dt \, + g(x(b)).$$

An x werden keine Randbedingungen gestellt. Wir werden aber gleich sehen, dass ein Minimierer bestimmte Randbedingungen erfüllen muss.

Die 1. Variation muss für **alle** $h \in \mathcal{PC}^1([a,b],\mathbb{R}^n)$ verschwinden:

$$\begin{aligned} 0 = \delta J(x,h) \quad &= \int_a^b \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} f(\,t, x(t) + \epsilon\, h(t), \dot{x}(t) + \epsilon \dot{h}(t))\,dt \; + \; \frac{d}{d\epsilon}\bigg|_{\epsilon=0} g(x(b) + \epsilon\, h(b)) \\ &\underset{(*)}{=} \int_a^b f_x(t,x(t),\dot{x}(t))\, h(t) \, + \, f_{\dot{x}}(t,x(t),\dot{x}(t))\, \dot{h}(t)\,dt \; + \; g_x(x(b))\, h(b) \\ &\underset{(**)}{=} \int_a^b \Big[f_x(t,x(t),\dot{x}(t)) - \frac{d}{dt} f_{\dot{x}}(t,x(t),\dot{x}(t)) \Big]\, h(t)\,dt \\ &\qquad + \big(\, f_{\dot{x}}(b,x(b),\dot{x}(b)) + g_x(x(b)) \,\big)\, h(b) - f_{\dot{x}}(a,x(a),\dot{x}(a))\, h(a). \end{aligned}$$

Die partielle Integration $(**)$ geht nur, wenn wir wissen, dass $\frac{d}{dt} f_{\dot{x}}$ existiert. $(*)$ gilt insbesondere für alle h mit $h(a) = h(b) = 0$. Für diese h entfällt der Randterm. Aus dem Fundamentallemma folgt, dass f die Euler-Lagrange-Gleichung löst. Damit ist der Klammerausdruck in $(**)$ Null und das Integral verschwindet für **alle** h. Übrig bleibt:

$$0 = \big(\, f_{\dot{x}}(b,x(b),\dot{x}(b)) + g_x(x(b)) \,\big)\, h(b) - f_{\dot{x}}(a,x(a),\dot{x}(a))\, h(a).$$

Durch geeignete Wahl verschiedener hs (Hausaufgabe) folgen die Randbedingungen:

$$f_{\dot{x}}(b,x(b),\dot{x}(b)) + g_x(x(b)) = 0, \qquad f_{\dot{x}}(a,x(a),\dot{x}(a)) = 0.$$

Fazit: wenn keine Randbedingungen vorgegeben sind, ergeben sie sich 'natürlich'.

Die 'zweite' Euler-Lagrange-Gleichung

Satz. Sei x_0 ein **starker** lokaler Minimierer für das Funktional

$$J : \mathcal{Z} \to \mathbb{R}, \qquad J(x) = \int_a^b f(t, x(t), \dot{x}(t))\, dt,$$

wobei

$$\mathcal{Z} \supseteq \{x \in \mathcal{PC}^1([a,b], \mathbb{R}^n) \mid x(a) = A, x(b) = B\}.$$

Sei f nach allen Variablen stetig differenzierbar. Dann erfüllt $x := x_0$ an allen Stetigkeitspunkten von $\dot{x}$ die **'zweite' Euler-Lagrange-Gleichung**

in Integralform:

$$f(t, x(t), \dot{x}(t)) - f_{\dot{x}}(t, x(t), \dot{x}(t))\, \dot{x}(t) = \int_a^t f_t(s, x(s), \dot{x}(s))\, ds + c. \qquad (*)$$

in differentieller Form:

$$\frac{d}{dt}\Big[f(t, x(t), \dot{x}(t)) - f_{\dot{x}}(t, x(t), \dot{x}(t))\, \dot{x}(t)\Big] = f_t(s, x(s), \dot{x}(s)) \qquad (**)$$

$(**)$ gilt auch, wenn x ein 2mal stetig differenzierbarer schwacher lokaler Minimierer ist.

Der Term auf der linken Seite von $(*)$ und $(**)$ wird übersichtlicher, wenn man die Argumente weglässt. Dann lautet er $f - f_{\dot{x}}\, \dot{x}$.

Folgerung aus dem Satz:
Wenn f nicht explizit von t abhängt, dann ist $f - f_{\dot{x}}\, \dot{x}$ konstant (nicht von t abhängig).

Herleitung der 2. Euler-Lagrange-Gleichung wenn $\ddot{x}$ existiert

Zur Übersichtlichkeit lassen wir die Argumente weg.

Wir nehmen 2malige stetige Differenzierbarkeit von x an.

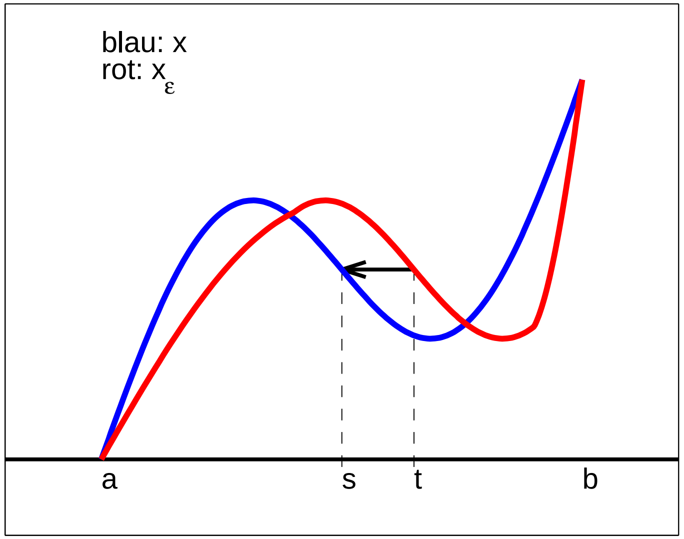
Dann folgt mit Kettenregel, Produktregel und der 1. Euler-Lagrange-Gleichung

$$\begin{aligned}\frac{d}{dt}(f - f_{\dot{x}}\,\dot{x}) &= f_t + f_x\,\dot{x} + f_{\dot{x}}\,\ddot{x} - \left(\frac{d}{dt}f_{\dot{x}}\right)\dot{x} - f_{\dot{x}}\ddot{x} \\ &= f_t + f_x\,\dot{x} + f_{\dot{x}}\,\ddot{x} - f_x\,\dot{x} - f_{\dot{x}}\ddot{x} \\ &= f_t.\end{aligned}$$

Herleitung der 2. Euler-Lagrange-Gleichung in Integralform

Sei $h \in \mathcal{PC}^1([a,b],\mathbb{R})$ mit $h(a) = h(b) = 0$.
Wenn $|\epsilon|$ hinreichend klein, dann ist die Abbildung

$$[a,b] \ni s \longmapsto t = g_\epsilon(s) := s + \epsilon\, h(s) \in [a,b]$$



umkehrbar, denn $|\dot{g}_\epsilon(s)| = |1 + \epsilon\, \dot{h}(s)| \geq 1 - |\epsilon|\, |\dot{h}(s)|$
ist dann positiv, also g_ϵ monoton wachsend.
Die (horizontale) Variation (Argumentverschiebung)

$$x_\epsilon(t) := x(g_\epsilon(t)^{-1}) = x(s)$$

ist daher wohldefiniert. Die Ableitung ist

$$\dot{x}_\epsilon(t) = \dot{x}(g_\epsilon^{-1}(t))\, \frac{d}{dt} g_\epsilon^{-1}(t) = \frac{\dot{x}(s)}{\dot{g}_\epsilon(s)} = \frac{\dot{x}(s)}{1 + \epsilon\, \dot{h}(s)}.$$

Mit Substitutionsregel folgt

$$J(x_\epsilon) \;=\; \int_a^b f(t, x_\epsilon(t), \dot{x}_\epsilon(t))\, dt = \int_a^b f\left(s + \epsilon\, h(s),\, x(s),\, \frac{\dot{x}(s)}{1 + \epsilon\, \dot{h}(s)}\right)(1 + \epsilon\, \dot{h}(s))\, ds$$

Da x Minimierer ist, folgt

$$0 = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} J(x_\epsilon) = \int_a^b f_t(s, x(s), \dot{x}(s))\, h(s) + \left(f(s, x(s), \dot{x}(s)) - f_{\dot{x}}(s, x(s), \dot{x}(s))\, \dot{x}(s)\right) \dot{h}(s)\, ds$$

Anwendung des Fundamentallemmas ergibt nun die Behauptung.

Das Hamiltonprinzip und die Euler-Lagrange-Gleichungen

Gegeben sei ein mechanisches System mit

- Freiheitsgraden (Zustand) $x \in \mathbb{R}^n$,
- kinetischer Energie $\frac{1}{2}\dot{x}^\top M\,\dot{x}$, $M \in \mathbb{R}^{n\times n}$ symmetrisch positiv definit (Massenmatrix),
- Potential $V(x)$.

Hamiltonprinzip:

Die Zeitentwicklung $x(t)$ des Zustands ist eine Extremale des Funktionals

$$J(x) = \int_a^b f(x(t), \dot{x}(t))\,dt, \qquad f(x,\dot{x}) = \frac{1}{2}\dot{x}^\top M\dot{x} - V(x), \qquad \text{(Wirkungsintegral).}$$

Die 'erste' Euler-Lagrange-Gleichung, $\frac{d}{dt}f_{\dot{x}} = f_x$, lautet dann (Spaltenvektoren):

$$M\ddot{x}(t) = -\nabla V(x) \qquad (\nabla = \text{Gradient})$$

Die 'zweite' Euler-Lagrange-Gleichung, $f - f_{\dot{x}}\dot{x} = const$, ergibt

$$-(\underbrace{\frac{1}{2}\dot{x}^\top M\dot{x} + V(x)}_{\text{Gesamtenergie}}) = const.$$

Details: Hausaufgabe

Die Weierstraßsche Exzess-Funktion (kurz: E-Funktion)

Die Exzess-Funktion ist folgendermaßen definiert:

$$E_f(t,x,u_0,u) := f(t,x,u) - f(t,x,u_0) - f_{\dot{x}}(t,x,u_0)(u-u_0).$$

Bedeutung: Wenn man die Gleichung umstellt, bekommt man

$$f(t,x,u) = f(t,x,u_0) + f_{\dot{x}}(t,x,u_0)(u-u_0) + E_f(t,x,u_0,u)$$

E_f ist also das Restglied in der Taylor-Entwicklung 1. Ordnung der Funktion

$$u \longmapsto f(t,x,u)$$

mit Entwicklungspunkt u_0.

Satz (Weierstraß-Kriterium):

Sei x_0 ein starker lokale Minimierer des Funktionals $J(x) = \int_a^b f(t,x(t),\dot{x}(t))\,dt$. Dann ist

$$E_f(t,x_0(t),\dot{x}_0(t-),u) \geq 0, \qquad E_f(t,x_0(t),\dot{x}_0(t+),u) \geq 0$$

für all $t \in [a,b]$ und alle u, für die $f(t,x_0(t),u)$ definiert ist. Wenn x_0 nur ein schwacher lokaler Minimierer ist, dann gelten diese Ungleichungen nur, wenn $\|u - \dot{x}_0(t\pm)\|$ hinreichend klein ist.

Beweis des Weierstraß-Kriteriums 1

Der Beweis wird für $t+$ geführt.

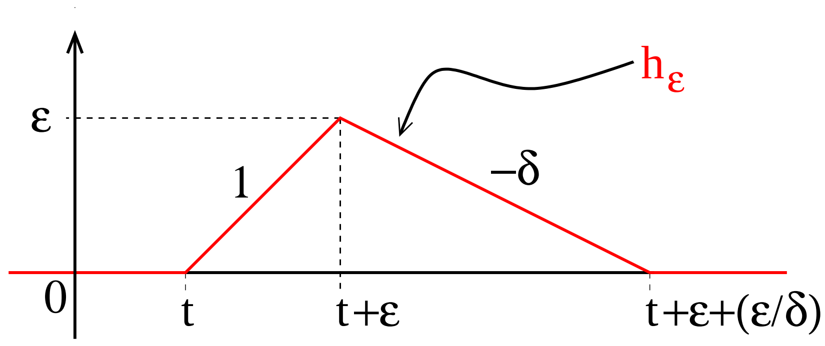
Für $t-$ geht alles analog.

Notationsvereinfachung: $x = x_0$.

Wir fixieren zunächst ein $\delta > 0$.

Für ein variables $\epsilon > 0$ sei

$h_\epsilon : [a, b] \to \mathbb{R}$ wie im Bild definiert.



Sei nun $x_\epsilon = x + h_\epsilon v$ mit $v = u - \dot{x}(t+)$. Da x Minimierer ist, hat man

$$0 \;\leq\; \frac{d}{d\epsilon}\bigg|_{\epsilon=0} J(x_\epsilon) = \lim_{\epsilon \downarrow 0} \frac{J(x_\epsilon) - J(x)}{\epsilon} \;=\; \lim_{\epsilon \downarrow 0} I_1(\epsilon) + \lim_{\epsilon \downarrow 0} I_2(\epsilon),$$

wobei

$$I_1(\epsilon) \;=\; \frac{1}{\epsilon} \int_t^{t+\epsilon} f(s, x_\epsilon(s), \dot{x}_\epsilon(s)) - f(s, x(s), \dot{x}(s))\, ds,$$

$$I_2(\epsilon) \;=\; \frac{1}{\epsilon} \int_{t+\epsilon}^{t+\epsilon+(\epsilon/\delta)} f(s, x_\epsilon(s), \dot{x}_\epsilon(s)) - f(s, x(s), \dot{x}(s))\, ds.$$

Die Grenzwerte von I_1 und I_2 werden im folgenden ausgerechnet (nächste Seite). Dabei wird ausgenutzt, dass

$$\dot{x}_\epsilon(s) = \dot{x}(s) + \dot{h}_\epsilon(s)\, v = \dot{x}(s) + \begin{cases} v & \text{wenn } s \in (t, t+\epsilon), \\ -\delta\, v & \text{wenn } s \in (t+\epsilon, t+\epsilon+(\epsilon/\delta)). \end{cases}$$

Beweis des Weierstraß-Kriteriums 2

Wir haben

$$\begin{aligned}
I_1(\epsilon) &= \frac{1}{\epsilon}\int_t^{t+\epsilon} f(s, x_\epsilon(s), \dot{x}_\epsilon(s)) - f(s, x(s), \dot{x}(s))\, ds \\
&= f(s_*, x_\epsilon(s_*), \dot{x}_\epsilon(s_*)) - f(s_*, x(s_*), \dot{x}(s_*)) \qquad s_* \in [t, t+\epsilon] \quad \text{(Mittelwertsatz)} \\
&= f(s_*, x_\epsilon(s_*), \dot{x}(s_*) + v) - f(s_*, x(s_*), \dot{x}(s_*)) \\
&\to f(t, x(t), \dot{x}(t+) + v) - f(t, x(t), \dot{x}(t+)) \qquad \text{für } \epsilon \downarrow 0.
\end{aligned}$$

$$\begin{aligned}
I_2(\epsilon) &= \frac{1}{\epsilon}\int_{t+\epsilon}^{t+\epsilon+(\epsilon/\delta)} f(s, x_\epsilon(s), \dot{x}_\epsilon(s)) - f(s, x(s), \dot{x}(s))\, ds \\
&= \frac{1}{(\epsilon/\delta)}\int_{t+\epsilon}^{t+\epsilon+(\epsilon/\delta)} \frac{f(s, x_\epsilon(s), \dot{x}_\epsilon(s)) - f(s, x(s), \dot{x}(s))}{\delta}\, ds \\
&= \frac{f(s_*, x_\epsilon(s_*), \dot{x}_\epsilon(s_*)) - f(s_*, x(s_*), \dot{x}(s_*))}{\delta} \qquad s_* \in [t+\epsilon, t+\epsilon+(\epsilon/\delta)] \\
&= \frac{f(s_*, x_\epsilon(s_*), \dot{x}(s_*) - \delta\, v) - f(s_*, x(s_*), \dot{x}(s_*))}{\delta} \\
&\to \frac{f(t, x(t), \dot{x}(t+) - \delta\, v) - f(t, x(t), \dot{x}(t+))}{\delta} \qquad \text{für } \epsilon \downarrow 0.
\end{aligned}$$

Beweis des Weierstraß-Kriteriums 3

Bisher gezeigt:

$$
\begin{aligned}
0 \;\leq\; \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} J(x_\epsilon) &= \lim_{\epsilon\downarrow 0} I_1(\epsilon) + \lim_{\epsilon\downarrow 0} I_2(\epsilon) \\
&= f(t,x(t),\dot{x}(t+)+v) - f(t,x(t),\dot{x}(t+)) \\
&\quad + \\
&\quad \frac{f(t,x(t),\dot{x}(t+)-\delta\, v) - f(t,x(t),\dot{x}(t+))}{\delta} \\
&= f(t,x(t),\dot{x}(t+)+v) - f(t,x(t),\dot{x}(t+)) \\
&\quad - \\
&\quad \frac{f(t,x(t),\dot{x}(t+)-\delta\, v) - f(t,x(t),\dot{x}(t+))}{-\delta}
\end{aligned}
$$

Wenn $\delta \to 0$, dann konvergiert der Quotient (setze $\tau = -\delta$) gegen

$$
\left.\frac{d}{d\tau}\right|_{\tau=0} f(t,x(t),\dot{x}(t+)+\tau\, v) = f_{\dot{x}}(t,x(t),\dot{x}(t+))\, v = f_{\dot{x}}(t,x(t),\dot{x}(t+))\,(u-\dot{x}(t+)).
$$

□