Extremale des Lagrange-Funktionals

Def. Ein Element $x \in \mathcal{K}$ heißt Extremalkurve oder Extremal eines differenzierbaren Funktionals F, falls in diesem Punkt die Ableitung D auf eigentlichen Variationen verschwindet (d.h. gleich 0 ist) auf h, die $h(t_0) = h(t_1) = \vec{0}$ erfüllen.
Satz 2. $\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = \vec{0}$ g.d., w $x(t) \in \mathcal{K}$ ein Extremal des Langange-Funktionals mit Lagrange-Funktion L ist.

Folgerung. Extremale des Lagrange-Funktionals sind die Lösungen der gewöhnlichen Differentialgleichung (GDG)

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = \vec{0}$$

Bemerkung. Die Gleichung heißt Euler-Lagrange-Gleichung und ist in der Physik von größter Bedeutung.

Die Euler-Lagrange-Gleichung ist ein System von n gewöhnlichen Differentialgleichungen auf n unbekannten Funktionen $(x_1(t), \dots, x_n(t))$, wie wir in der Erklärung zu den Sätzen 1 und 2 gesehen haben und wie wir auch in weiteren Beispielen sehen werden.

Extremale des Längenfunktionals

Wir betrachten das Längenfunktional

$$\mathcal{L}(x) = \int_{t_0}^{t_1} \sqrt{\dot{x}_1(t)^2 + \ldots + \dot{x}_n(t)^2}\,dt.$$

Das ist ein Lagrange-Funktional mit
$L(x, \dot{x}, t) = |\dot{x}| = \sqrt{\dot{x}_1(t)^2 + \ldots + \dot{x}_n(t)^2}$.
Aus geometrischen Gründen erwarten wir, dass Geradenstücke die Extremale dieses Funktionals sind, weil die Geraden kürzeste Verbindungskurven sind. Die E-L-Gleichungen bestätigen das: Wir haben

$$\frac{\partial L}{\partial x} \equiv \vec{0}, \quad \frac{\partial L}{\partial \dot{x}} = \frac{1}{\sqrt{\dot{x}_1(t)^2 + \ldots + \dot{x}_n(t)^2}} \begin{pmatrix} \dot{x}_1(t) \\ \vdots \\ \dot{x}_n(t) \end{pmatrix}$$

und daher lauten die E-L-Gleichungen

$$\tfrac{d}{dt}\frac{\partial L}{\partial \dot{x}} = \tfrac{d}{dt}\left(\frac{1}{\sqrt{\dot{x}_1(t)^2 + \ldots + \dot{x}_n(t)^2}} \begin{pmatrix} \dot{x}_1(t) \\ \vdots \\ \dot{x}_n(t) \end{pmatrix} \right) = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

$$\boxed{\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = \frac{d}{dt}\left(\frac{1}{\sqrt{\dot{x}_1(t)^2 + \ldots + \dot{x}_n(t)^2}}\begin{pmatrix}\dot{x}_1(t)\\ \vdots \\ \dot{x}_n(t)\end{pmatrix}\right) = \begin{pmatrix}0\\ \vdots \\ 0\end{pmatrix}.}$$

Wenn eine Ableitung gleich 0 ist, so ist die Funktion eine Konstante, also hat die Lösung die Eigenschaft

$$\frac{1}{\sqrt{\dot{x}_1(t)^2 + \ldots + \dot{x}_n(t)^2}}\begin{pmatrix}\dot{x}_1(t)\\ \vdots \\ \dot{x}_n(t)\end{pmatrix} = \begin{pmatrix}v_1\\ \vdots \\ v_n\end{pmatrix}$$

für einen (konstanten) Vektor $v = (v_1, \ldots, v_n)^T$.
Man sieht übrigens sofort, dass $|v| = 1$, sonst gäbe es gar keine Lösung.
Die Lösungen sind also die Kurven, die in jedem Punkt den Vektor v tangieren. Damit sind sie (beliebig umparameterisierte) Geraden.

Extremale der Wirkungsfunktionals

$$\boxed{\textbf{Wirkungsfunktional.}\ W(x) := \int_{t_0}^{t_1} |\dot{x}(t)|^2 dt = \int_{t_0}^{t_1} (\dot{x}_1(t)^2 + \ldots + \dot{x}_n(t)^2) dt}$$

Das ist ein Lagrange-Funktional mit
$L(x, \dot{x}, t) = |\dot{x}|^2 = \dot{x}_1(t)^2 + \ldots + \dot{x}_n(t)^2$. Wir haben

$$\frac{\partial L}{\partial x} \equiv \vec{0}, \quad \frac{\partial L}{\partial \dot{x}} = \begin{pmatrix} 2\dot{x}_1(t) \\ \vdots \\ 2\dot{x}_n(t) \end{pmatrix},$$

und daher lauten die E-L-Gleichungen

$$\tfrac{d}{dt} \frac{\partial L}{\partial \dot{x}} = \tfrac{d}{dt} \left(\begin{pmatrix} 2\dot{x}_1(t) \\ \vdots \\ 2\dot{x}_n(t) \end{pmatrix} \right) = \begin{pmatrix} 2\ddot{x}_1(t) \\ \vdots \\ 2\ddot{x}_n(t) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Die Lösungen sind also affin parametrisierte Geradenstücke

$$x(t) = x_0 + tv.$$

Wesentliche Unterschiede zwischen Extremalen von Längen- und Wirkungsfunktional

Wir haben gesehen, dass die Bahnen von Lösungen gleich sind und dass es Geradenstücke sind. Die Extremalen sind aber verschieden: Für das Wirkungsfunktional sind sie affin parametrisierte Geraden, für das Längenfunktional sind sie beliebig (also nicht nur affin) parametrisierte Geraden. Wir sehen also, dass die Lösungsmenge für das Längenfunktional viel größer ist als für das Wirkungsfunktional.

Auf dem Level von Differentialgleichungen hat das damit zu tun, dass die E-L-Differentialgleichungen für das Wirkungsfunktional folgende Eigenschaft haben: Alle höchsten Ableitungen der Unbekannten $x_1(t), ..., x_n(t)$ (im unserem Fall die zweiten Ableitungen $\ddot{x}_1(t), ..., \ddot{x}_n$) sind Funktion niedrigerer Ableitungen (einschließlich der Unbekannten selbst sowie der Koordinaten $x_1, ..., x_n$).

Aus der Theorie der GDG wissen wir, dass Differentialgleichungen mit dieser Eigenschaft eindeutig durch die Anfangsbedingungen bestimmt sind, in unserem Fall durch Anfangspunkt und Anfangsvektor.

Die E-L-Gleichungen für das Längenfunktional haben diese Eigenschaft nicht (versuchen sie es rechnerisch nachzuweisen, z.B. in kleineren Dimensionen)

Die Natur sucht optimale Wege.

Mathematische Umformulierung: Mehrere (oder sogar fast alle) physikalischen Prozesse verlaufen entlang von Extremalen eines Lagrange-Funktionals.

Bemerkung. Die Formulierung ist absichtlich vage – man kann diese experimentale Beobachtung nicht präziser formulieren. Um kompliziertere physikalische Prozesse zu beschreiben, muss man auch die Definition des Lagrange-Funktionals erweitern: Z.B. muss man eine mehrdimensionale "Zeit" erlauben, nicht-glatte Lagrange-Funktionale gestatten und so weiter.

Bemerkung. Außerdem ist es nicht immer trivial, zu einem physikalischen Prozess ein geeignetes Lagrange-Funktional zu finden.

In Rahmen von klassischen Mechanik kommen die oben genannten Schwierigkeiten nicht vor: Es gibt bestimmte Regeln zur Konstruktion des Lagrange-Funktionals.

Beschreibung des 1. und 2. Newtonschen Gesetzes im Rahmen der Lagrange-Mechanik

Lex prima. Ein Körper verharrt im Zustand der Ruhe oder der gleichförmigen Translation, sofern er nicht durch einwirkende Kräfte zur Änderung seines Zustands gezwungen wird.

Lex secunda. Die Änderung der Bewegung ist der Einwirkung der bewegenden Kraft proportional und geschieht in Richtung derjenigen geraden Linie, in welcher jene Kraft wirkt.

Fakt. Wegen des Energieerhaltungsgesetzes (dies ist eine weitere experimentelle Beobachtung) ist das Kraftfeld immer ein (negatives) Gradientenvektorfeld für eine Funktion. In der Physik nennt man diese Funktion die potentielle Energie:

$$F = -\operatorname{grad}(U) = -\left(\frac{\partial U}{\partial x_1}, \dots, \frac{\partial U}{\partial x_n}\right)^T.$$

Bewegung eines Teilchens im Kraftfeld.

Die Bewegung freier Teilchen im Kraftfeld der potentiellen Energie U kann man mit Hilfe der Lagrange-Funktion

$$L = K - U = \frac{1}{2}m|\dot{x}|^2 - U(x) = \frac{1}{2}m(\dot{x}_1(t)^2 + \dot{x}_2(t)^2 + \dot{x}_3(t)^2) - U(x).$$

beschreiben.

Also hängt die Funktion $L(x, y, t)$ nicht von t ab und lautet

$$L(x, y, t) = \frac{1}{2}m(y_1(t)^2 + y_2(t)^2 + y_3(t)^2) - U(x_1, x_2, x_3).$$

Wir schreiben jetzt die E-L-Gleichungen für dieses Lagrangefunktional auf. Offensichtlich stimmen sie mit den Newton-Gleichungen aus der **lex secunda** überein. Wir haben

$$\frac{\partial L}{\partial x} = -\begin{pmatrix} \frac{\partial U}{\partial x_1} \\ \frac{\partial U}{\partial x_2} \\ \frac{\partial U}{\partial x_3} \end{pmatrix}, \quad \frac{\partial L}{\partial \dot{x}} = \begin{pmatrix} m\dot{x}_1(t) \\ m\dot{x}_2(t) \\ m\dot{x}_3(t) \end{pmatrix},$$

also lauten die E-L-Gleichungen:

$$m\begin{pmatrix} \ddot{x}_1(t) \\ \ddot{x}_2(t) \\ \ddot{x}_3(t) \end{pmatrix} = -\begin{pmatrix} \frac{\partial U}{\partial x_1} \\ \frac{\partial U}{\partial x_2} \\ \frac{\partial U}{\partial x_3} \end{pmatrix}.$$

Die E-L-Gleichungen für die Bewegung eines freien Teilchens im Kraftfeld der potentiellen Energie U lauten:

$$m \begin{pmatrix} \ddot{x}_1(t) \\ \ddot{x}_2(t) \\ \ddot{x}_3(t) \end{pmatrix} = - \begin{pmatrix} \frac{\partial U}{\partial x_1} \\ \frac{\partial U}{\partial x_2} \\ \frac{\partial U}{\partial x_3} \end{pmatrix}$$

Da nach unserer Definition der potentiellen Energie das (negative) Gradientenfeld

$$- \begin{pmatrix} \frac{\partial U}{\partial x_1} \\ \frac{\partial U}{\partial x_2} \\ \frac{\partial U}{\partial x_3} \end{pmatrix}$$

mit der Kraft übereinstimmt, muss die Beschleunigung, also $\ddot{x}$, zur Kraft proportional sein, wie in der **lex secunda.**

Lex prima

Wir betrachten jetzt den Spezialfall $U \equiv \mathrm{const}$, d.h. es wirken keine Kräfte. Dieser Fall ist (bis auf den Koeffizienten $\frac{m}{2}$) identisch mit dem oben betrachteten Fall des Wirkungsfunktionals. So sind die Extremalen etwa die affin parametrisierten Geraden. Folglich:

Lex prima. Ein Körper verharrt im Zustand der Ruhe oder der gleichförmigen Translation, sofern er nicht durch einwirkende Kräfte zur Änderung seines Zustands gezwungen wird.

Mehrere Teilchen

Wir werden mit 3 Teilchen arbeiten und in Dimension 2. Das ist aber nicht wesentlich. Sie werden sehen, dass man dieselbe Methode für beliebig viele Teilchen und in beliebiger Dimension anwenden kann.
Die Bewegung von Teilchen (der Massen m_1, m_2, m_3) werden wir durch die vektorwertigen Funktionen

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix},\ y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix},\ z(t) = \begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix}$$

beschreiben, also ist der Vektor, der die Lage aller Teilchen beschreibt, 6-dimensional:

$$(x_1(t), x_2(t), y_1(t), y_2(t), z_1(t), z_2(t))^T.$$

Wir nehmen an, dass die Teichen über die potentielle Energie $U(x, y, z)$ irgendwie miteinander wechselwirken. Bspw. wäre möglich:

$$U(x, y, z) = \frac{m_1 m_2}{|x-y|} + \frac{m_1 m_3}{|x-z|} + \frac{m_1 m_3}{|y-z|}.$$

Wir können dann die Veränderung des Systems mit Hilfe des Lagrangefunktionals von $L = \frac{1}{2} m_1 |\dot{x}|^2 + \frac{1}{2} m_2 |\dot{y}|^2 + \frac{1}{2} m_3 |\dot{z}|^2 - U(x, y, z)$ beschreiben.

Auch hier werden wir sehen (wenigstens für den Spezialfall $U(x, y, z) = U_1(y, z) + U_2(x, z) + U_3(x, y)$), dass die E-L-Gleichung äquivalent zur Newton-Gleichung ist.

Bemerkung. Der Spezialfall $U(x, y, z) = U_1(y, z) + U_2(x, z) + U_3(x, y)$ entspricht der Situation, dass die Wechselwirkung zweier Teilchen nicht von einem dritten Teilchen abhängt.

Wir können dann die Veränderung des Systems mit Hilfe des Lagrangefunktionals von $L = \frac{1}{2} m_1 |\dot{x}|^2 + \frac{1}{2} m_2 |\dot{y}|^2 + \frac{1}{2} m_3 |\dot{z}|^2 - U(x, y, z)$ beschreiben.

Für diese Lagrangefunktion sieht die E-L-Gleichung wie folgt aus: Die Dimension ist $n = 3 \times 2 = 6$ und es ist

$$\frac{d}{dt} \begin{pmatrix} \dot{m}_1 x_1(t) \\ m_1 \dot{x}_2(t) \\ m_2 \dot{y}_1(t) \\ m_2 \dot{y}_2(t) \\ m_3 \dot{z}_1(t) \\ m_3 \dot{z}_2(t) \end{pmatrix} = \begin{pmatrix} m_1 \ddot{x}_1(t) \\ m_1 \ddot{x}_2(t) \\ m_2 \ddot{y}_1(t) \\ m_2 \ddot{y}_2(t) \\ m_3 \ddot{z}_1(t) \\ m_3 \ddot{z}_2(t) \end{pmatrix} = - \begin{pmatrix} \frac{\partial U}{\partial x_1} \\ \frac{\partial U}{\partial x_2} \\ \frac{\partial U}{\partial y_1} \\ \frac{\partial U}{\partial y_2} \\ \frac{\partial U}{\partial z_1} \\ \frac{\partial U}{\partial z_2} \end{pmatrix} .$$

Wenn z.B. $U(x, y, z) = U_1(y, z) + U_2(x, z) + U_3(x, y)$ gilt, sehen die ersten 2 Gleichungen wie folgt aus:

$$\begin{pmatrix} m_1 \ddot{x}_1 \\ m_1 \ddot{x}_2 \end{pmatrix} = - \begin{pmatrix} \frac{\partial U}{\partial x_1} \\ \frac{\partial U}{\partial x_2} \end{pmatrix} = - \begin{pmatrix} \frac{\partial U_2}{\partial x_1} + \frac{\partial U_3}{\partial x_1} \\ \frac{\partial U_2}{\partial x_2} + \frac{\partial U_3}{\partial x_2} \end{pmatrix},$$

und der Vektor $- \begin{pmatrix} \frac{\partial U_2}{\partial x_1} + \frac{\partial U_3}{\partial x_1} \\ \frac{\partial U_2}{\partial x_2} + \frac{\partial U_3}{\partial x_2} \end{pmatrix}$ ist der Vektor der Kräfte, welche die Teichen Z und Y auf das Teilchen X ausüben.

Allgemeine Regel zur Auswahl der Lagrange-Funktion bei mechanischen Systemen

- $L = K - U, \quad$ wobei K die kinetische und U die potenzielle Energie ist
  - Dabei sind sowohl die kinetische, als auch die potentielle Energie additiv: Wenn ein Teilsystem potentielle Energie U_1 und kinetische Energie K_1 hat und ein anderes disjunktes Teilsystem die potentielle bzw. kinetische Energie U_2 bzw. K_2, dann hat die Vereinigung der Systeme die Energien $U = U_1 + U_2$ bzw. $K = K_1 + K_2$.
  - In euklidischen Koordinaten hat das Teilchen der Masse m die kinetische Energie $K = \frac{m}{2}|\dot{x}|^2$.

Beispiel: Pendel

Ein Massenpunkt bewege sich reibungslos auf der Kurve $y = -\sqrt{1-x^2}$ im Feld der Schwerkraft.
In diesem Fall ist die Dimension 1, da die x-Koordinate des Punktes die Lage des Punktes eindeutig bestimmt. Wenn die x-Koordinate sich gemäß einer Funktion $x(t)$ ändert, bewegt sich das Teilchen in $\mathbb{R}^2$ auf der Kurve

$$\begin{pmatrix} x(t) \\ -\sqrt{1-x^2} \end{pmatrix}$$

und der Geschwindigkeitsvektor in 2-dimensionalen kartesischen Koordinaten ist

$$\frac{d}{dt}\begin{pmatrix} x(t) \\ -\sqrt{1-x^2} \end{pmatrix} = \begin{pmatrix} \dot{x} \\ \frac{\dot{x}x}{\sqrt{1-x^2}} \end{pmatrix} \text{, sodass } K = \frac{m}{2}\dot{x}^2\left(1+\frac{x^2}{1-x^2}\right) = \frac{m}{2}\frac{\dot{x}^2}{1-x^2}.$$

Die potentielle Energie U ist offensichtlich $-mg\sqrt{1-x^2}$, also ist

$$L(x,\dot{x},t) = \frac{m}{2}\frac{\dot{x}^2}{1-x^2} + mg\sqrt{1-x^2}.$$

$$\boxed{L(x,\dot{x},t) = \frac{m}{2}\frac{\dot{x}^2}{1-x^2} + mg\sqrt{1-x^2}}$$

Dann lauten die E-L-Gleichungen (man dividiert noch durch m):

$$\frac{\ddot{x}}{1-x^2} + {\frac{4\dot{x}^2 x}{1-x^2}}^2 = {\frac{4\dot{x}^2 x}{1-x^2}}^2 - g\frac{x}{\sqrt{1-x^2}}.$$

Wir erhalten schließlich die Differentialgleichung für die Beschreibung der Bewegung:

$$\ddot{x} = -gx\sqrt{1-x^2}.$$

(An der Tafel werde ich geometrisch erklären, dass man diese Gleichung einfach mittels der Newton-Gesetzte bekommen kann).

Beispiel dafür, dass die GDG für die Bewegung einfacher aus der E-L-Gleichung zu bekommen ist als aus den Newton-Gesetzen

Ein Massenpunkt bewege sich im Feld der Schwerkraft reibungslos auf der Parabel $y = x^2$.
Die Dimension ist in diesem Fall 1, da die x-Koordinate des Punktes die Lage des Punktes eindeutig beschreibt. Wenn die x-Koordinate sich gemäß einer Funktion $x(t)$ ändert, bewegt sich das Teilchen in $\mathbb{R}^2$ auf der Kurve

$$\begin{pmatrix} x(t) \\ x(t)^2 \end{pmatrix}$$

und der Geschwindigkeitsvektor lautet in 2-dimensionalen kartesischen Koordinaten

$$\tfrac{d}{dt}\begin{pmatrix} x(t) \\ x(t)^2 \end{pmatrix} = \begin{pmatrix} \dot{x} \\ 2\dot{x}x \end{pmatrix}.$$ Es ist also $K = \frac{m}{2}\dot{x}^2(1+4x^2)$.

Die potentielle Energie U ist offensichtlich $mgx(t)^2$. Also ist

$$L(x, \dot{x}, t) = \tfrac{m}{2}\dot{x}^2(1+4x^2) - mgx^2.$$

$$\boxed{L(x, \dot{x}, t) = \frac{m}{2}\dot{x}^2(1 + 4x^2) - mgx^2}$$

Die E-L-Gleichungen lauten daher

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = \frac{d}{dt}(m\dot{x}(1+4x^2)) = m\ddot{x}(1+4x^2)+4m\dot{x}^2x = \frac{\partial L}{\partial x} = 4m\dot{x}^2x-2mgx.$$

Nach algebraischen Manipulationen bekommen wir:

$$\ddot{x} = -\frac{2gx}{1+4x^2},$$

was wieder den Newton-Gesetzen entspricht.Jedoch ist es in diesem Fall eine nicht-triviale Aufgabe, die Newtonschen Gesetze zu verwenden.

Die Lagrange-Gleichungen hängen nicht vom Koordinatensystem ab.

Die Newton-Gesetzte sind dagegen nur in "inertialen" Koordinatensystemen anwendbar!

Def. Sei $U_1, U_2 \in \mathbb{R}^n$. Eine Abbildung $\phi : U_1 \to U_2$ heißt ein (C^∞-)Diffeomorphismus, wenn sie bijektiv ist, C^∞-glatt, und wenn die Umkehrabbildung $f^{-1} : U_2 \to U_1$ ebenfalls differenzierbar ist (und dann automatisch C^∞-glatt nach dem Satz über die Umkehrfunktion).

Bsp. Polarkoodinaten kann man als Diffeomorphimus

$$\phi(r, \theta) = (r\cos(\theta), r\sin(\theta)).$$

auffassen. Dieser Diffeomorphismus ist auf $U_1 := \{(r, \phi) \in \mathbb{R}^2 \mid r > 0, \phi \in (-\pi, \pi)\}$ definiert. Sein Wertebereich ist $U_2 := \mathbb{R}^2 \setminus \{(x, y) \in \mid x \leq 0, y = 0\}$.

Def. Sei $L : \mathbb{R}^n \times \mathbb{R}^n \times R \to \mathbb{R}$ eine Lagrange-Funktion und ϕ ein Diffeomorphismus. Definiere den Pull-back der Lagrange-Funktion als Funktion $\overline{L}$, gegeben durch

$$\overline{L}(x, y, t) := L(\phi(x), D_x(y), t).$$

Dabei ist das Differential D_x die lineare Abbildung, die $\phi(x + \xi) - \phi(x)$ bis zur 2. Ordnung (in ξ) approximiert:

$$\phi(x + \xi) = \phi(x) + D_x(\xi) + O(|\xi|^2).$$

Wiederholung. Aus der Analysis II folgt die Existenz dieser linearen Abbildung D_x. Sie ist eine Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^n$, deren Matrix (die sogenannte Jacobi-Matrix) die Einträge $\left[\frac{\partial \phi_i}{\partial x_j}(x)\right]$ hat.

Bsp. Für $\phi(r, \theta) = (r\cos(\theta), r\sin(\theta))$ ist

$$D_{(r,\phi)}(\xi) = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \cos(\theta)\xi_1 - r\sin(\theta)\xi_2 \\ \sin(\theta)\xi_1 + r\cos(\theta)\xi_2 \end{pmatrix}.$$

Bsp. Sei $L(x, y, t) = \frac{m}{2}|y|^2$ und $\phi(r, \theta) = (r\cos(\theta), r\sin(\theta))$. Dann ist

$$\overline{L}((r, \theta), \xi_1, \xi_2, t) = \frac{m}{2} \left| \begin{pmatrix} \cos(\theta)\xi_1 - r\sin(\theta)\xi_2 \\ \sin(\theta)\xi_1 + r\cos(\theta)\xi_2 \end{pmatrix} \right|^2 = \frac{m}{2}\left(\xi_1^2 + r^2\xi_2^2\right).$$

Differential als lineare Abbildung (Analysis II)

Die partiellen Ableitungen (z.B. in Dimension 2) kann man als Komponenten der Matrix $\begin{pmatrix} \frac{\partial \phi_1}{\partial x} & \frac{\partial \phi_1}{\partial y} \\ \frac{\partial \phi_2}{\partial x} & \frac{\partial \phi_2}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial y} \end{pmatrix}$ auffassen. Diese Matrix definiert (für jedes $p = (x, y) \in U$) eine lineare Abbildung

$$D\phi_p : \mathbb{R}^2 \to \mathbb{R}^2,\ D\phi_p(\xi) = \begin{pmatrix} \frac{\partial \phi_1}{\partial x} & \frac{\partial \phi_1}{\partial y} \\ \frac{\partial \phi_2}{\partial x} & \frac{\partial \phi_2}{\partial y} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \xi_1 \frac{\partial \phi_1}{\partial x} + \xi_2 \frac{\partial \phi_1}{\partial y} \\ \xi_1 \frac{\partial \phi_2}{\partial x} + \xi_2 \frac{\partial \phi_2}{\partial y} \end{pmatrix}.$$

Diese Abbildung heißt Differential von f im Punkt p und ist diejenige lineare Abbildung, welche die Funktion $\phi(x) - \phi(p)$ in der Umgebung des Punktes p am besten approximiert. Das bedeutet:

$$\phi(x) = \phi(p) + D_p\phi(x - p) + o(x - p).$$

Insbesondere gilt: Für jedes $p \in U$ und jeden Vektor $X \in \mathbb{R}^2$ ist die Richtungsableitung im Punkt p in Richtung X der Vektor

$$X\phi(p) := \tfrac{d}{dt}\phi(p + tX)_{|t=0} = d_p\phi(X).$$

Kettenregel für Differentiale

Selbstverständlich definiert man das Differential u.s.w. nicht nur für die Abbildungen von $\mathbb{R}^2$ nach $\mathbb{R}^2$, sondern auch für die Abbildungen f von $U^n \subseteq \mathbb{R}^n$ nach $\mathbb{R}^m$. In diesem Fall ist die Jacobi-Matrix eine $m \times n$ Matrix $\left(\frac{\partial f_i}{\partial x_j}\right)$ $(i = 1, ..., m; j = 1, ..., n)$ und das Differential df_p eine lineare Abbildung von $\mathbb{R}^n \to \mathbb{R}^m$.

Die Kettenregel aus Analysis II besagt, dass das Differential (im Punkt p) der Verkettung der Abbildungen h und g, $f = h \circ g$, mit der Verkettung der Differenzial-Abbildungen übereinstimmt: $D_p f = D_{g(p)} h \circ D_p g$. Letzteres wird in Koordinaten durch das Matrizenprodukt der Jacobi-Matrix von h (im Punkt $g(p)$) und der Jacobi-Matrix von g (im Punkt p) dargestellt.

Satz 3. Sei $x(t)$ eine Extremalkurve des Funktionals $\mathcal{L}$ und ϕ ein Diffeomorphismus. Dann ist $\phi \circ x(t)$ ein Extremal des Funktionals $\bar{\mathcal{L}}$ für den Pull-back $\bar{L}$ der Lagrange-Funktion L.

Bemerkung. Der Beweis im Skript von Stilianos Louca ist fehlerhaft.
Beweis. Nach Definition des Pull-back und des Differentials $D_x\phi$ ist

$$\bar{\mathcal{L}}(\phi \circ x) = \int_{t_0}^{t_1} \bar{L}(\phi(x(t)), \phi(\dot{x(t)}), t)dt$$

$$= \int_{t_0}^{t_1} L(x(t), \dot{x}(t), t)dt.$$

Für jede eigentliche Variation $x + h$ von x ist $\phi(x + h)$ auch eine eigentliche Variation von $\phi(x)$. Die entsprechende Funktion $\bar{h}$ ist dann

$$\bar{h} = \phi(x + h) - \phi(x).$$

Glattheit von $\bar{h}$ ist offensichtlich und es ist
$\bar{h}(t_0) = \phi(x(t_0) - h(t_0)) - \phi(x(t_0)) = \vec{0}$ ($\bar{h}(t_1) = \vec{0}$ entsprechend), sodass $\bar{h}$ eine eigentliche Variation ist.

Weil das Intervall $[t_0, t_1]$ kompakt ist, gibt es eine Konstante C, sodass der ε-kleinen Störung h eine $C\varepsilon$-kleine Störung von $\bar{h}$ entspricht.
Wenn das Differential D auf eigentlichen Variationen verschwindet, ist $\mathcal{L}(x + h) = \mathcal{L}(x)$ bis auf Terme, die quadratisch klein bzgl. $\|h\|$ sind. Dann ist aber $\bar{\mathcal{L}}(\phi(x + h)) = \mathcal{L}(\phi(x))$ bis auf Terme, die quadratisch klein bzgl. $\|\bar{h}\|$ sind, also ist auch $\phi(x)$ ein Extremal. $\square$